

Social Variation in the English of the Southampton Area

Kate Elizabeth Wallace

Submitted in accordance with the requirements for the degree of PhD
The University of Leeds School of English
July 2007

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Acknowledgements

I wish to acknowledge the generous grant awarded to me by the AHRC which enabled me to undertake this research. I would like to thank Dr Clive Upton, my supervisor, for his advice and support, not only during my PhD studies but throughout my time at the University of Leeds. I wish to acknowledge Esther Asprey of the University of Leeds, and Lourdes Burbano Elizondo of the University of Sheffield with whom I worked on revising the Survey of Regional English methodology. I would also like to thank Dr Paul D. Baxter of the Department of Statistics, University of Leeds, for his advice on the statistics in my thesis, and Professor Allan Fordy of the Department of Applied Mathematics, University of Leeds, for introducing me to Dr Baxter.

My thanks go to all of those in the Southampton area who introduced me to my informants. In particular, I would like to acknowledge the help of Jan Church of Taunton's Sixth Form College, Southampton, who allowed me to interview her students. Special thanks to all of my informants, who made this study possible.

This thesis is dedicated to my family, Janet, Helen, Naomi, Thomas and Doris, my fiancé, Dom, and is in memory of my dad, Brian. Without you, I could not have done this.

Abstract

This study provides an introduction to social variation in the English of the Southampton area. As the central Southern coast of Great Britain, in which Southampton is situated, has been largely neglected by dialect research, very little is known about the linguistic situation in the Southampton area. The present study seeks to rectify this situation.

Lexical, grammatical and phonological data have been collected using an adapted version of the Survey of Regional English methodology. This study describes the Survey's means of data elicitation, and how it has been revised for use in the Southampton area. The present study also employs an identity questionnaire, a grammar questionnaire, and a word list to collect linguistic and attitudinal data.

Linguistic data from sixty informants are analysed according to the speaker variables of sex, age, social class and identity, all of which have been shown in previous social dialectological research to influence language use. These linguistic data are also considered in the light of qualitative attitudinal data provided by informants. The matter of identity is of particular interest given Southampton's central location on the southern coastline, and use of a range of linguistic features is examined according to informants' feelings of identity with both the Southampton area, and the South East and South West of England.

Table of Contents

List of Tables and Figures	i
Abbreviations	vi
Chapter 1	1
1.0 Introduction	1
1.1 The City of Southampton	1
1.2 Defining the Sampling Universe	2
1.2.1 Links between Southampton and Eastleigh Borough	2
1.2.1.1 Geographical Links between Southampton and Eastleigh Borough	3
1.2.1.2 Economic Links between Southampton and Eastleigh Borough	3
1.2.1.3 Attitudinal Links between Southampton and Eastleigh Borough	4
1.3 Social Class in the Southampton Area	4
1.3.1 Indices of Deprivation 2004	5
1.4 Linguistic Perspectives on the Location of the Southampton Area	6
1.5 Identity in the Southampton Area	8
1.6 Extralinguistic Variables	9
1.7 Linguistic Variables	10
1.7.1 Dialect Levelling	10
1.7.2 Selection of Variables for Data Collection	12
1.7.3 Phonological Variables	13
1.7.3.1 Consonants	13
1.7.3.2 Vowels	18
1.7.4 Grammatical Variables	20

1.7.4.1	Verb Forms	21
1.7.4.2	Negation	26
1.7.4.3	Adverbs	28
1.7.4.4	Nominal Constructions	30
1.7.5	Lexical Variables	34
1.7.5.1	Lexis and the Issue of Dialect Versus Slang	34
Chapter 2		37
2.0	Literature Review	37
2.1	The Speaker Variable of Sex	37
2.1.1	Linguistic Sex Differentiation	37
2.1.2	Feminist Criticism	42
2.2	The Speaker Variable of Age	46
2.2.1	Empirical Research on the Linguistic Effect of Speaker Age	47
2.2.1.1	Age-Grading	47
2.2.1.2	Change in Progress – Variant on the Increase	49
2.2.1.3	Change in Progress – Variant on the Decline	50
2.3	The Speaker Variable of Social Class	51
2.3.1	Empirical Approaches to the Speaker Variable of Social Class	52
2.3.1.1	Labov on New York City	52
2.3.1.2	Trudgill on Norwich	54
2.3.1.3	Macaulay on Glasgow	57
2.3.1.4	Milroy and Milroy on Belfast	59
2.3.1.5	F Eckert on Belten High School, Detroit	60

2.4 The Speaker Variable of Identity	62
2.4.1 Gal on Oberwart, Austria	62
2.4.2 Trudgill on British Pop Song Pronunciation	63
2.4.3 Dyer on Corby, Northamptonshire	64
Chapter 3	66
3.0 The Southampton Area Study Methodology	66
3.1 The Survey of Regional English	66
3.2 The Survey of Regional English Methodology	67
3.3 Refinement of the Original Survey of Regional English Methodology	69
3.3.1 Alterations Made to the Untitled Instructions Sheet	69
3.3.1.1 Repetition between the Instructions Sheet and the <i>Biographical Information</i> Sheet	69
3.3.1.2 An Addition to the Instructions Sheet – <i>What is SuRE?</i>	70
3.3.1.3 Changes to the Wording of the Instructions Sheet	70
3.3.1.4 Aesthetic Alterations to the Instructions Sheet	71
3.3.2 Alterations Made to the <i>Biographical Information</i> Sheet	72
3.3.2.1 <i>Title</i>	72
3.3.2.2 <i>Name</i>	72
3.3.2.3 <i>Sex</i>	72
3.3.2.4 <i>Age (dob)</i>	72
3.3.2.5 <i>Place of birth</i>	73
3.3.2.6 <i>Birthplace of mother and birthplace of father</i>	74
3.3.2.7 <i>Birthplace of grandmothers and birthplace of grandfathers</i>	75
3.3.2.8 <i>Ethnic group</i>	75

3.3.2.9	<i>Occupation (current/usual)</i>	76
3.3.2.10	<i>Assessment of social class</i>	77
3.3.2.11	<i>Housing</i>	78
3.3.2.12	<i>Education</i>	78
3.3.3	Alterations Made to the Sense Relation Networks	79
3.3.3.1	Renaming of the Sense Relation Network Sheets	79
3.3.3.2	Alterations to the Sense Relation Network Subdivisions	79
3.3.3.3	Removal of the <i>any others</i> Sections from the Sense Relation Network Sheets	80
3.3.3.4	Removal of Notion Words Believed to be Unproductive	81
3.3.3.5	Alterations to Existing Notion Words	81
3.3.3.6	Addition of New Notion Words	83
3.3.3.7	Alterations to the Presentation of the Sense Relation Network Sheets	83
3.3.4	<i>Confidentiality and Consent Form</i>	85
3.4	The Revised Core SuRE Methodology Sheets	86
3.5	Additions Made to the Core SuRE Methodology for the Southampton Area Study	92
3.5.1	The Southampton Area Identity Questionnaire	92
3.5.1.1	A Copy of the Southampton Area Identity Questionnaire	95
3.5.2	The Southampton Area Language Questionnaire	96
3.5.2.1	A Copy of the Southampton Area Language Questionnaire	98
3.5.3	The Southampton Area Word List	101
3.5.3.1	A Copy of the Southampton Area Word List	103

3.6 Speaker Selection	104
3.7 The Interview	106
3.8 Sample Size	108
3.9 Social Stratification of Southampton Area Informants	110
3.9.1 The National Statistics Socioeconomic Classification (NS-SEC) and its Use in the Southampton Area Study	111
3.9.2 The NS-SEC and Full-Time Students	115
3.9.3 Unclassified Informants	117
3.9.4 Analytic Class Sizes	118
3.10 Data Analysis	118
3.10.1 Data Analysis - Lexis	118
3.10.2 Data Analysis – Grammar	122
3.10.3 Data Analysis – Phonology	126
3.11 Analysis of Data by Multiple Extralinguistic Variables	127
3.12 Statistical Testing	128
3.13 Selection of Linguistic Variables for Examination	129
Chapter 4	133
4.0 Lexis	133
4.1 <i>Pregnant</i>	133
4.1.1 <i>Pregnant</i> and Speaker Sex	133
4.1.2 <i>Pregnant</i> and Speaker Social Class	137
4.2 <i>Toilet</i>	141
4.2.1 <i>Toilet</i> and Speaker Sex	141
4.2.2 <i>Toilet</i> and Speaker Age	144
4.2.3 <i>Toilet</i> and Speaker Social Class	148

4.3	<i>Grandmother and Grandfather</i>	150
4.3.1	<i>Grandmother and Speaker Sex</i>	150
4.3.2	<i>Grandfather and Speaker Sex</i>	153
4.3.3	<i>Grandmother and Speaker Social Class</i>	154
4.3.4	<i>Grandfather and Speaker Social Class</i>	156
4.3.5	<i>Grandmother and Grandfather – A Summary</i>	158
4.4	<i>Long seat in main room of house</i>	159
4.4.1	<i>Long seat in main room of house and Speaker Age</i>	159
4.4.2	<i>Long seat in main room of house and Speaker Social Class</i>	162
4.5	<i>Left-handed</i>	163
4.5.1	<i>Left-handed and Speaker Age</i>	163
4.5.2	<i>Left-handed and Speaker Social Class</i>	166
4.6	Lexis – A Summary	168
	Chapter 5	170
5.0	Grammar	170
5.1	Grammar Scores	170
5.1.1	Grammar Scores and Speaker Sex	171
5.1.2	Grammar Scores and Speaker Age	173
5.1.3	Grammar Scores and Speaker Social Class	175
5.2	<i>Ain't</i>	176
5.2.1	<i>Ain't and Speaker Age</i>	176
5.2.2	<i>Ain't and Speaker Social Class</i>	178
5.3	Prepositions	180
5.3.1	Prepositions and Speaker Sex	180

5.3.2	Prepositions and Speaker Age	182
5.3.3	Prepositions and Speaker Social Class	183
5.4	Nouns of Measurement with Zero Plurals	186
5.4.1	Nouns of Measurement with Zero Plurals and Speaker Age	187
5.4.2	Nouns of Measurement with Zero Plurals and Speaker Social Class	188
5.5	<i>There was</i> with a Plural Notional Subject	189
5.5.1	<i>There was</i> with a Plural Notional Subject and Speaker Sex	190
5.5.2	<i>There was</i> with a Plural Notional Subject and Speaker Social Class	191
5.6	Speaker Self-reporting Versus Actual Use	192
5.6.1	Present Tense Verb Morphology	193
5.6.2	Adverbs	204
5.7	Grammar – A Summary	211
	Chapter 6	214
6.0	Phonology	214
6.1	L Vocalization	215
6.1.1	L Vocalization and Speaker Sex	215
6.1.2	L Vocalization and Speaker Age	217
6.1.3	L Vocalization and Speaker Social Class	218
6.2	Rhoticity	220
6.2.1	Rhoticity and Speaker Sex	220
6.2.2	Rhoticity and Speaker Age	221
6.2.3	Rhoticity and Speaker Social Class	222
6.3	The BATH Vowel	223
6.3.1	The BATH Vowel and Speaker Sex	223

6.3.2	The BATH Vowel and Speaker Age	225
6.3.3	The BATH Vowel and Speaker Sex and Age	226
6.3.4	The BATH Vowel and Speaker Social Class	228
6.3.5	The BATH Vowel and Speaker Sex, Age and Social Class	229
6.4	The PRICE Vowel	231
6.4.1	The PRICE Vowel and Speaker Sex	232
6.4.2	The PRICE Vowel and Speaker Age	233
6.4.3	The PRICE Vowel and Speaker Sex and Age	236
6.4.4	The PRICE Vowel and Speaker Social Class	238
6.5	Phonology – A Summary	240
Chapter 7		242
7.0	The Speaker Variable of Identity in the Southampton Area	242
7.1	Findings of the Southampton Area Identity Questionnaire	242
7.1.1	Question One	243
7.1.2	Question Two	246
7.1.3	Question Three	248
7.1.4	Question Four	251
7.1.5	Southampton Area Identity Scores	253
7.2	Data Presentation Issues	255
7.3	Identity and Lexis	258
7.3.1	<i>Long seat in main room of house</i>	258
7.3.2	<i>Grandmother and Grandfather</i>	261
7.3.2.1	<i>Grandmother</i>	261
7.3.2.2	<i>Grandfather</i>	263
7.3.3	<i>Pregnant</i>	264

7.4 Identity and Grammar	266
7.4.1 <i>Ain't</i>	267
7.4.2 <i>There was</i> with a Plural Notional Subject	269
7.4.3 Nouns of Measurement with Zero Plurals	271
7.5 Identity and Phonology	272
7.5.1 L Vocalization	272
7.5.2 Rhoticity	273
7.5.3 The BATH Vowel	275
7.5.4 The PRICE Vowel	276
7.6 East-west Identity in the Southampton Area	278
7.6.1 East-west Identity and Phonology	281
7.6.1.1 L Vocalization	281
7.6.1.2 Rhoticity	283
7.6.1.3 The BATH Vowel	284
7.7 The Speaker Variable of Identity in the Southampton Area – A Summary	286
Chapter 8	288
8.0 Conclusion	288
8.1 A Critique of the Adapted SuRE Methodology	288
8.1.1 Lexical Data	288
8.1.2 Grammatical Data	291
8.2 The Speaker Variables of Sex, Age and Social Class	296
8.3 The Speaker Variable of Identity	297
8.4 Suggestions for Further Work	299
8.5 Closing Comments	299
Appendix 1 – The original core Survey of Regional English methodology	300

Appendix 2 - Variants recorded by Southampton area speakers on the SuRE Sense Relation Network sheets, listed in decreasing order of frequency	304
Appendix 3 – Standardised lexical variants and alternative spellings	314
Appendix 4 – Results of the Southampton area language questionnaire	319
Bibliography	322

List of Tables and Figures

Table 3.1 Speaker cells for the Southampton area survey	108
Figure 3.1 Operational Categories and Sub-categories of the National Statistics Socioeconomic Classification (NS-SEC)	112
Table 3.2 National Statistics Socioeconomic Classification Operational Categories and Analytic Classes of Southampton Area Informants	114- 115
Table 3.3 Allocation of Students to Analytic Classes Based on Occupation of Higher Ranking Parent	117
Table 3.4 Social Class Composition of Southampton Area Study	118
Figure 4.1 Most frequently reported variants for the notion word <i>pregnant</i> according to the speaker variable of sex	134
Figure 4.2 Most frequently reported variants for the notion word <i>pregnant</i> according to the speaker variable of social class	138
Figure 4.3 Most frequently reported variants for the notion word <i>toilet</i> according to the speaker variable of sex	141
Figure 4.4 Most frequently reported variants for the notion word <i>toilet</i> according to the speaker variable of age	145
Figure 4.5 Most frequently reported variants for the notion word <i>toilet</i> according to the speaker variable of social class	149
Figure 4.6 Most frequently reported variants for the notion word <i>grandmother</i> according to the speaker variable of sex	151
Figure 4.7 Most frequently reported variants for the notion word <i>grandfather</i> according to the speaker variable of sex	153
Figure 4.8 Most frequently reported variants for the notion word <i>grandmother</i> according to the speaker variable of social class	155
Figure 4.9 Most frequently reported variants for the notion word <i>grandfather</i> according to the speaker variable of social class	157
Figure 4.10 Most frequently reported variants for the notion phrase <i>long seat in main room of house</i> according to the speaker variable of age	160
Figure 4.11 Most frequently reported variants for the notion phrase <i>long seat in main room of house</i> according to the speaker variable of social class	162

Figure 4.12 Most frequently reported variants for the notion word <i>left-handed</i> according to the speaker variable of age	164
Figure 4.13 Most frequently reported variants for the notion word <i>left-handed</i> according to the speaker variable of social class	167
Figure 5.1 Mean grammar scores according to the speaker variable of sex	172
Figure 5.2 Individual grammar scores according to the speaker variable of	173
Figure 5.3 Mean grammar scores according to the speaker variable of age	174
Figure 5.4 Mean grammar scores according to the speaker variable of social class	175
Figure 5.5 Reported use of <i>ain't</i> when talking to a friend according to the speaker variable of age	177
Figure 5.6 Reported use of <i>ain't</i> when talking to a friend according to the speaker variable of social class	179
Figure 5.7 Reported use of nonstandard simple and complex prepositions when talking to a friend according to the speaker variable of sex	181
Figure 5.8 Reported use of nonstandard simple and complex prepositions when talking to a friend according to the speaker variable of age	182
Figure 5.9 Reported use of nonstandard simple and complex prepositions when talking to a friend according to the speaker variable of social class	184
Table 5.1 Reported use of <i>I'm going down my friend's house later</i> when talking to a friend according to the speaker variables of social class and age	186
Table 5.2 Reported use of <i>I'm going up my friend's house later</i> when talking to a friend according to the speaker variables of social class and sex	186
Figure 5.10 Reported use of nouns of measurement with zero plurals when talking to a friend according to the speaker variable of age	187
Figure 5.11 Reported use of nouns of measurement with zero plurals when talking to a friend according to the speaker variable of social class	188
Figure 5.12 Reported use of <i>there was</i> with a plural notional subject when talking to a friend according to the speaker variable of sex	190
Figure 5.13 Reported use of <i>there was</i> with a plural notional subject when talking to a friend according to the speaker variable of social class	192
Table 5.3 Reported use of standard and nonstandard present tense verb morphology	193

Table 5.4	Reported use of standard and nonstandard present tense verb morphology according to the speaker variables of age and sex	194
Table 5.5	Reported use of standard and nonstandard present tense verb morphology according to the speaker variable of social class	195
Table 5.6	Actual use of standard and nonstandard present tense verb morphology	197
Table 5.7	Speaker self-reporting of present tense verb morphology	197
Table 5.8	Under- and over-reporting of present tense verb morphology use according to the speaker variables of age and sex	199
Table 5.9	Under- and over-reporting of present tense verb morphology use according to the speaker variable of social class	201
Table 5.10	Reported use of <i>-ly</i> adverbs	204
Table 5.11	Reported use of standard and nonstandard adverbs according to the speaker variables of age and sex	204
Table 5.12	Reported use of standard and nonstandard adverbs according to the speaker variable of social class	205
Table 5.13	Actual use of standard and nonstandard <i>-ly</i> adverbs	206
Table 5.14	Speaker self-reporting of adverbial constructions	206
Figure 5.14	Under- and over-reporting of standard and nonstandard adverbial constructions according to the speaker variables of age and sex	208
Table 5.15	Under- and over-reporting of standard and nonstandard adverbial constructions according to the speaker variable of social class	210
Figure 6.1	L Vocalization according to the speaker variable of sex	215
Figure 6.2	L Vocalization according to the speaker variable of age	217
Figure 6.3	L Vocalization according to the speaker variable of social class	219
Figure 6.4	Rhoticity according to the speaker variable of sex	220
Figure 6.5	Rhoticity according to the speaker variable of age	221
Figure 6.6	Rhoticity according to the speaker variable of social class	223

Figure 6.7 Realisations of the BATH vowel according to the speaker variable of sex	224
Figure 6.8 Realisations of the BATH vowel according to the speaker variable of age	225
Figure 6.9 Realisations of the BATH vowel according to the speaker variables of sex and age	227
Figure 6.10 Realisations of the BATH vowel according to the speaker variable of social class	229
Table 6.1 Realisations of the BATH vowel according to the speaker variables of sex, age and social class	230
Figure 6.11 Most frequently occurring realisations of the PRICE vowel according to the speaker variable of sex	232
Figure 6.12 Most frequently occurring realisations of the PRICE vowel according to the speaker variable of age	234
Figure 6.13 Most frequently occurring realisations of the PRICE vowel according to the speaker variables of sex and age	237
Figure 6.14 Most frequently occurring realisations of the PRICE vowel according to the speaker variable of social class	238
Table 7.1 Responses to question one of the Southampton area identity questionnaire	243
Table 7.2 Responses to question two of the Southampton area identity questionnaire	247
Table 7.3 Responses to question three of the Southampton area identity questionnaire	249
Table 7.4 Responses to question four of the Southampton area identity questionnaire	251
Figure 7.1 Southampton area identity scores	254

Figure 7.2 Most frequently reported variants for the notion phrase <i>long seat in main room of house</i> according to level of speaker identity	259
Figure 7.3 Most frequently reported variants for the notion word <i>grandmother</i> according to level of speaker identity	262
Figure 7.4 Most frequently reported variants for the notion word <i>grandfather</i> according to level of speaker identity	263
Figure 7.5 Most frequently reported variants for the notion word <i>pregnant</i> according to level of speaker identity	265
Figure 7.6 Mean grammar scores according to level of speaker identity	267
Figure 7.7 Reported use of <i>ain't</i> when talking to a friend according to level of speaker identity	268
Figure 7.8 Reported use of <i>there was</i> with a plural subject when talking to a friend according to level of speaker identity	270
Figure 7.9 Reported use of nouns of measurement with zero plurals when talking to a friend according to level of speaker identity	271
Figure 7.10 L Vocalization according to level of speaker identity	273
Figure 7.11 Rhoticity according to level of speaker identity	274
Figure 7.12 Realisations of the BATH vowel according to level of speaker identity	275
Figure 7.13 Most frequently occurring realisations of the PRICE vowel according to level of speaker identity	277
Figure 7.14 L Vocalization according to speaker feelings of east-west	282
Figure 7.15 Rhoticity according to speaker feelings of east-west identity	283
Figure 7.16 Realisations of the BATH vowel according to speaker feelings of east-west identity	285

Abbreviations

<i>OED</i>	<i>Oxford English Dictionary</i>
RP	Received Pronunciation
SED	Survey of English Dialects
SRN	Sense Relation Network
SRNs	Sense Relation Networks
SuRE	Survey of Regional English

Chapter 1

1.0 Introduction

Southampton was chosen for investigation because no notable linguistic studies have been carried out in the area. Edwards et al. (1984: 33), in their survey of the grammar of English dialect, refer to the central southern coast of England as having been neglected by dialect research, and this deficiency has not been redressed. A further reason for selecting Southampton for investigation is that as a native of the area, I am well placed to conduct linguistic research there. Trudgill (1983: 41) states that ‘wherever possible fieldworkers should be natives of the area’. Edwards et al. support this notion, arguing that:

[t]he native speaker intuitions of the ‘insider’ [...] often offer valuable insights and short cuts, and it is certainly easier for a member of a given community to make contact with and record suitable informants without having to worry so much about the sociolinguist’s perennial problem – the observer’s paradox.

(Edwards et al. 1984: 33)

The present study provides an introduction to the dialect of a previously unresearched area.

1.1 The City of Southampton

The city of Southampton in the county of Hampshire is situated approximately 100 kilometres south-west of London, on the central southern coast of Great Britain. It is ‘one of the major cities in the south outside London’ (Southampton City Council 2006), with a population at the time of the 2001 census of 217,445, living in an area covering 4,984 hectares (Office for National Statistics 2007).

Southampton’s industries are varied, though influenced by its waterside location. There is a long history of shipbuilding in the area, which continues to the

present day. The port, one of the biggest in the country, is 'the UK's principal cruise port', and the city's container terminal is the second largest in the United Kingdom (Associated British Ports 2007). Southampton is also home to major national and international businesses, such as British American Tobacco, Vosper Thornycroft, and Ford.

1.2 Defining the Sampling Universe

For the purpose of this research, the city of Southampton and the neighbouring borough of Eastleigh have been amalgamated to form a sampling universe referred to as the Southampton area. Situated to the north of Southampton, Eastleigh comprises coastline, countryside, villages and towns. For administrative purposes, the borough is divided into five local areas, these being: Chandler's Ford and Hiltingbury; Eastleigh; Bishopstoke, Fair Oak and Horton Heath; Botley, Hedge End and West End; and Bursledon, Hamble and Hound (Eastleigh Borough Council 2007). At the time of the 2001 census, it had a total population of 116,169, living in an area covering 7,978 hectares (Office for National Statistics 2007). Once a thriving railway area, the borough's principal occupations now include retail and manufacturing (Office for National Statistics 2007).

1.2.1 Links between Southampton and Eastleigh Borough

A variety of links exist between the city of Southampton and Eastleigh Borough which have been key factors in the decision to amalgamate the two areas for the purpose of the present study. These connections are detailed in the following sections.

1.2.1.1 Geographical Links between Southampton and Eastleigh Borough

Strong links exist between Southampton and the borough of Eastleigh. Not only do they abut one another, but parts of Eastleigh Borough have been included as subdivisions of the Southampton urban area for census purposes (Office for National Statistics 2004). Up to and including 1971, towns or cities were defined for census purposes according to their administrative boundaries. There were, however, 'serious disadvantages' in the use of such administrative boundaries, which 'changed infrequently, and often did not reflect the expansion of urban areas' (Office for National Statistics 2004: 1). Since the 1981 census, urban areas have been defined according to 'the extent of urban development indicated on Ordnance Survey (OS) maps' (Office for National Statistics 2004: 2). The incorporation of parts of Eastleigh Borough into the Southampton urban area for the 2001 census highlights the changes to population and social composition in the two conurbations, caused by out-migration from Southampton, and the resultant links between the two locations.

1.2.1.2 Economic Links between Southampton and Eastleigh Borough

Eastleigh Borough Council (2006) states that '[a]round 12,600 people commute to Southampton [for work] from Eastleigh and about 11,700 travel the other way'. Along with other concerned parties, Southampton City Council and Eastleigh Borough Council have made common cause in seeking to improve the economic situation in the south Hampshire area (Partnership for Urban South Hampshire 2007). The Councils of Southampton and Eastleigh are clearly aware that the economic health of their areas is closely linked, and are keen to collaborate in order to ensure that local residents have the relevant skills to secure jobs, and thus decrease unemployment levels.

1.2.1.3 Attitudinal Links between Southampton and Eastleigh Borough

Attitudinal links between Southampton and Eastleigh Borough are also clear. When debates took place in the early 1990s as to how best to structure local government in Hampshire, both Eastleigh Borough Council and Southampton City Council accepted that a merger between the two areas, which would result in the creation of an Eastleigh/Southampton unitary council, was a ‘viable and acceptable option’ (Local Government Commission for England 1994: 19). Ultimately, Eastleigh retained its two-tier system of County and District councils, while Southampton became a unitary authority. However, the acceptance of a merger, in principle at least, suggests strong attitudinal ties between the two areas. The range and strength of links between the city of Southampton and the borough of Eastleigh, as well as the complementary social composition of the two areas, provide sound justification for grouping these neighbouring conurbations to form the Southampton area for the purpose of this research.

1.3 Social Class in the Southampton Area

My decision to amalgamate the city of Southampton and the borough of Eastleigh is also based on the social class of their inhabitants. It is important to consider class in planning the speaker sample, since to fail to take into account social class is to ignore what can be argued to be ‘the most important social factor underlying [linguistic] changes in progress’ (Foulkes and Docherty 1999: 15).

Discussing her Belfast survey, Milroy poses the question as to:

whether it is always reasonable to take the population of an urban area as a sampling universe, when in fact a high proportion of the higher-status people who work in that city actually reside in neighbouring towns.

(Milroy 1987b: 85)

Many middle-class families, Milroy (1987b: 85) goes on to state, had moved out of the Belfast area to neighbouring dormitory towns, a migratory pattern 'in line with changes in the demography of British cities generally for the last two decades or so'.

Southampton's population has followed a similar pattern to that of Belfast, with many families formerly resident in the city having moved to the neighbouring areas of Eastleigh, Winchester, and the New Forest (Patterson 1970: 210; Directorate of Strategy and Development 1993). It is because of this migratory pattern especially that Eastleigh Borough and Southampton City have been amalgamated for the Southampton area study.

1.3.1 Indices of Deprivation 2004

Many of Southampton's wards fare badly in the Indices of Deprivation 2004 (Office of the Deputy Prime Minister 2004a). These national rankings are constructed using seven weighted indices, several of which have been used in sociolinguistic studies to determine the social class of speakers (see, for example, Trudgill 1974 and Macaulay 1977). The indices are: income; employment; health and disability; education, skills and training; housing and services; living environment; and crime (Office of the Deputy Prime Minister 2004a). Data for the Indices of Deprivation 2004 are presented using Super Output Areas (SOAs), each containing approximately 1500 people (Office of the Deputy Prime Minister 2004a). Southampton has 146 SOAs, and the city's average SOA score is ranked 96th out of the 354 local authorities in England, a rank of 1 being the most deprived (Office of the Deputy Prime Minister 2004a). The most deprived of Southampton's SOAs is ranked at 1393 out of the 32,482 Super Output Areas in England, and the least deprived at 28,299 out of 32,482 (Office of the Deputy Prime Minister 2004b). 34.25% of SOAs in Southampton are ranked between 1 and 10,000. 44.52% between 10,001 and 20,000.

and 21.23% between 20,001 and 32,482 (Office of the Deputy Prime Minister 2004b). Though accounting for only around a fifth of the total number of SOAs in Southampton, those Southampton Super Output Areas ranked between 20,001 and 32,482 are nonetheless significant, indicating that Southampton is not a homogeneous working class area. Alone, however, these SOAs would not yield a comparable number of higher class speakers to measure the effects of social class on language. In order to access such a comparable number of speakers from higher social rankings, it is necessary to look beyond Southampton to the borough of Eastleigh, an area that fares much better in the Indices of Deprivation rankings. Eastleigh Borough is divided into 77 Super Output Areas, and its average SOA score is ranked 311th out of the 354 local authorities in England (Office of the Deputy Prime Minister 2004a). The most deprived SOA is ranked at 8514 out of 32,482, and the least deprived at 32,368 out of 32,482 (Office of the Deputy Prime Minister 2004b). 1.30% of Super Output Areas in Eastleigh are ranked between 1 and 10,000, 19.48% between 10,001 and 20,000 and 79.22% between 20,001 and 32,482 (Office of the Deputy Prime Minister 2004b).

1.4 Linguistic Perspectives on the Location of the Southampton Area

Linguistically, there exists a variety of opinions concerning the Southampton area's location in relation to the South East and the South West. Wakelin (1986: 1) acknowledges the difficulties of dividing the South of England into dialect groups. He is well aware that dialect areas as such do not exist. Rather, 'dialects fade into other regional varieties' (Wakelin 1986: 1-3). Wakelin employs administrative boundaries as a means by which to define the South West, and, by extension, the South East. He defines the South West for the purposes of dialect research as being

comprised of Cornwall, Devon, Somerset, South Avon, Wiltshire, and Dorset. The West of Hampshire is described as forming part of a transitional area. The East of Hampshire in which Southampton lies is then, by extension, part of the South East, or at least part of central southern England, from Wakelin's dialectal point of view.

Wells (1982: 335) describes the South as being comprised of three accent areas outside London: the Home Counties, East Anglia, and the West Country. Hampshire, the county in which the Southampton area is situated, is described as being part of Wessex, the transitional area of the West Country, which bridges the gap between the West Country and the Home Counties (Wells 1982: 335-36).

Hughes et al. (2005: 69), on the other hand, see the South of England as falling into five accent areas: the Western South West; the Eastern South West; the South East; the South Midlands; and East Anglia. They categorise Hampshire and Southampton as being part of the Eastern South West. They distinguish the two South Western areas from the other Southern areas on the grounds that they are rhotic and lack the distinction between /a/ and /ɑ:/, and the two South Western areas from one another on the grounds that Long Mid Diphthonging (the nineteenth century change in the FACE vowel from [e:] to [eɪ] and in the GOAT vowel from [o:] to [ou] (Wells 1982: 210)) is absent in the Western South West (Hughes et al. 2005: 69-70).

Viereck (1980: 28-30), using lexical and grammatical data collected by Lowman in the 1930s, plots a dividing line between the South East and the South West 'running from Portsmouth northwards through Buckinghamshire, then turning west in the direction of Birmingham'. According to this definition, the Southampton area is in the South West.

1.5 Identity in the Southampton Area

The issue of the city's location, discussed in 1.4. and the questions of linguistic identity that this raises, is of great interest. Though restricted, being a solely phonological study with a limited speaker sample, existing research on speaker self-evaluation and identity in the Southampton area nevertheless indicates sharply polarised opinion among residents as to whether Southampton is in South East or South West England and, by extension, whether Sotonians identify with the East or the West (Wallace 2000). That study highlights a mixture of phonological forms typical of the South East and of the South West in the speech of the area.

Le Page and Tabouret-Keller's Acts of Identity Theory states, subject to a number of riders, that:

individuals create the patterns for their linguistic behaviour so as to resemble those of the group or groups with which from time to time they wish to be identified.

(Le Page and Tabouret-Keller 1985: 18)

Such groups are often geographically based: I will examine the influence of the South East and South West on the phonology of the inhabitants of the central southern Southampton area in light of the Theory. The decision has been taken to focus on phonology as opposed to grammar or lexis as, in general, more is known about the phonology of British English, including the geographical spread of particular phonological features, than is known about its grammar or lexis. Johnson argues that:

[L]inguists have generally tended to regard content words as unworthy of study as part of language, preferring to study the phonological components of words or the behavior of function words and morphemes.

(Johnson 1996: 83)

Cheshire et al. (1989: 185) describe our knowledge of the morphology and syntax of British English dialects as 'lag[ging] far behind our knowledge of their phonology'.

Trudgill (1986: 54) states that ‘London-based innovation is making its way into Norwich and other East Anglian centres’, offering ‘attitudinal factors’ as one explanation for this. His work highlights the influence of London on more distant urban areas. My earlier findings (Wallace 2000) suggest it is reasonable to suppose that Southampton could be one such urban area influenced by London. If attitudinal factors are key to the spread of phonological features as Trudgill argues, it is also reasonable to suppose that influences other than that of London might be at work on Southampton speech.

In light of Trudgill’s (1986) work on the influence of London on more distant urban areas, London will be the location in the South East studied for the purpose of linguistic comparison with the Southampton area. Tollfree’s (1999) study provides detailed phonological information on the speech of some of the capital’s inhabitants. The South West is more problematic, however, since it has no clear centre. Phonological data are, however, available on the South West in general (Wells 1982; Hughes et al. 2005), and these data, along with Tollfree’s, will facilitate identification of those variables which characterise the speech of London and the South West for comparison with linguistic data collected in the Southampton area. Ultimately, correlations might be expected between claimed speaker identity with an area (the Southampton area, the South East, the South West) and the employment of linguistic variables typical of that area.

1.6 Extralinguistic Variables

The present study analyses lexical, grammatical and phonological data according to the speaker variables of sex, age, social class, and identity. All four of these extralinguistic variables, which are discussed at greater length in chapter 2, have

been found to influence language use (see, for example, Trudgill 1974 and Cheshire 1982 on the influence of sex, Labov 1972b and Macaulay 1977 on age, Labov 1966 and Trudgill 1974 on social class, and Trudgill 1983 and Le Page and Tabouret-Keller 1985 on identity). Chapters 4, 5 and 6 are comprised of analysis of lexical, grammatical and phonological data respectively, according to speaker sex, age and social class. Lexis, grammar and phonology are also analysed in chapter 7, which is concerned with the essential issue of identity in the Southampton area.

1.7 Linguistic Variables

1.7.1 Dialect Levelling

Dialect levelling, and, by extension, accent levelling, is the process:

whereby differences between regional varieties are reduced, features which make varieties distinctive disappear, and new features emerge and are adopted by speakers over a wide geographical area.

(Williams and Kerswill 1999: 149)

This phenomenon can be explained in part by an increase in geographical mobility, such as that seen in Great Britain over the past 40 years. Former city dwellers have in increasing numbers moved to neighbouring towns and suburbs, as has been the case in the Southampton area, the populations of smaller cities and towns have increased, and new towns such as Milton Keynes have been created (Williams and Kerswill 1999: 149). This geographical mobility, coupled with an increase in social mobility, are believed to lead to the decline of the 'close-knit social networks associated with traditional working-class communities and thought to be influential in maintaining local linguistic norms' (Williams and Kerswill 1999: 149). As such, 'levelling has been shown to occur in mobile populations where there is a high level of dialect contact' (Williams and Kerswill 1999: 150).

Many levelled phonological forms are argued by some to have their roots in London English (Williams and Kerswill 1999: 159). On accent levelling, Wells states:

By the end of the [20th] century [...] some new non-localizable but more democratic standard may have arisen from the ashes of RP [Received Pronunciation]: if so, it seems likely to be based on popular London English.

(Wells 1982: 118)

Llamas (1998: 97) cites T Glottalling, TH Fronting, and the use of the labiodental approximant [ʋ] as a replacement for the alveolar approximant [ɹ] as having been identified as ‘part of ‘popular London English’-based levelling changes’. Wells (1982: 301) argues that the working-class London accent is ‘today the most influential source of phonological innovation in England and perhaps in the whole English-speaking world’. It should be noted, however, that arguments for the widespread influence of London English are not firmly established. It is hoped that the Southampton area study might add to the debate, though, given that Southampton, like London, is in the South of England, the findings of the study cannot be taken as indicative of what is happening elsewhere in the country.

In their survey of dialect grammar, Cheshire et al. (1989: 187) describe ‘shared dialect features in the major urban centres of Britain’, acknowledging that these features could be indicative of a move towards a ‘levelled nonstandard dialect’. Cheshire (1982: 128-29) makes the point in her study of Reading adolescents that none of the nonstandard grammatical features under examination ‘is peculiar to Reading. Some features [...] occur in most, if not all, British and American varieties. Others are more localised’. Similarly, while the lexical, grammatical and phonological variables under examination in this study are features of the speech of the Southampton area, they are not specific to Southampton. Indeed, many of these

variants are found country-wide. This could be argued to be a result of dialect levelling. Equally, it could be the result of another factor, such as chance.

1.7.2 Selection of Variables for Data Collection

Milroy (1987b: 115) states that '[s]ince [linguistic] variables pattern differently in different places, their initial identification is by no means automatic and might not even be particularly easy'. She argues that 'those who have grown up as native speakers of a dialect [...] may have intuitions about its structure' (Milroy 1987b: 3), and that they are 'likely to be able to articulate and use intuitions about relevant variables' (Milroy 1987b: 115). Trudgill (1974: 80), citing 'native knowledge of the speech of the area' as grounds for the selection of phonological variables in his original Norwich survey, is a prime example of the native speaker Milroy describes. However, the employment of native intuition as a means by which to select linguistic variables is not without its dangers. Milroy (1987b: 116) argues that it is an often selective and sometimes misleading method, and Trudgill (1983: 41) argues that whilst native speaker preconceptions 'are less likely to be wrong, [...] if they are wrong they are more likely to be adhered to'.

As a Sotonian, I have compiled the lists of phonological and grammatical features below which I believe to be characteristic of speech in the Southampton area. Data on these features have been collected using the methods detailed in chapter 3, and a range of the features selected for analysis by the extralinguistic variables of sex, age, social class and identity. Chapter 3 provides details of the lexical features investigated by the present study, from which a selection has been chosen for examination by these same extralinguistic variables. In the case of several of the phonological variables, I have been able to employ earlier research conducted on the phonology of the Southampton area (Wallace 2000) as a means by

which to ensure the salience of variables investigated. In all other cases, I have employed native speaker intuition. The compilation of these lists did not preclude, however, the possibility that other salient variables might occur during fieldwork, nor did it imply that all features would be analysed by the extralinguistic variables under examination.

It should again be emphasised that neither the phonological nor the grammatical variables under investigation are specific to the Southampton area, but are employed to varying extents in many British English dialects. Indeed, the majority of these forms, both grammatical and phonological, occur extensively country-wide. What is of interest is not simply the occurrence of particular variables and their nonstandard variants – empirical research such as that by Cheshire (1982) and Cheshire et al. (1989) on grammar, and by Trudgill (1974) on phonology strongly suggest that they will occur – but rather how these variables correlate with sex, age, social class and identity. It is here that the purpose of the investigation of these features lies, since an examination of the relationship between selected linguistic and extralinguistic variables will enable a picture of the Southampton area dialect to be constructed, and, in the case of some of the phonological features, will facilitate discussion of east-west identity in the area.

1.7.3 Phonological Variables

1.7.3.1 Consonants

- **T Glottalling**

T Glottalling is the realisation of /t/ with a glottal stop, a glottal stop being ‘the sound (or, to be more exact, the lack of sound) that occurs when the vocal cords are held tightly together’ (Ladefoged 1993: 52). The glottal stop variant ‘can occur in any

non-initial post-tonic position' (Chambers 2003: 208), though it occurs more frequently in some phonological contexts than in others:

most frequent	word-finally before a consonant	<i>that man</i>
	before a syllabic nasal	<i>button</i>
	word-finally before a vowel	<i>that apple</i>
	before syllabic [l]	<i>bottle</i>
least frequent	word-internally before a vowel	<i>better</i>

(Hughes et al. 2005: 66)

The glottal stop is a 'stigmatized feature', a 'shibboleth of British W[orking]C[lass] speech' (Chambers 2003: 208-209), frequently commented upon unfavourably by the media (Foulkes and Docherty 1999: 11). Wells (1982: 323) describes the glottal stop as being 'widely regarded as a sound particularly characteristic of Cockney', Cockney being the 'traditional working-class dialect of London' (Wells 1982: 301-02). He argues:

It is certainly plausible to suppose that one of the principal factors contributing to the apparently recent geographical spread of T Glottalling is the influence of London English, where it is indeed very common.

(Wells 1982: 323)

T Glottalling has been found, however, in urban areas distant from London, such as Norwich (Trudgill 1999: 132), Hull (Williams and Kerswill 1999: 147), Newcastle upon Tyne (Watt and Milroy 1999: 29), Derby (Milroy et al. 1999: 39), and Glasgow (Macaulay 1977: 45). Glottal stops have 'a long independent history in Scotland [...], and there is [...] some evidence that they have been present in R[ecieved] P[ronunciation] for longer than has usually been credited' (Foulkes and Docherty 1999: 11).

T Glottalling has been found to correlate with the speaker variables of sex, social class and age. In their survey of London schoolchildren, Hudson and Holloway (cited in Wells 1982: 325) found that for /t/ in the intervocalic

environment, ‘the percentage of [ʔ] realizations for middle-class children was well under 10, for working-class girls 40, and for working-class boys 80’. In his Glasgow study, Macaulay (1977: 47) found the percentage use of the [ʔ] variant among middle-class informants to decrease with increasing age.

Previous research in the Southampton area has shown T Glottalling to be a widespread feature of local speech (Wallace 2000: 30-31). In the present study, data have been collected on this feature in intervocalic, preconsonantal, and word-final positions.

- TH Fronting

TH Fronting is the use of the labiodental fricative forms [f] and [v] instead of the dental fricatives [θ] and [ð] (Foulkes and Docherty 1999: 11). Wells states that the dental fricatives [θ] and [ð] are:

relatively unnatural segment types [...] learnt late by children. Compared with them, the labiodental fricatives, [f] and [v], are more natural; and children [...] readily substitute them for the difficult dentals.

(Wells 1982: 96)

TH Fronting is thus described by Wells (1982: 96) as a ‘persistent infantilism’ in the speech of the adult working-class Londoners who employ it.

While Trudgill (1999: 137) argues that TH Fronting was ‘formerly confined to the London area and to Bristol’, records exist which show it was present in Mid-Yorkshire in 1876 (Robinson 1876). Like T Glottalling, TH Fronting is found country-wide. Trudgill (1999: 138) reports it as being absent from his 1968 Norwich survey, but as making inroads into the local dialect by the time of his 1983 survey. It has also been reported in cities such as Derby (Docherty and Foulkes 1999: 51), Hull (Williams and Kerswill 1999: 147) and Glasgow (Stuart-Smith 1999: 209).

In the present study, data have been collected on TH Fronting, both of the voiceless interdental fricative [θ] and of the voiced interdental fricative [ð], in word-initial, intervocalic, and word-final positions.

- (ng)

The variable (ng) is found in the verbal ending *-ing* (*calling, trying, stopping*), as well as in the *-ing* of nouns such as *ceiling, morning, shilling, pudding*, and of adjectives such as *cunning* (Wells 1982: 262). Wells (1982: 262) states that (ng) is subject to social or stylistic alternation in almost all English-speaking communities. He cites the word *running* as an example, whereby the form with the velar nasal, [rʌnɪŋ], is associated with higher social class and more formal speech, and that with the alveolar, [rʌnɪn], with lower social class and less formal speech (Wells 1982: 262). Wells (1982: 262-63) makes the point, however, that there exists ‘geographical variation in respect of the point in social or stylistic stratification at which the changeover occurs’. He states:

In Birmingham, England, it appears that the velar form extends well down into working-class speech, while in Birmingham, Alabama, the alveolar form extends well up into middle-class or educated speech.
(Wells 1982: 263)

Wells (1982: 262) also argues that ‘it is probably not correct to regard [n] for [ŋ] as an innovation’, as both ‘alveolar and velar forms are to be found in early Middle English’.

- Rhoticity

Wells states:

In [...] **rhotic** accents [...] /r/ can occur, with an overt phonetic realization, in a wide variety of phonetic contexts, including pre-consonantal and absolute-final environments, thus *farm* [farm], *far* || [far]. In [...] **non-rhotic** accents /r/ is excluded from pre-consonantal

and absolute-final environments, thus [fɑ:m], *far* || [fɑ:] [original emphasis].

(Wells 1982: 75-76)

Rhoticity is found in most south-western accents (Altendorf and Watt 2004: 200-01).

The Survey of English Dialects (SED) reported its occurrence in the word *farm* in the South West and Hampshire (Orton and Wakelin 1967-68: 78), and it is in its role as a south-western, as opposed to a south-eastern, phonological variable that data have been collected on rhoticity in the Southampton area.

- L Vocalization

Wells (1982: 258) describes L Vocalization as the process by which the lateral [ɫ] is converted into a non-syllabic back vocoid, [ɣ], or its rounded equivalent, [o], (or [ʊ], since the precise quality varies). Thus, *milk* is pronounced [mɪʊk], as opposed to Received Pronunciation (RP) [mɪɫk], *shelf* is pronounced [ʃeʊf], as opposed to RP [ʃɛɫf], and *middle* pronounced [mɪdo], as opposed to RP [mɪdɫ]. Though Wells (1982: 258) limits the occurrence of L Vocalization to non-prevocalic positions, Tollfree (1999: 174) in her study of South East London English attests to vocalization in word-final intervocalic position in the speech of informants aged between 15 and 30, citing the example of *legal info* [liɣw^wɪnfɛɣ].

Though there have been ‘droppings or vocalizations of /l/ in various environments in the earlier history of English’, as in the words *calm* and *walk*, Wells (1982: 259) describes L Vocalization as a relatively recent feature, probably less than a century old in London. It is, however, widespread, having been reported as far afield as Derby (Docherty and Foulkes 1999: 53) and Glasgow (Macafee 1983: 38). Though a feature ‘to which low prestige is attached’ (Wells 1982: 20), Wells (1982:

259) predicts that it will become standard in English over the course of the next century.

In their study of pupils at a London secondary school, Hudson and Holloway (cited in Wells 1982: 20) found L Vocalization to correlate with both the speaker variable of sex and that of social class, with working-class children employing this feature more than middle-class children, and boys more than girls. Wells (1982: 259) describes L Vocalization as having its origins in London and the surrounding counties, and it is in its role as a south-eastern, as opposed to a south-western, phonological feature that data have been collected on L Vocalization in the Southampton area in word-final, word-final intervocalic, word-medial, and preconsonantal positions.

1.7.3.2 Vowels

Wells's (1982: xviii) standard lexical set keywords are used to describe those vowels under examination. Data have been collected on all vowels listed below in all those phonetic environments in which they occur.

- The FACE Vowel

The standard lexical set FACE is comprised of 'those words whose citation form in R[ecceived] P[ronunciation] has the stressed vowel /eɪ/' (Wells 1982: 141). It is an unrounded vowel, and is either a front narrow closing diphthong or, more unusually, a front half-close monophthong (Wells 1982: 141). In accents other than RP, there exist:

two other main types of quality for this vowel: monophthongs in the [e:] area, and wide diphthongs such as [ɛɪ, æɪ, ʌɪ]. [...] Wide diphthongs are southern, whether in the United States or [...] in England.

(Wells 1982: 142)

The [ʌɪ] realisation of the FACE vowel is typical of London speech (Altendorf and Watt 2004: 187). In the far South West, however, [e:] is often found in old-fashioned Devon and Cornwall speech, though there is a tendency nowadays for this monophthong to be replaced by diphthongs (Wells 1982: 347). In Hampshire, the Survey of English Dialects recorded the following realisations of the FACE vowel as found in the word *play*: [aɪ]; [e:]; [eɪ]; and [ɛɪ] (Orton and Wakelin 1967-68: 1003). Earlier research in the Southampton area showed the predominant local realisation of the FACE vowel to be [ɛɪ] (Wallace 2000: 39).

- The PRICE Vowel

The *Oxford English Dictionary (OED)* states that the standard lexical set PRICE is comprised of those words whose citation form in Received Pronunciation has the stressed vowel /ʌɪ/. Of variability in the realisation of the PRICE vowel, Wells states:

Very back starting-points, [aɪ - ɔɪ] are characteristic of the urban south of England [...]. A starting-point that is not fully open, [ɐɪ - ʌɪ - əɪ], is typical of the rural south of England.

(Wells 1982: 149)

Research conducted in the Southampton area (Wallace 2000) showed local realisations of the PRICE vowel to be: [ɔɪ], a realisation also found in the speech of the West Country (Wells 1982: 347); [aɪ], again found in West Country speech and also in London speech (Altendorf and Watt 2004: 187, 198); and [ɔɪ]. Both the [ɔɪ] and the [ɔɪ] realisation of the PRICE vowel were recorded in the word *lights* in Hampshire by the Survey of English Dialects (Orton and Wakelin 1967-68: 356).

- The BATH Vowel

The standard lexical set BATH is comprised of those words whose citation form contains the stressed vowel /ɑ:/ in RP (*OED*). It is a 'fully open unrounded vowel

lying between back and central' (Wells 1982: 158). BATH words 'belong phonetically with [...] PALM and START in RP' (Wells 1982: 133-34). In Southampton, however, Wells (1982: 346) reports that BATH words have '[æ:] in working-class speech, varying sociolinguistically with a backer vowel, towards [ɑ:], among the middle class'. The open unrounded backed variant [ɑ:] is typical of the South East, while [a] is typical of the South West (Hughes et al. 2005: 69). SED recorded the following realisations of the BATH vowel, as found in the word *path*, in Hampshire: [a:]; [a]; and [ɑ:] (Orton and Wakelin 1967-68: 425).

Data have been collected on this variable as it is a vowel which, according to Wells, is subject to social class differentiation in the Southampton area, and one which has clearly marked south-eastern and south-western variants.

1.7.4 Grammatical Variables

Using the questionnaire employed in the Survey of British Dialect Grammar (Cheshire et al. 1989), grammatical forms that native speaker intuition suggests are to be found in the Southampton area have been selected in order that data on these features can be collected. It should be noted, however, that in the case of both phonological and grammatical features, not all variables presented here will be analysed. Concentration will be on those significant to the interests of the Southampton area survey and these features will be examined according to the speaker variables of sex, age, social class and identity.

1.7.4.1 Verb Forms

- Present Tense Verb Form Endings - Nonstandard -s

On the matter of subject-verb concord, Biber et al. state:

The subject and the verb phrase agree in number and person [...]. The basic grammatical rule is that the *s*-form of lexical verbs and the primary auxiliaries [...] is used with a third person singular subject in the present tense indicative.

(Biber et al. 1999: 180)

The extension of the present tense third person singular suffix *-s* to other persons is described by Edwards et al. (1984: 18) as the most widespread feature of nonstandard grammar. This extension or generalisation, as in *they knows my secrets* and *I hates it when she does that*, has been reported as common practice in ‘Scotland, parts of Northern England, Herefordshire, parts of South Wales and parts of Southern (particularly South Western) England’ (Edwards et al. 1984: 18). In her Reading study, Cheshire (1982: 31) states that the nonstandard suffix occurs with most verbs, including the irregular verbs *say*, *have*, and *do*, and argues that its occurrence can be attributed to the previous influence of northern varieties of English. She states:

[W]e know that the Northumbrian dialect of Old English, for example, had an *-s* suffix throughout the present tense paradigm, and that this pattern was extended in the Middle English period to Midlands areas [...]. We also know that until about 1640 the suffix was used in standard English with singular subjects and occasionally with plural subjects [...], and we can assume that use of the suffix spread to Southwestern varieties of English also.

(Cheshire 1982: 31)

In Hampshire, SED recorded examples of nonstandard *-s* in both the first person singular form, *I eats* (Orton and Wakelin 1967-68: 679), and the third person plural, *they keeps chickens* (Orton and Wakelin 1967-68: 445).

- Past Tense Verb Forms - *To Be*

Cheshire states:

In standard English, past tense forms of BE differ from all other past tense verb forms by having more than one morphological form: *was* is used with first and third person singular subjects, and *were* is used elsewhere.

(Cheshire 1982: 44)

In their review of research, Edwards et al. note that many nonstandard dialects simplify the conjugation of the verb *to be* in the past tense. They state:

The most common tendency seems to be the use of *was* with all persons ([for example in] S[outh] W[est] England [...]). Generalisation of *were* to all persons were claimed for only two localities (Yorkshire, Dorset), but generalisation to the single persons only – with *was* in the plural – was claimed more often ([for example in] Cockney). [...] Finally, some areas seem to allow generalisation of both *was* and *were* ([for example in] Hampshire [...]).

(Edwards et al. 1984: 20)

The responses to the Survey of British Dialect Grammar indicated the use of nonstandard *was* in urban centres country-wide, though it occurred less frequently in Scotland and the North of England than elsewhere (Cheshire et al. 1989: 201).

Nonstandard *were* was also reported as ‘cooccurring with nonstandard *was* by schools in the South (though less frequently than nonstandard *was*)’ (Cheshire et al. 1989: 201). Both nonstandard *was* and nonstandard *were* were found in Hampshire by SED (Orton and Wakelin 1967-68: 1038-40), and data on both forms have been collected in the Southampton area.

- Past Tense *Done*

In some dialects, ‘the past participle form [of certain verbs is] generalized to the past tense function’ (Biber et al. 1999: 1124). This results in past tense *done*, as in the example *he done that wrong*. Cheshire et al. (1989: 207) state that this form ‘is thought to be used only for the lexical verb DO’. Past tense *done* is not believed to be used with auxiliary *do*, as in sentences such as *he did read my diary*. In the Survey of British Dialect Grammar, past tense *done* was reported by virtually all participating schools in the South of England, though it was less widespread

elsewhere (Cheshire et al. 1989: 207). This was noted with interest by Cheshire et al. (1989: 207), since ‘a tendency for this form to be more widespread in the south [...] than elsewhere has not previously been recognised’.

- *Sat* and *Stood* as Present Participles

Prior to the Survey of British Dialect Grammar, the use of the past participles *sat* and *stood* as present participles where Standard English uses *sitting* and *standing* had not been considered as a widespread feature of regional grammar (Cheshire et al. 1989: 200). Cheshire et al. state:

Hughes and Trudgill [...] say that these forms are ‘widely used in parts of the north and west of England’, and Edwards and Weltens [...] found them reported in only five areas of Britain (Manchester, North Lancashire, West Wirral, Herefordshire and Reading).

(Cheshire et al. 1989: 200)

They go on to argue that ‘we should have been alerted to the fact that these features are [...] widespread [...] by the fact that they have been noticed by prescriptivists’ (Cheshire et al. 1989: 200). Burchfield (1985: 54-55), for example, discussing his guide for the BBC, considers *he was sat there* as ‘unacceptable [...] in any circumstances’. Examples of these nonstandard forms can, however, be found in educated spoken English and in written, semi-formal English (Cheshire et al. 1989: 200).

Sat and *stood* as present participles were reported by the majority of schools taking part in the Survey of British Dialect Grammar. All but two schools in the North report their occurrence, along with 67% of schools in the Midlands, and 81% of schools in the South (Cheshire et al. 1989: 200). The preferences which this regional distribution reveals point perhaps to a ‘recent diffusion of these features from the north and west of England’, and might indicate that, ‘although they once

had a regional distribution. they are now becoming characteristic of a general nonstandard or semistandard variety of English' (Cheshire et al. 1989: 200).

- *There was* with a Plural Notional Subject

In Standard English:

the verb phrase combining with existential *there* takes its number from the **notional subject** [...]. A plural form is generally used with plural noun phrases; a singular form otherwise [original emphasis].

(Biber et al. 1999: 185-186)

The notional subject of a clause with existential *there* is the noun phrase which follows the verb *be* (Biber et al. 1999: 944). The following sentence features examples of existential *there* constructions, and was collected in Milton Keynes:

‘I went to Gemma’s house and [...] of course **there were** boys staying over as well [...] so **there was** like all our friends as well’

(cited in Cheshire 1999: 70)

In the first existential *there* construction in this sentence, *there were boys*, there is, as one would expect in Standard English, agreement between the form of *be*, *were*, and a plural notional subject, *boys*. In the second token, *there was like all our friends*, the singular form of the past tense of *be* is used with a notional subject containing a plural head noun. Biber et al. (1999: 944) argue that in existential *there* clauses, there is a ‘strong tendency in conversation to use a singular verb regardless of the number of the notional subject’. The Survey of British Dialect Grammar found *there was* with a plural notional subject to be ‘very widespread indeed’ (Cheshire et al. 1989: 199), so much so that Cheshire et al. (1989: 200) argue that it is ‘best seen as a stylistic feature of English, characteristic of colloquial, informal speech, rather than as a nonstandard feature’. They go on to state that *there was* with a plural notional subject:

can be [...] seen as an oral strategy, similar to French *il y a* and German *es gibt*, which is used to signal to the addressee that new information is being introduced in the thematic position in a sentence.

(Cheshire et al. 1989: 200)

Research conducted by Cheshire (1999: 70) in Milton Keynes shows that, like *il y a* and *es gibt*, '*there was* does indeed show signs of being a lexical unit'. This feature does not show the kind of sharp variation exhibited by features such as multiple negation and is frequently used by both working class and middle class female and male speakers (Cheshire 1999: 70).

- *Should of*

Should of, corresponding to Standard English *should have*, was included in the Survey of British Dialect Grammar as it was known to occur in the writing of schoolchildren in the South of England, and Cheshire et al. (1989: 196) wanted to know if it was widespread throughout Britain. The Survey responses show that it is (Cheshire et al. 1989: 196).

Cheshire et al. (1989: 196) believe *should of* to be fairly recent in origin, since it is not mentioned in early works on dialect, nor in prescriptive handbooks. Its etymology is unknown, though it seems likely, according to Cheshire et al., that:

should of (and parallel forms such as *must of*, *could of* and *better of*) derives from the phonetic reduction in informal spoken English of unstressed *have* to /əv/, which is phonetically identical to unstressed *of* (in, for example, *some of that* [sʌm əv ðæt]). The full form in both cases is then produced as [ɒv], with the original syntactic derivation of the verbal construction apparently forgotten.

(Cheshire et al. 1989: 197)

The Survey of British Dialect Grammar investigated *should of* in a full phrase (*you should of left an hour ago*) and also in an ellipted form (*you should of*). Data have also been collected on *should of* in both full and ellipted forms in the Southampton area study.

1.7.4.2 Negation

- Multiple Negation

Also referred to as negative concord, multiple negation is a ‘common and widespread’ feature of English nonstandard dialects (Chambers 2003: 129). Biber et al. (1999: 178) state that multiple negation occurs when ‘[t]wo or more negative forms [...] co-occur within the same clause to express a single negative meaning’. An example of multiple negation is *I ain’t done nothing*, corresponding to the Standard English *I haven’t done anything*.

Edwards et al. (1984: 17) state that ‘the use of more than one negative is a matter of concord, not a means of intensification’. Biber et al. (1999: 178), however, argue that ‘[b]ecause of the repetition of the negative forms, this type of negation appears to have a strengthening effect’. In the Survey of British Dialect Grammar, multiple negation was reported ‘less frequently in the North of Britain than in the Midlands, and most frequently in the South’ (Cheshire et al. 1989: 205-06).

Cheshire et al. (1989: 206) offer several suggestions by way of explanation for the relative scarcity of this feature in the North and the Midlands. It may be the result of some flaw in the questionnaire item employed by the Survey. Another explanation is that ‘the feature is recessive in some of the urban centres of the country’.

Alternatively, Cheshire et al. (1989: 206) suggest that multiple negation may be ‘so heavily stigmatised that despite our best efforts to raise the status of dialect in the classroom’ the pupils participating in the Survey ‘simply failed to report it’. The authors suggest that ‘only empirical research can determine what really is happening to multiple negation in urban varieties of English’ (Cheshire et al. 1989: 206).

- Never as Past Tense Negative

Cheshire states that in present-day English:

never continues to be used as a negative marker, sometimes with the possibility of the literal meaning of universal temporal negation, [...] and sometimes without universal temporal reference, [...] where the utterance refers to the past.

(Cheshire 1998: 128)

An example of the use of *never* with universal temporal meaning is *I've got two left feet so I never dance*. An example of *never* without universal temporal meaning where the utterance refers to the past is *I never broke it*. *Never* as it is used in the latter example 'functions in much the same way as *didn't*, but it is followed by a past tense verb form not by an infinitive' (Edwards et al. 1984: 18). Cheshire argues that, from a prescriptivist point of view:

[i]t is [...] incorrect to use *never* when referring to one occasion. *Never* can only be used in continuous contexts [...]. Nevertheless, *never* does still occur with simple past tense tenses with the meaning 'not'.

(Cheshire 1998: 129)

Though the use of *never* as a 'negative particle has been a favoured strategy of negation throughout the history of English', the use of *never* with past tense verbs is 'proscribed by prescriptive grammarians' (Cheshire et al. 1989: 197):

Fowler 1965 (*Modern English Usage*): 'this use of *never*, however illogical, is idiomatic, at least colloquially'.

Collins Dictionary of the English Language (1981): 'in good usage, *never* is not used with simple past tenses to meant [sic] *not*'.

Wood 1981 (*Current English Usage*): '*never* means 'not ever, on no occasion''.

(cited in Cheshire 1998: 129)

Cheshire (1998: 129-30) argues, however, that the use of *never* with simple past tenses with the meaning 'not' is not just a feature of nonstandard grammar. It also occurs in what is usually considered to be 'standard English – in other words, in published written prose and in the speech of people who consider themselves to be educated' (Cheshire 1998: 129-30). *Never* as a past tense negative was 'amongst the

most widely reported features' in the Survey of British Dialect Grammar (Cheshire et al. 1989: 197).

- Ain't

Cheshire describes *ain't* as follows:

The non-standard form corresponds to several standard English verb forms. It occurs as the negative present tense contracted form of *be*, both as the copula:

1. We've got a park near us, but there ain't nothing over there and as the auxiliary:

2. How come that ain't working?

It is also used as the negative present tense contracted form of the auxiliary *have*:

3. I ain't got one single flea in my hair, they're all married.

(Cheshire 1991: 54)

Though 'widely regarded as non-standard', *ain't* is 'relatively widespread in use' (Biber et al. 1999: 167). Edwards et al. (1984: 18) report it as being particularly prevalent in the dialects of southern England. It was recorded by the SED in Hampshire as a negative present tense contracted form of both *be* (Orton and Wakelin 1967-68: 1135-37) and *have* (Orton and Wakelin 1967-68: 1118-19), and data have been collected on these forms in the Southampton area by the present study.

1.7.4.3 Adverbs

- Adverb Formation

In Standard English, many adverbs are formed by adding the suffix *-ly* to the base form of an adjective, as in the examples *quick – quickly* and *cold – coldly* (Biber et al. 1999: 539). The addition of this suffix is, however, 'optional in virtually all dialects' (Edwards et al. 1984: 24), and the Survey of British Dialect Grammar shows the zero form to be widespread (Cheshire et al. 1989: 202). Use of adverbs without the *-ly* suffix 'is associated with non-standard [...] or colloquial language' (Tagliamonte and Ito 2002: 237), and Cheshire et al. (1989: 202) suggest it is

possible that some adverbial forms without the *-ly* suffix ‘express intensity or emphasis’.

In Tagliamonte and Ito’s (2002) study of adverbs in the York English corpus, the *-ly* form was found to be dominant, accounting for 85% of dual form adverbs (Tagliamonte and Ito 2002: 248). However, all speakers in the corpus used the zero adverb some of the time (Tagliamonte and Ito 2002: 252). Tagliamonte and Ito (2002: 236) state that ‘[i]n contemporary spoken corpora, variability [in the marking of adverbs] is rampant even in the speech of the same individual in the same conversation’.

The adverb *really*, which accounted for the majority of the total number of adverbs in the York corpus, was found to behave differently to other adverbs. *Really* was used far more frequently by younger speakers than it was by their older counterparts, who favoured the zero form *real*, regardless of level of speaker education or sex (Tagliamonte and Ito 2002: 251-52). According to Tagliamonte and Ito (2002: 238), this apparent time change in the use of *really* was ‘not due to adverb formation processes’ but was ‘the result of changes in fashion amongst English intensifiers’.

By contrast, age was not found to be significant in the case of the other adverbs in the corpus (Tagliamonte and Ito 2002: 251). Less educated men were the main users of zero adverbs other than *real*, which were shown to be stable sociolinguistic markers (Tagliamonte and Ito 2002: 252).

However, what Tagliamonte and Ito (2002: 258) describe as ‘the most significant factor contributing to variation in adverb formation’ was ‘a propensity for *-ly* with abstract meanings [for example, *if I remember rightly*] and zero with concrete meaning [for example, *I’ve walked upstairs dead quick*]’.

1.7.4.4 Nominal Constructions

- Prepositions

Edwards and Weltens (1985: 114), and Hughes et al. (2005: 32) note that there is a wide range of variation where prepositional use is concerned. The Survey of British Dialect Grammar investigated:

both the use of simple prepositions where St[andard] E[n]glish has complex prepositions ([...] *'I'm going up/down/over my friend's house later'*) and the use of a complex preposition where St[andard] E[n]glish has a simple preposition ([...] *he knocks his hat off of his head*).
(Cheshire et al. 1989: 206)

These nonstandard forms were reported more frequently in the South than in the North or the Midlands (Cheshire et al. 1989: 206). Data concerning the use of both nonstandard simple and nonstandard complex prepositions of the type investigated by Cheshire et al. (1989: 206) have been collected in the Southampton area study.

- Demonstrative *them*

Biber et al. (1999: 347) state that '[i]n addition to marking something as known, [...] demonstrative forms specify whether the referent is near or distant in relation to the addressee'. The Standard English demonstratives are *this*, *that*, *these* and *those*. In the case of nonstandard British English dialects, *them* corresponds to *those*, resulting in sentences such as *'If you walks across, there's a load of them bus stops'* (example collected in Reading, cited in Cheshire 1982: 78). Demonstrative *them* was the most widely reported feature in the Survey of British Dialect Grammar, with 97.50% of schools reporting it (Cheshire et al. 1989: 195). In her Reading study, Cheshire (1982: 78) found the frequency of occurrence of demonstrative *them* to be very high in the speech of the male adolescent participants, but lower in that of the female adolescents.

- Comparative and Superlative Adjectives

Biber et al. (1999: 521) state that '[a]djectives capable of representing degrees of a characteristic are said to be **gradable** [original emphasis]'. An example of a gradable adjective is *pretty* – *prettier* – *prettiest*, compared to the arguably ungradable adjective *atomic*, **more atomic*. Gradable adjectives can take comparative and superlative forms when used to compare two or more things. In Standard English, this comparison can be marked either inflectionally, with the suffixes *-er* and *-est*, or phrasally, using the degree adverbs *more* and *most* (Biber et al. 1999: 521-22). Exceptions to these means of comparison are the adjectives *good* and *bad*, which take the forms *good* – *better* – *best*, and *bad* – *worse* – *worst* respectively. Except for a few forms, among them *right*, *wrong*, and *real*, monosyllabic adjectives take the inflectional suffix, whereas longer adjectives usually take phrasal comparison (Biber et al. 1999: 522).

In nonstandard dialects throughout Britain, however, *-er/-est* 'may be added to all adjectives' to produce comparative and superlative forms such as, for example, *beautifuller* and *beautifullest* (Edwards et al. 1984: 28). A feature referred to by Edwards et al. (1984: 28) as 'double comparison' can also be employed to form the comparative and superlative forms of gradable adjectives. Double comparison is 'the simultaneous application of both the periphrastic comparison and the addition of *-er/-est*', and results in sentences such as '*I've never seen a more beautifuller one*', and '*This is the most beautifullest house I've seen*' (Cheshire et al. 1989: 224). It is not, however, a recent innovation. Wakelin states:

In early N[ew]E[nglish] it was possible to add *more*, *most* to the comparative and superlative pleonastically, thus in Shakespeare 'This was the most unkindest cut of all' (Julius Caesar, III.ii.185), and this custom survives in the dialects.

(Wakelin 1977: 117)

What might also be considered to be a type of double comparison 'is the use of inflected suppletive forms such as *worser* and *leastest*', though this is less common than 'ordinary' double comparison (Edwards et al. 1984: 28). Data have been collected in the Southampton area on this feature, along with the more common form of double comparison, and the addition of *-er/-est* to all adjectives.

- Regularisation of Reflexive Pronouns

In Standard English, four of the eight reflexive pronouns (*myself*, *yourself*, *yourselves*, *ourselves*,) are formed by added the suffix *-self* (singular) or *-selves* (plural) to the possessive determiner (Biber et al. 1999: 328). In the case of *himself*, *itself* and *themselves*, *-self* or *-selves* is added to the accusative form of the personal pronoun (Biber et al. 1999: 328). *Herself* could be argued to be based on either the possessive determiner or the accusative form of the personal pronoun.

Hughes et al. (2005: 30-31) state that '[m]any non-standard dialects have regularized the reflexive pronoun system so that [...] all forms are based on the possessives'. This results in *hissself* and *theirselves*, corresponding to the Standard English *himself* and *themselves*. In the Survey of British Dialect Grammar, *hissself* was reported by 63.20% of the participating urban schools, whilst *theirselves* was reported by 77%, and there appeared to be no geographical pattern to the distribution of responses (Cheshire et al. 1989: 207). Data have been collected in the Southampton area on both *hissself* and *theirselves*.

- Relative Pronouns

Biber et al. (1999: 608) state that in Standard English there are five relative pronouns: *which*; *who*; *whom*; *whose*; and *that*. They add that '[t]he relative pronoun can sometimes be omitted altogether' (Biber et al. 1999: 608). In nonstandard dialects, however, other relative pronouns can be used. These include *what*, *as* and

at, as in sentences such as: *the films what I like best are horror films: the films as I like best are horror films*; and *the films at I like best are horror films* (Cheshire et al. 1989: 222). Though Edwards and Weltens's (1985: 116) review of research concluded that *what* was not found in the North of England and was less common than other nonstandard relative pronouns such as *as*, the Survey of British Dialect Grammar showed *what* to occur frequently country-wide (Cheshire et al. 1989: 198). *As* and *at* were reported very infrequently in the Survey of British Dialect Grammar, leaving Cheshire et al. (1989: 198) to conclude that *what* seems to be 'the preferred relative pronoun in the urban centres of Britain today'. SED recorded the nonstandard relative pronouns *as* and *what* in Hampshire (Orton and Wakelin 1967-68: 288-92), and data on both these and nonstandard *at* have been collected in the Southampton area.

- Nouns of Measurement with Zero Plurals

In Standard English, the plural ending of a noun is *-s* or, if the noun ends in *s*, *z*, *x*, *sh*, or *ch*, the ending takes the form *-es* (Biber et al. 1999: 285). Edwards et al. (1984: 25) state that '[i]n British dialects it is almost a universal rule that, after numerals, nouns of measurement and quantity retain their singular form'. Examples of this phenomenon include *two inch*, *three pound*, and *four mile*. In the Survey of British Dialect Grammar, nouns of measurement with zero plurals were reported as being very widespread in urban varieties of English, but some nouns of measurement seemed 'more likely to have no plural marking than others' (Cheshire et al. 1989: 198). *Twenty mile*, for example, was reported by 75.90% of the participating schools, whereas *three inch* was reported by only 43.70%. A variety of nouns of measurement have been investigated in the Southampton area.

1.7.5 Lexical Variables

The Southampton area survey employs the methodology devised for the Survey of Regional English (SuRE) as the primary means by which to collect lexical, grammatical and phonological data (Upton and Llamas 1999). SuRE and its methodology are discussed in detail in chapter 3.

1.7.5.1 Lexis and the Issue of Dialect Versus Slang

Like the grouping of questions by subject matter in the questionnaire employed by the Survey of English Dialects (Orton et al. 1962-1971), the Sense Relation Networks (SRNs) which form the basis of the SuRE methodology are built around semantic fields (Llamas 1999: 98). Though drawing on the SED in terms of content design, the SRNs employ only a limited number of notion words investigated by the Survey of English Dialects. Llamas states:

When selecting standard notion words, the wish to include the same standard notion word as the SED where possible and appropriate was borne in mind, as a direct comparison could reveal potential real time change. Due to the urban bias of the proposed survey [...], however, this proved inappropriate in most cases, with few SED notion words remaining.

(Llamas 1999: 102)

Though similar in that both are concerned with capturing linguistic variation in England, in the case of the SED, and in the British Isles as a whole, in the case of SuRE, in many other respects, the aims of the two Surveys are very different. The informants chosen for the SED were predominantly males over the age of 60, in particular those who were, or had been, involved in farming, 'for it is amongst the rural populations that the traditional types of vernacular English are best preserved to-day' (Orton 1962: 14). The SED has a strongly philological bent, the emphasis being on the recording of traditional dialect. SuRE, on the other hand, does not aim to search out such 'traditional' dialect, though some may, of course, be found.

Grammar and phonology aside, the SRNs are designed to allow the collection of a wide range of lexis used by males and females, young and old, in Britain today. The SuRE methodology is not concerned with notions of dialect pedigree. Whereas grammar and phonology are opened in their entirety to 'dialectal' investigation, lexis, by virtue of the existence of apparently tidy boxes of 'dialect' and 'slang' into which items can be placed, has found, until this point, its 'dialectal' aspect circumscribed. Lexical items defined as 'slang' have previously been ignored. There exists, however, great overlap between definitions of dialect and slang. The *Oxford English Dictionary* defines dialect as a:

manner of speech peculiar to, or characteristic of, a particular person or class [or as]
 [o]ne of the subordinate forms or varieties of a language arising from local peculiarities of vocabulary, pronunciation, and idiom.
 (*Oxford English Dictionary*)

Slang, it defines as:

[t]he special vocabulary or phraseology of a particular calling or profession; the cant or jargon of a certain class or period [or as]
 [l]anguage of a highly colloquial type, considered as below the level of standard educated speech, and consisting either of new words or of current words employed in some special sense.
 (*Oxford English Dictionary*)

Both definitions are of nonstandard lexical items used as a marker of group identity, whether this be regional, class-, age- or profession-based. As such, allocation of lexical items to one or the other of these groupings can be problematic. Are, for example, words and phrases coined by adolescents in a particular region slang, by virtue of their novelty, or dialect because of their regional slant or their espousal by a social group? In light of the overlap between the definitions of the two terms, to ignore lexical items defined as slang can be seen to be a questionable practice, and one which results in the neglect of many interesting lexical items. It is for these

reasons that the SuRE methodology does not concern itself with notions of dialect and slang, but instead prompts the recording of all examples of lexical variation.

Chapter 2

2.0 Literature Review

In this chapter, literature pertaining to the extralinguistic variables of speaker sex, age, social class and identity are detailed and discussed.

2.1 The Speaker Variable of Sex

In examining the speaker variable of sex, it is necessary to distinguish between this and the concept of speaker gender. Whilst sex is a biological given, gender is a cultural construct. Since the two concepts are closely linked – Chambers (2003: 118) states that ‘[g]ender differences are partly based on sex differences’ – sociolinguists do not always make this distinction. Eckert argues that:

we have been examining the interaction between gender and variation by correlating variables with sex rather than gender differences. This has been done because although an individual’s gender-related place in society is a multi-dimensional complex that can only be characterized through careful analysis, his or her sex is generally a readily observable binary variable.

(Eckert, cited in Chambers 2003: 118)

Gender-based variability, Chambers (2003: 119) argues, emerges ‘when gender roles differ in terms of the mobility of women and men in a community’. Sex-based variability, on the other hand, occurs ‘even in the absence of well-defined gender roles’ (Chambers 2003: 119). In the Southampton area study, speakers will be selected according to their sex, since, as Eckert states, this is an easily identifiable binary variable.

2.1.1 Linguistic Sex Differentiation

Chambers states:

In virtually all sociolinguistic studies that include a sample of males and females, there is evidence [...] women use fewer

stigmatized and non-standard variants than do men of the same social group in the same circumstances.

(Chambers 2003: 116)

There exists no single or wholly accepted explanation as to why women should approximate more closely than men to the standard, nor does consensus exist as to why linguistic sex differentiation in general occurs. Labov (1966: 312) suggests that the trend among female participants in his New York study to adopt a wider range of variants than their male counterparts can be called 'hypercorrection'. He later argues that, possibly as a result of their influence on children in the first stages of language acquisition (Labov 1972b: 302-03), women 'are more sensitive than men to overt sociolinguistic values' (Labov 1972b: 243).

Trudgill (1972: 182) offers two inter-connected theories by way of explanation for the greater use of forms associated with the prestige standard among the female informants in his Norwich survey. The first of these focuses on male-female power relations and, like Labov's theory, on female linguistic sensitivity. Trudgill states that 'women in our society are more status-conscious than men, generally speaking [...], and are therefore more aware of the social significance of linguistic variables', and argues that there are two possible reasons for this status consciousness and resultant use of a greater proportion of forms approximating to the standard among females (Trudgill 1972: 182-83). The first is the need for women, subordinate to men and with a less secure social position than them, to signal their status linguistically in the absence of any other means, such as employment, by which to do this. The second is the fact that women tend to be rated on how they appear, rather than on what they do. Whilst men can be rated socially according to their occupation and earning power, the same is often not true of women, for whom other status signals, such as speech, are therefore more important. Trudgill's (1972:

183) second theory is concerned with working class speech and the differing expectations surrounding male and female behaviour. He states:

W[orking] C[lass] speech, like other aspects of WC culture, appears [...] to have connotations of masculinity [...], probably because it is associated with the roughness and toughness supposedly characteristic of WC life which are, to a certain extent, considered to be desirable masculine attributes. They are not [...] considered to be desirable feminine characteristics. On the contrary, features such as refinement and sophistication are much preferred.

(Trudgill 1972: 183)

According to this theory, women tend to use more standard forms than men because to do otherwise would be to behave in an unsuitable manner for their sex.

Like Trudgill's first theory, Deuchar's explanation for linguistic sex differentiation also focuses on the notion of power. Again assuming, like Trudgill, that women are relatively powerless speakers, Deuchar, employing Brown and Levinson's notion of 'face', 'the public self-image that every member wants to claim for himself', states that:

the use of standard speech, with its connotations of prestige, appears suitable for protecting the face of a relatively powerless speaker without attacking that of the addressee.

(Deuchar 1988: 31)

The explanations of all three sociolinguists have been criticised, however. Chambers (2003: 147) argues that, like Labov's and Trudgill's explanations, Deuchar's 'explanation in terms of face-saving is essentially a negative attribute'. He goes on to state:

The presupposition of all three explanations is that women are somehow compensating for shortcomings. To Trudgill, they are affecting the trappings of social status that they otherwise do not have; to Labov, they are exceeding the norms appropriate to their stations; and to Deuchar, they are offering tacit apologies to their overlords. All of these are basically negative motives.

(Chambers 2003: 147)

Instead, Chambers (2003: 147) argues that empirical evidence ‘shows women to be much more able performers than men in the whole spectrum of sociolinguistic situations’, commanding a wider range of linguistic variants and having the ability to alter their speech as situations warrant.

Though Cameron (1988: 11) supports Deuchar’s theory on the grounds that it examines gender differentiation in terms of power, ‘a point which can hardly be overemphasised’, she, like Chambers, criticises Labov and Trudgill. While acknowledging that when Labov and Trudgill were conducting their fieldwork in the late 1960s, their explanations for the differences in male and female speech ‘must have seemed reasonable enough’, Cameron states that the social role assigned to women according to Labov and Trudgill:

involved paying attention to appearances and superficial aspects of behaviour to a higher degree than was expected of men [...]. It involved responsibility for transmitting the norms of speech to children (something which would make women especially sensitive to correctness [...]) and finally, it denied women the opportunity to pursue social status through *work* in the same way men did [original italics].
(Cameron 1988: 5)

Feminists have criticised ‘almost every aspect of the quantitative paradigm’s dealings with women’, arguing that bias and stereotype are inherent in sociolinguistics, in its ‘methodology, measuring instruments and scoring systems, theoretical assumptions and individual interpretations’ (Cameron 1988: 5). This bias must not be ignored, Cameron (1988: 5-6) argues, since ‘studies of ‘difference’ are not just disinterested quests for the truth, but in an unequal society inevitably have a political dimension’.

Eckert and McConnell-Ginet (1999: 197) take the debate a step further, however, arguing that generalisations such as ‘‘‘Women use more standard forms, are more polite’’’ are ‘inadequate characterizations’. There are many exceptions to such

generalisations, and Eckert and McConnell-Ginet (1999: 193) state that ‘the real nature of valid gender generalizations depends on fully accounting for these “exceptions” as well as for the “typical” pattern’.

Eckert and McConnell-Ginet (1999: 185) approach the matter of gender from the perspective of the Community of Practice (CofP), ‘a group whose joint engagement in some activity or enterprise is sufficiently intensive to give rise over time to a repertoire of shared practices’. They see gender as being part of a person’s identity, and argue that identity is negotiated by participation in a range of Communities of Practice. Eckert and McConnell-Ginet state:

Gender emerges, in large measure, from differentiation in the kinds of CofP in which males and females tend to participate, and from the differentiated forms of participation that males and females tend to develop in mixed-gender communities of practice.

(Eckert and McConnell-Ginet 1999: 188)

In Eckert’s study of adolescents in Detroit, the girls’ use of the variables under examination showed greater variability than the boys’ (Eckert and McConnell-Ginet 1999: 195). The middle class jock girls were the most standard in their pronunciation of the variables, and the working class burnout girls the least. Though their male counterparts’ use of the variables followed the same pattern, with jock boys using more standard forms than burnout boys, the variability in male use of the variables was smaller than the female variability. Eckert and McConnell-Ginet state:

In Eckert’s research, standard language usage seems to be actively pursued by those young women who identify themselves with the school’s corporate culture (and the middle-class aspirations it supports); it is roundly avoided by those who reject such an identification.

(Eckert and McConnell-Ginet 1999: 195)

The authors recommend that linguists move ‘away from properties that women and men might have, and toward their social practices and social relations’ (Eckert and Mc-Connell-Ginet 1999: 198). Eckert and McConnell-Ginet do not believe that

binary statements about differences between men and women's speech are sufficient, and argue that linguists must look to the Communities of Practice in which men and women participate in order to explain adequately their observations.

2.1.2 Feminist Criticism

Feminists have criticised sociolinguists' dealings with women on several grounds. Their first criticism concerns the 'invisibility' of women, the fact that they are often 'excluded from research or at best, defined as peripheral and 'deviant'' (Cameron 1988: 6). Sociolinguistics in Britain inherited 'a tradition of work in dialectology from which women informants were almost completely absent' (Cameron 1988: 7). The Survey of English Dialects tended to select non-mobile, older, rural males as informants, on the grounds that 'men speak vernacular more frequently, more consistently, and more genuinely than women' (Orton 1962: 15). In the United States, Labov et al.'s (1968: 57) study of peer group speech in the Black community was based solely on male adolescent subjects. Like Orton, they too have argued that 'males are the chief exemplars of the vernacular culture' (Labov et al. 1968: 41), yet Conklin (cited in Cameron 1988: 7) claims that 'no conclusive evidence has been presented' to support this.

In her Belfast survey of three working class communities, Milroy (1987a) shows that it is not always the case that women use fewer stigmatised and nonstandard variants than their male counterparts. Examining the variables of age, neighbourhood and social network in addition to the speaker variable of sex, Milroy (1987a: 149) discovered that, in some cases, the young female participants in the Clonard district of Belfast had higher vernacular scores than their male peers. She argues that 'personal network structure is in these communities of very great importance in predicting language use', including linguistic sex differentiation

(Milroy 1987a: 160). Dense, multiplex networks are traditionally associated with men, particularly those living in working class communities (Milroy 1987a: 144), and this kind of tightly-knit network has ‘the capacity to impose [...] *linguistic* norms upon its members [original italics]’ (Milroy 1987a: 136). Unlike men, women have tended to participate in less cohesive social networks, these looser-knit networks having less capacity to enforce linguistic norms. These traditional network structures were not evident in Clonard, however. At a time of high male unemployment, the young women from Clonard both worked and socialised together, and thus had tighter-knit networks than either young Clonard men or women from other districts (Milroy 1987a: 148). This in turn influenced their use of nonstandard variables.

In Reading, Cheshire (1982: 86) found that, while adolescent male informants employed a higher proportion of nonstandard variants for the majority of grammatical variables under investigation, female informants used the nonstandard present tense third person singular form *do*, as in the example *it bothers me more than it do her*, to a greater extent than their male counterparts.

More significant for this study, perhaps, are Cheshire’s (1982: 92-107) findings when she further subdivided her male and female groupings. She constructed a ‘vernacular culture index’ for the male participants, which rated them according to six factors: carrying of weapons; style; job; criminal activities; skill at fighting; and swearing. The boys were then classified into four groups on the basis of their index scores. Female participants in the study ‘did not belong to a closely-knit group’, nor did they have ‘a clearly defined system of cultural values’ (Cheshire 1982: 107). It was not appropriate, therefore, to use the ‘vernacular culture index’ to measure their vernacular loyalty. Instead, the girls were divided between those who

‘showed some degree of adherence to a culture other than the mass ‘legitimate’ teenage culture [...] and the three girls who did not’ (Cheshire 1982: 107). The frequency with which a range of the study’s nonstandard grammatical variables occurred in the speech of the members of each of the male and female groupings was then analysed in order to ascertain which linguistic features functioned as markers of vernacular loyalty. Some grammatical variables were found to serve as markers of vernacular loyalty for girls but not for boys, and vice versa (Cheshire 1982: 109).

Cheshire states:

The main point that emerges from this analysis [...] is not that girls are more susceptible to the overt norms governing the use of standard English features (though this is certainly *to some extent* true), but that different linguistic features are used in different ways by boys and girls [original italics].

(Cheshire 1982: 110)

Just as with Milroy’s Clonard findings, had Cheshire stopped at a straightforward analysis of sex differentiation, this point would have been missed. By looking beyond the sex of the speakers to their social networks and gender roles, both sociolinguists have been able to challenge long-held assumptions regarding linguistic sex differentiation. Their findings do not, however, go as far as to negate the traditional assumptions underlying adult gender roles.

Further feminist criticism has been levelled at the measuring instruments and scoring systems employed by sociolinguists. In his original Norwich survey, Trudgill (1974: 35-36) employed a social class index, SCI, in order to stratify his informants. Six indicators were used in constructing the SCI: occupation; income; education; housing; locality; and father’s occupation. For the category of occupation, Trudgill (1974: 38) rated married women and widows on their husbands’ occupation, and unmarried women on that of their fathers’ because ‘with the still limited employment opportunities for women – that is, especially in professional and

administrative roles – occupation is not a satisfactory index of social status for women in our society’ (Glass and Hall, cited in Trudgill 1974: 38). Where, however, a working woman had an occupation with a higher status than her husband’s or father’s job, her own occupation was used (Trudgill 1974: 38-39). In the case of income, ‘women [...] were ranked as for occupation’ (Trudgill 1974: 39). Cameron and Coates state:

The underlying assumption [...] is that the whole family takes its position from the status of the father, who is assumed to be the main breadwinner. This latter assumption is by no means obvious in a society where male unemployment is widespread, and where divorce often results in single-parent families headed by women. Furthermore, Trudgill is prepared to ignore it in certain cases: he classifies married women by their own occupations if these outrank the husband’s job on the Registrar-General’s scale.

(Cameron and Coates 1988: 18)

They argue that ‘this model – on which sex-difference findings depend – itself uses sex-differentiated criteria’ (Cameron and Coates 1988: 18).

Though Cameron and Coates (1988: 18) object to these kinds of ‘inconsistencies and absurdities’ in stratification studies, on the grounds that they may prejudice findings, it should again be noted that Trudgill was working over thirty years ago in a very different social context to that of Cameron and Coates. In the late 1960s and early 1970s, there was no absurdity in Trudgill’s assumptions regarding male-female roles and power relations. For very many they were – and still are – a matter of fact. Indeed, Eckert (2000: 107) states that since fewer than half of the mothers of the adolescent subjects in her Belten High study worked outside the home, it was ‘impossible to use mother’s occupation as an indication of socioeconomic status’.

It is clear that the speaker variable of sex is the subject of much debate. It is also clear that it is of great importance, having been shown to correlate with

linguistic variation in research conducted in a number of locations (Labov 1966 in New York City; Trudgill 1974 in Norwich; Cheshire 1982 in Reading; Milroy 1987a in Belfast). There is every reason to suppose that linguistic sex-differentiation will be found in the speech of the Southampton area.

2.2 The Speaker Variable of Age

Not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity.
(Weinreich et al. 1968: 188)

Like the speaker variable of sex, age has been shown to influence language use significantly (Labov 1972b; Trudgill 1974; Macaulay 1977; Milroy 1987a).

Differences in language use between age groups are often indicative of change in progress. Chambers states:

Where change is involved, a certain variant will occur in the speech of children though it is absent in the speech of their parents. or, more typically, a variant in the parents' speech will occur in the speech of their children with greater frequency, and in the speech of their grandchildren with even greater frequency. In the community at large, successive generations will show incremental frequencies in the use of the innovative variant. The

logical conclusion, as time goes by, will be the categorical use of that new variant and the elimination of older variants.
(Chambers 2003: 203)

On some occasions, however, differences between the speech of one generation and another are not indicative of changes in progress, but rather a result of age-graded changes. Chambers states:

Age-graded changes are usually thought of as changes in the use of a variant that recur at a particular age in successive generations. They are, then, regular and predictable changes that might be thought of as marking a developmental stage in the individual's life.
(Chambers 2003: 206)

Labov (1994: 84) states that ‘if individuals change their linguistic behavior throughout their lifetimes, but the community as a whole does not change’. then the pattern is one of age-grading. Though less common than changes in progress (Chambers 2003: 206), many well-established sociolinguistic variables, Labov (1994: 73) argues, exhibit such age-grading, particularly ‘where adolescents and young adults use stigmatized variants more freely than middle-aged speakers’.

2.2.1 Empirical Research on the Linguistic Effect of Speaker Age

2.2.1.1 Age-Grading

- Trudgill on Norwich – (ng)

In Trudgill’s survey of Norwich English, /ŋ/, the final consonant in words such as *laughing* and *morning*, was found to be subject to age-grading (Chambers and Trudgill 1998: 79). The youngest and oldest speakers in the survey had the highest (ng) index scores, in that they used the highest levels of the [n] variant, with the middle-aged speakers having significantly lower (ng) scores. Chambers and Trudgill (1998: 79) explain this pattern by analysing the social situation of each of the three age groups. In childhood and adolescence, they argue, ‘the most important social pressures come from the peer group’ and younger speakers ‘are more strongly influenced by their friends’ than by the standard (Chambers and Trudgill 1998: 79). Use of nonstandard or stigmatised variants therefore results. As speakers get older and begin working they:

move into wider and less cohesive social networks [...], and are more influenced by mainstream societal values and, perhaps, by the need to impress, succeed and make social and economic progress.

(Chambers and Trudgill 1998: 79)

It is also during this period that many people raise a family, and so may feel the need to ‘set a good example’, linguistically and otherwise, to their children. For these

reasons, they often employ more prestige linguistic forms to achieve their goals. For older people, ‘social pressures are again less, success has already been achieved (or not, as the case may be), and social networks may [...] be narrower’ (Chambers and Trudgill 1998: 79). They are under less influence from the standard, and therefore employ more nonstandard or stigmatised forms.

- Macaulay on Glasgow

Predicting ‘that language would vary according to the age, sex [and] social class [...] of the informant’, Macaulay’s (1977: 18) speaker sample for his survey of the speech of Glasgow comprised equal cells of informants of both sexes from three age groups in four social class categories. The age groups were 10, 15, and adult, and the social class categories were based on the Registrar-General’s classification of occupations: class I (professional and managerial); class IIa (white-collar, intermediate non-manual); class IIb (skilled manual); and class III (semi-skilled and unskilled manual). A number of phonological variables were investigated, one of these being the glottal stop, which was subsequently found to be subject to age-grading among classes I and IIa (roughly equating to the middle and lower-middle classes). Macaulay states:

In class I [...] there are considerable differences between the age groups, with the 10-year-olds using almost six times as many glottal stops as the adults. Class IIa [...] also [exhibits] substantial age differences [...], with fifteen-year-old boys showing a particularly high percentage of glottal stops. [...] In class III there are only very slight age and sex differences with all speakers using a very high percentage of glottal stops.

(Macaulay 1977: 47)

Macaulay argues that ‘of all features of Glasgow speech the most notorious is the glottal stop’. It is ‘the feature most frequently singled out by teachers as characteristic of a Glasgow accent’, and ‘is the most openly stigmatised feature of Glasgow speech’ (Macaulay 1977: 45-47). As such, it is unsurprising that use of glottal stops decreases with age amongst classes I and IIa. With increasing age,

‘M[iddle]C[lass] Glaswegians learn to control the use of the W[orking]C[lass] shibboleth in their speech’ (Chambers 2003: 211). It is similarly unsurprising that most of the 10- and 15-year-olds in Macaulay’s study use the glottal stop more frequently than their adult counterparts, as research into phonological and grammatical variation has shown that adolescents lead all other age groups ‘in the use of vernacular forms’ (Eckert 2003: 391).

2.2.1.2 Change in Progress – Variant on the Increase

- Labov on Martha’s Vineyard

On the island of Martha’s Vineyard, Massachusetts, Labov (1972b: 9) examined differences in the height of the first element of the diphthongs [aɪ] and [aʊ] in the PRICE and MOUTH vowels. He states:

Instead of the common southeast New England standard [aɪ] and [aʊ], one frequently hears on Martha’s Vineyard [ɐɪ] and [ɐʊ], or even [əɪ] and [əʊ]. This feature of centralized diphthongs is salient for the linguist, but not for most speakers; it is apparently quite immune to conscious distortion, as the native Vineyarders are not aware of it, nor are they able to control it consciously.

(Labov 1972b: 9)

Degree of centralisation was found to vary with speaker age, with centralisation of (ay) and (aw) showing a ‘regular increase in successive age levels, reaching a peak in the 31 to 45 group’ (Labov 1972b: 21). Labov was able to explain this increase by examining the social situation of islanders. In addition to age, centralisation was found to vary with speakers’ attitudes to the summer people who flocked every year to the island to holiday, with high centralisation correlating closely with ‘expressions of strong resistance to the incursions’ of these tourists (Labov 1972b: 28). Labov states:

It is apparent that the immediate meaning of this phonetic feature is ‘Vineyarder.’ When a man says [rɛɪt] or [hɛʊs], he is unconsciously

establishing the fact that he belongs to the island: that he is one of the natives to whom the island really belongs.

(Labov 1972b: 36)

Whaling, fishing, and farming were all in decline on Martha's Vineyard, and the island's economy was dependent to a large extent on the seasonal income from the tourist trade. Many of the inhabitants of Martha's Vineyard were experiencing unemployment and poverty. Centralisation was highest amongst those between 31 and 45 as it was this group which had been most affected by the declining economy and influx of tourists, and who therefore felt the need to protect their identity as 'Vineyarders', linguistically and otherwise, most strongly. Centralisation was also on the increase among younger people planning to remain on the island. Labov's study shows clearly a change in progress on Martha's Vineyard and the co-variation between age and identity. The matter of identity is discussed further in section 2.4 and in chapter 7.

- Trudgill on Norwich – (e)

In Trudgill's survey of Norwich English, (e), the DRESS vowel was found to have three main variants: [ɛ]; [ɜ]; and [ʌ], with the most extreme Norwich pronunciation of *hell*, for example, being identical with *hull* (Chambers and Trudgill 1998: 80).

Chambers and Trudgill (1998: 80) state that use of [ʌ] 'is very much on the increase, with speakers under thirty showing a very high level of centralised vowels', and argue that 'a linguistic change is currently taking place'.

2.2.1.3 Change in Progress – Variant on the Decline

- Trudgill on Norwich – (ir)

This variable is the NURSE vowel, and in Norwich was found to vary from the RP-like [ɜ:] to local pronunciations such as [ɛ:], [a:], and [ɐ] (Chambers and Trudgill

1998: 81). These last pronunciations are, however, dying out, being found only in the speech of older speakers. Chambers and Trudgill (1998: 81) state that 'it will not be too long before this particular relic form has disappeared altogether'.

2.3 The Speaker Variable of Social Class

The importance of social class in society, and the influence of class on language use, are widely acknowledged by linguists. Macaulay states:

[I]f linguists aspire to social relevance, they cannot avoid a concept so crucial to society as that of class, however awkward it may seem to them.

(Macaulay 1976: 187)

Milroy and Gordon concur:

[S]ocial class is a variable which plays so prominent a role in language variation, at least in industrialized countries, that a socially accountable researcher cannot avoid considering it at least at some level of the analysis.

(Milroy and Gordon 2003: 40)

Despite its significance, however, the variable of social class is a highly problematic one, perhaps more so than other speaker variables in social dialectology. Chambers (2003: 41) states that '[t]he notion of social class is inherently fuzzy'. Though well aware that the social significance of the speaker variables of age and sex varies between societies, Chambers (2003: 41-42) argues that these speaker variables, unlike social class, are at least easy to determine; a person is either male or female and is a particular age. Social class, on the other hand, is far harder to define. Lesley Milroy (1987b: 29) states that class is 'a variable which is [...] obviously relevant to language variation in a modern western urban community', yet acknowledges that it has also been a variable which has presented, and continues to present, 'pervasive difficulties' in both definition and interpretation (Milroy 1987b: 97).

Milroy (1987b: 97) comments that linguists tend 'to use the concept of social class rather unreflectingly, primarily as a means of imposing some order on variable linguistic data'. Her concern is that linguists do not define what they mean by social class. To do so, she acknowledges, is by no means a simple thing. Of social class as it is used in stratificational studies, Milroy (1987b: 101) concludes that it is a '*proxy* variable covering distinctions in life-style, attitude and belief, as well as differential access to wealth, power and prestige [original italics]'

2.3.1 Empirical Approaches to the Speaker Variable of Social Class

Some of the approaches taken by linguists to the speaker variable of social class are detailed in the following sections. These approaches are discussed in relation to the selection of a means of determining social class for the Southampton area study in section 3.9.

2.3.1.1 Labov on New York City

- **The Social Stratification of (r) in New York City Department Stores**

As part of his survey of the social stratification of English in New York City, Labov investigated rhoticity in the speech of sales assistants in three department stores of differing statuses in New York. He hypothesised that 'if any two sub-groups of New York City speakers are ranked in a scale of social stratification, then they will be ranked in the same order by their differential use of (r)' (Labov 1966: 64). Though Labov could have tested this theory by comparing the speech of people from two different occupational groups, such groups being 'among the most important indexes of social stratification', he opted instead to focus on a single occupational group in which there existed 'a subtle case of stratification' (Labov 1966: 64). Labov's (1966: 73-74) hypothesis was correct, and a strong correlation was apparent between

use of (r), and social class as defined by status within a particular occupational group.

Labov's department store survey emphasises the evaluative element of social class. Though all informants in this part of his New York study were shop assistants, and would, therefore, have been ranked equally on a purely economic scale, the way in which they were perceived by others, and indeed perceived themselves, as a result of the relative status of their places of work differed, highlighting the fact that 'social stratification is the product of social differentiation and social evaluation' (Labov 1966: 63).

- The Survey of the Lower East Side

Labov's 'principal device' (Labov 1966: 154) for the study of the social stratification of English in New York City was a survey of the Lower East Side. A social class index (SCI) based on that used by the Mobilization for Youth Program, which had taken place in the area shortly before Labov began his research, was employed to stratify informants (Labov 1966: 211). Milroy states:

It is [the] grouping of persons into loosely defined bodies, intersubjectively perceived as occupying positions relative to each other, which linguists have tried to capture quantitatively by means of a *social class index score* [original italics].

(Milroy 1987b: 30)

Three indicators of social class were selected: occupation; education; and income (Labov 1966: 213). Though Labov (1966: 212-13) states that '[a] single indicator, such as occupation or education, might have been used for the social class index', the decision to use three equally-weighted indicators was 'based on considerations of accuracy and reliability' (multiple-item indices are discussed further in section 2.3.1.2). In the case of occupation, informants were ranked according to the occupation of the main wage earner in their family (Labov 1966: 213). However, in

the case of education, the education of the individual was used to ascertain their position in the social class index (Labov 1966: 214). Ranking according to income took into consideration total family income and the number of adults and children in the family (Labov 1966: 214-15). Based on these three factors, individuals were placed on a 10-point scale. This scale was then divided into class groupings according to speaker use of individual linguistic variables, on the grounds that ‘language is a measure of class behavior’ (Labov 1966: 237) (the division of informants into social class groups according to their language use is discussed further in section 2.3.1.2). Though a different approach to determining social class than the one employed in the New York department stores survey, this method was also successful in highlighting class-based differences in language use.

2.3.1.2 Trudgill on Norwich

In his original Norwich survey, Trudgill (1974) employed a multiple-item index of social class in order that ‘co-variation between linguistic behaviour and social status could be studied accurately’ (Trudgill 1974: 35). Like Labov (1966: 64), Trudgill (1974: 36) emphasises the value of occupation in determining social class, stating that ‘occupation is probably the most important stratifying element in British society’. Despite this, Trudgill employed six indicators for his social class index, these being: occupation; income; education; housing; locality; and father’s occupation. He argues for the use of such an index on the grounds of its objectivity, stating that:

by increasing the number of indicators of social class involved, [a multiple-item index] is a much more refined and reliable means of measuring social class [than a single-item index].

(Trudgill 1974: 36)

The objectivity of this scale is in fact questionable, by Trudgill's own admission. His ranking of the localities in Norwich, one of the six indicators used in the SCI, was anything but objective. Trudgill states:

The social status of different localities within a city can to some extent be measured objectively [...]. In this work, however, no attempt was made to achieve this kind of measurement. Instead, the different areas investigated in the survey were ranked subjectively – a much quicker and simpler process. This ranking was based on knowledge, acquired during many

years' residence in the city, of the status significance of different neighbourhoods.

(Trudgill 1974: 40)

Milroy (1987b: 31) also questions the objectivity of SCIs, arguing that 'arbitrariness is one of the principal problems in the use of a social class index' as 'different indicators are perceived as important by different investigators; moreover, perception of their *relative* importance seems to vary [original italics]'. This is certainly true if Trudgill's SCI, which employs six indicators, is compared to those used by Labov in the Lower East Side (see section 2.3.1.1) and Macaulay in Glasgow (see section 2.3.1.3), which feature three indicators and one indicator respectively. Macaulay also expresses unease regarding the use of multiple-item SCIs:

In the absence of evidence by which occupational, educational, and other factors can be ranked on equal interval scales, the degree of distortion produced through adding them to each other is totally unpredictable. [...] In other words, it should not be taken as self-evident that the use of several indicators necessarily provides a more accurate measure of social stratification than the use of a single one.

(Macaulay 1976: 185)

This is a view echoed by Chambers, who argues that:

[w]hen several class indicators are used, each one increases the fuzziness of the individual index [...]; the fuzzier the scores – or, in other words, the less discrete the class continuum – the vaguer the correlations.

(Chambers 2003: 53)

Trudgill has also been criticised regarding the methodology used to divide the scale created by the SCI into discrete groups for the purposes of linguistic analysis. For each of the six indicators employed by Trudgill, informants were awarded a score of between 0 and 5, resulting in possible social class index scores of between 0 and 30 (Trudgill 1974: 38). The continuum produced by the SCI was divided into five discrete class groupings according to informants' use of the third person singular marker (Trudgill 1974: 55-61). Trudgill was not, however, the first linguist to use linguistic data to divide a social stratification scale into discrete classes or groupings; Labov (1966: 237) used this technique in his survey of the Lower East Side (see section 2.3.1.1). Of the use of this technique by these researchers and others, Macaulay states:

This [...] approach ensures that the groupings will be to a certain degree linguistically homogeneous, but unfortunately it does not guarantee that the resultant groupings will correspond to sociologically meaningful divisions of the social stratification continuum (though it does not preclude this either).

(Macaulay 1976: 185)

Similarly, to divide a social class scale into groups according to linguistic data for the purpose of analysing the relationship between social class and the use of particular linguistic variables is self-referential, and might not, therefore, be very reliable. In order to overcome such problems, Macaulay (1976: 186) argues that '[i]n dividing up the social stratification continuum into social classes, the sociolinguist [...] ought to seek extra-linguistic evidence to justify any classification based primarily on the linguistic analysis'.

Despite these criticisms, and those levelled by feminists regarding the ranking of female informants according to their husbands' or fathers' occupation and income (see section 2.1.2), Trudgill's SCI was nonetheless successful in highlighting social class differentiation in the use of the linguistic features examined.

2.3.1.3 Macaulay on Glasgow

Macaulay's (1977) Glasgow study was based on three sets of interviews: a 'Community Sample' designed to collect a wide range of Glasgow speech; a 'Teachers Sample' designed to investigate teachers' attitudes to language and the language of their pupils; and an 'Employers Sample' to investigate the attention paid to language when choosing among job applicants (Macaulay 1977: 18). Informants for the 'Community Sample' of the survey were selected using occupation as the criterion for social class membership (Macaulay 1977: 18). Aware of the controversy surrounding class as a variable in linguistic surveys, and of the variety of ways of determining an informant's class, Macaulay argues for the use of occupation alone for a variety of reasons, both pragmatic and theoretical. He states that 'information about occupation is relatively easy to obtain, and it is not as potentially embarrassing a topic for informants as questions about family income' (Macaulay 1977: 57). Use of occupation as the sole indicator of social class also facilitated the selection of informants, as details of fathers' occupations were kept on school records, making it quick for Macaulay to select a balanced sample for his study (Macaulay 1977: 57). He also acknowledges that there existed at the time no demographic study of Glasgow which would have allowed for the objective weighting of other factors that might have functioned as indicators of social class (Macaulay 1977: 57). Perhaps the key justification for Macaulay's decision for the use of occupation alone, however, is the fact that it is 'generally regarded as the best *single* indicator of social class [original italics]' (Macaulay 1977: 57).

Macaulay (1977: 18) employed a quota sample 'to provide an equal representation of both sexes from three age-groups in four social class categories [...] based on the Registrar-General's classification of occupations'. Children were

selected on the basis of their father's occupation, and married women on the basis of their husband's occupation, unless the woman was or had been in an occupation higher up the Registrar-General's scale than her husband's. in which case her own ranking was used (Macaulay 1977: 18).

As part of the community sample, Macaulay asked his adult informants four questions concerning social class:

People often talk about there being different (social) classes – do you think this is true in Glasgow?

How many classes would you say there were in Glasgow?

Where would you put yourself in a class system?

What is it that determines which class you belong to – family, education, money, or job?

(Macaulay 1977: 60-62)

The informants' responses were varied but indicate awareness among respondents of their own social position and provide, according to Macaulay (1977: 64-65), support for the use of occupation as the sole indicator of social class. While Macaulay's study is relatively small, focusing as it does on sixteen adults and thirty-two children (Macaulay 1977: 19), his results show a strong correlation between the five phonological variables under analysis and social class distinctions as determined by occupation. When Macaulay (1977: 59) ranked his adult informants by their linguistic indices for the four vowel variants investigated, there was 'not a single individual out of place according to the order predicted on the basis of occupation'. Among the children, there were only four anomalies in the rankings.

Though Chambers (2003: 53) argues that Macaulay's Glasgow survey 'offers empirical support for simplifying class as an independent variable in sociolinguistic studies by using occupation as the sole indicator', he does state that this methodology 'must obviously be tested in other cities'. There is the possibility that there exists some feature peculiar to the social makeup of Glasgow which permits the type of

correlation between linguistic features and social class defined by occupation found by Macaulay. Macaulay (1977: 65) states that ‘Glasgow’s economy has been stagnant or declining since the end of World War I [...] and this may have contributed to reduced social mobility’, which in turn might have made social stratification less problematic. He also highlights the fact that the average age of his adult informants was 48, ‘so it is reasonable to suppose that they had all reached a stage where their occupational status was unlikely to change radically’ (Macaulay 1977: 65). Macaulay suggests that the inclusion of adults aged between 20 and 30 might have ‘complicated’ the use of occupation as the sole indicator of social class. Though Macaulay (1977: 65) himself acknowledges that in a larger sample, ‘the use of additional information on income, education, type of residence, etc., might have proved useful in helping to determine social class membership’, the use of occupation alone to determine social class proved to be highly successful in Glasgow.

2.3.1.4 Milroy and Milroy on Belfast

In their survey of three Belfast neighbourhoods, James and Lesley Milroy relied on their intuitions when determining the social class of their informants. They state:

The project investigates three communities in the inner city [...]. All three are decayed ‘core working-class’ areas with a high incidence of unemployment and other kinds of social malaise.

(Milroy and Milroy 1978: 21)

It was assumed that all the subjects for the Belfast study were working class, as the areas in question were supposedly inhabited solely by working class people.

On the kind of judgement sampling employed by the Milroys, Chambers states:

Intuitions about the class membership of individuals are reliable only under the most favorable conditions, as when one is judging prototypical

individuals rather than fringe members, or is intimately familiar with the community, or is dealing with homogeneous neighborhoods. and so on.
(Chambers 2003: 47)

Milroy herself argues:

Certainly, in view of the problems associated with strict representative sampling, it may be more realistic for researchers conducting, for example, an urban dialect survey, to judgement sample *on the basis of specifiable and defensible principles* than to aim for true representativeness [original italics].

(Milroy 1987b: 28)

It is presumably the Milroys' intimate knowledge of the areas under investigation that formed the basis of the 'specifiable and defensible principles' on which their own survey was based. Judgement samples have also been used by other linguists, including Chambers (2003: 44) in his examination of the North Toronto neighbourhood in Canada, and Tollfree (1999: 163-164) in her survey of South East London English. Indeed Trudgill's (1974) original Norwich survey employs judgement sampling in ranking localities in Norwich for the purpose of his SCI (see section 2.3.1.2). Chambers (2003: 44) argues that selecting informants from 'well-defined neighborhoods [...] carries few risks', provided that researchers have a good knowledge of the sampling area.

2.3.1.5 Eckert on Belten High School, Detroit

Linguists such as Labov and Trudgill who, consciously or unconsciously, adopt a functionalist view of social class, see it as being consensual, based 'on the values *shared* within and between classes [original italics]' (Milroy 1987b: 99). However, the jock-burnout opposition studied by Eckert in Belten High, Detroit, 'foregrounds conflict models of social class' (Eckert 2000: 16). Eckert states:

The jocks and burnouts constitute middle class and working class cultures respectively – they are the instantiation of class in the adolescent life stage, and serve as trajectories to adulthood. Representing opposing orientation to school and to the local area, the

jocks are an institutional, corporate culture while the burnouts are a personal, locally oriented culture.

(Eckert 2000: 2-3)

In her exploration of jocks and burnouts, Eckert raises a number of interesting points regarding social class. She draws attention to the way in which 'measures of socioeconomic class are based on school-leaving, employment, adult consumption patterns [...], and adult life style', meaning that 'social theorizing related to variation is most complete for this age group' (Eckert 2000: 7). Eckert's data call into question sampling techniques which categorise young informants according to their parents' occupation or education. The Belten High data show the 'small extent' to which the speech of the students reflects the social class of their parents (Eckert 2000: 108), raising the question as to how best to categorise younger informants according to social class. Eckert (2000: 21) argues that in employing SCIs based primarily on occupation and education, or in using occupation as the sole indicator of social class, it is the 'degree of engagement in the standard language market' that is being assessed. A conflict view of social class, however, suggests 'that there are alternative linguistic markets, within which forms other than the global standard constitute the norm' (Eckert 2000: 18).

For her Belten High study, Eckert attempts to overcome some of the problems she details, by basing her choice of speakers 'on place in social networks and particularly in relation to social categories, foregrounding the speakers' own socioeconomically related choices' (Eckert 2000: 111). She examines her informants' language use in relation to the groups to which they themselves have chosen to belong. The attitudes, fashions and hobbies of these groups represent typically working or middle class behaviours. Speakers are not assessed on their parents' occupations or level of education, but on their own choices and behaviour.

Eckert's study is, however, an ethnographic one (Eckert 2000: 2), and her means of assessing the social class of informants might not, therefore, be replicable by researchers who study speakers in relative isolation, rather than as part of the wider communities to which they belong.

2.4 The Speaker Variable of Identity

2.4.1 Gal on Oberwart, Austria

Gal's (1978, 1979) study of the Hungarian-German bilingual town of Oberwart, Austria, shows the far-reaching effects of identity on language use. When the language choices of members of the Oberwart community were examined, Gal (1978: 13) observed that young females were leading a change towards predominant use of German. They did not identify with the lifestyle, a peasant agricultural existence, associated with the speaking of Hungarian, but instead aligned themselves with the worker lifestyle associated with German. Young women's rejection of 'the social identity of peasant wife' (Gal 1978: 14) resulted not only in the need for peasant men in Oberwart to look outside their own community to the surrounding German monolingual towns to find a wife, but also, by extension, in the demise of Hungarian in Oberwart (Gal 1978: 13-14). Gal states:

Because the children of marriages between monolingual German speakers and bilingual Hungarian-German speakers in Oberwart rarely if ever learn Hungarian, in an indirect way the present generation of young women is limiting the language possibilities of the next generation.
(Gal 1978: 14)

Gal's study shows not only the interaction between identity and speaker sex, but also the far-reaching consequences of identity-based linguistic choices.

2.4.2 Trudgill on British Pop Song Pronunciation

Trudgill (1983) has applied Le Page and Tabouret-Keller's Acts of Identity Theory (see section 1.5) to the pronunciation of British pop songs. British pop singers,

Trudgill (1983: 144) argues, were attempting to modify their pronunciation in order to be identified with Americans. He suggests that this was because Americans dominated twentieth-century popular music, and such domination led to imitation.

The first of the four riders to the Acts of Identity Theory is that an individual's ability to modify their verbal behaviour so as to resemble the group with which they wish to be identified is dependent on the extent to which they are able to identify the said group (Le Page and Tabouret-Keller 1985: 182). Trudgill (1983: 145-46) states that the British pop singers were not entirely successful in their language modification as they had not identified exactly which Americans they were trying to copy. While attempting to imitate Southern and/or Black singers, British pop singers adopted rhotic pronunciations, when the speech of their target group(s) was typically non-rhotic (Trudgill 1983: 146-47).

The second of Le Page and Tabouret-Keller's (1985: 182) riders states that it is necessary to have adequate access to the group with which one wishes to be identified, and the ability to analyse their behaviour. The British pop singers are judged by Trudgill (1983: 148) to have been unsuccessful in analysing the behaviour of the group(s) of Americans with which they wished to be identified. Contact with speakers of American English made no difference, since the Americans were 'unlikely to offer any overt, accurate correction' (Trudgill 1983: 149).

The third Acts of Identity Theory rider (Le Page and Tabouret-Keller 1985: 182) is that the individual modifying their linguistic behaviour must have sufficiently powerful motivation to join the group they are attempting to imitate. Analysing use

of non-prevocalic /r/ and voiced alveolar flaps for intervocalic /t/ in the songs of The Beatles and The Rolling Stones, Trudgill (1983: 150-54) shows that use of these typically American forms declined from 1964 onwards, and that this was the result of a diminishing motivation on the part of these groups to sound American. One possible explanation for this diminishing motivation could have been a change in the genre of the music played, in the case of The Beatles, a move from rock-and-roll to more poetic lyrics dealing with British themes and places (Trudgill 1983: 153). Alternatively, it could have been a result of the success of British groups both at home and abroad, which ‘led to a change in the pattern of cultural domination’ (Trudgill 1983: 153).

The fourth and final rider to Le Page and Tabouret-Keller’s (1983: 182) Acts of Identity Theory hinges on the ability of the speaker to modify their behaviour. Trudgill (1983: 149) states that the majority of the modifications in pronunciation made by British singers were ‘variable, irregular, and inconsistent’, and suggests that they lacked the ability to imitate successfully the group(s) with which they wished to be identified.

2.4.3 Dyer on Corby, Northamptonshire

The matter of speaker identity is central to Dyer’s (2002) work on Corby, a town in Northamptonshire which had a large influx of inhabitants from Scotland between the 1930s and the 1970s. Linguistic data were collected from three generations of Corby inhabitants (Dyer 2002: 102). Of the eight first generation speakers, the oldest group in the study, half were born in England, and half in Scotland. Only one of the eight second generation speakers was Scottish-born, and all of the eleven third generation speakers were English-born.

Analysis of the LOT/THOUGHT lexical sets, which in Scottish English often have a single common phoneme but which in Anglo-English are distinguished between, revealed a tendency in the speech of the second and third generations to favour the Anglo-English norm (Dyer 2002: 105-106). Dyer (2002: 106) argues that these results 'suggest that the younger speakers are avoiding distinctly Scottish features in their speech and adopting established Anglo-English features'. Analysis of the GOAT lexical set, however, did not support this assertion, as the historically Scottish monophthongal variant was found to be in use by some of the younger English-born Corby inhabitants (Dyer 2002: 106). While it is possible that the use by third generation speakers of an historically Scottish feature could be the result of their identification with Scotland, Dyer (2002:112) instead argues that speakers' employment of the monophthongal variant signals their wish to distinguish themselves from the inhabitants of nearby Kettering. She states that in Corby 'there has been a shift from orientation towards ethnic group to orientation towards town community' (Dyer 2002: 112).

Chapter 3

3.0 The Southampton Area Study Methodology

This chapter provides an overview of the Survey of Regional English methodology, as it has been used in the present study. It details how this methodology has been refined and adapted, and examines the issues arising from the use of such a data elicitation technique. The chapter also details the ways in which linguistic variables have been selected for detailed analysis, and the means by which this analysis is undertaken.

3.1 The Survey of Regional English

Whilst acknowledging the very great importance of existing network surveys such as the Survey of English Dialects (Orton et al. 1962-71), the Survey of Anglo-Welsh Dialects (Parry 1977, 1979), the Linguistic Survey of Scotland (Mather et al. 1975-86), and the Tape-Recorded Survey of Hiberno-English Speech (Barry 1981), Kerswill et al. (1999: 258) point out that these surveys ‘are of different times, have different aims, and employ different methodologies’. The surveys were conducted some time ago, and ‘do not amount to a cohesive record of the dialects’ (Kerswill et al. 1999: 258). As such, a deficiency exists in our knowledge of contemporary English language variation in the British Isles. The Survey of Regional English (SuRE) (Upton and Llamas 1999) seeks to rectify this situation. Of the Survey

Llamas states:

The basic intention of the SuRE project is to create a computer-held database of consistently-collected material from a planned network of British localities which will record and document the facts of linguistic variation throughout Britain, permitting detailed analyses of issues concerning the diffusion of language change and the spread of current vernacular changes in British English.

(Llamas 1999: 96)

A joint undertaking by the Universities of Leeds and Sheffield, its aim is to create a lasting database documenting British English dialects as they exist at the turn of the twenty-first century and beyond.

3.2 The Survey of Regional English Methodology

Kerswill et al. state:

In order to undertake a large-scale survey of regional variation in contemporary spoken British English, data must be obtained which can be analysed on three levels of possible variation; phonological, grammatical and lexical. Although it is difficult to combine the three levels, to discount any would be to obtain an incomplete picture of the regional linguistic variation found in the British Isles at the turn of the Millennium. The phonological, grammatical and lexical data must be comparable across the localities to be studied, permitting quantitative analyses of the different levels of regional and social variation.

(Kerswill et al. 1999: 260-61)

Given the aims of SuRE, and in light of the limited time and economic resources available to most fieldworkers for data collection, the use of data elicitation techniques employed in existing surveys was ruled out. Whilst allowing free conversation between informants (cf. Docherty et al. 1997; Llamas 1998) or eliciting personal narratives (cf. Labov 1972b) would produce large amounts of speech for analysis, 'the possibility of obtaining any comparable data on lexical variation would be almost completely removed' (Kerswill et al. 1999: 261). Similarly, the use of a Survey of English Dialects-style questionnaire to elicit specific lexical and grammatical data was also deemed to be inappropriate, on the grounds that it was too formal and time-consuming (Kerswill et al. 1999: 261).

With these considerations in mind, a new methodology was developed which was first used by Llamas (2001) in her study of Middlesbrough English. In light of the need for greater levels of comparability between social dialectological studies, and since 'knowledge of current regional and social lexical variation in the British

Isles is extremely sparse', the SuRE methodology employs a set of core, and therefore comparable, notion words and phrases as a means by which to elicit data (Llamas 1999: 95-98). The principal tools of the methodology are three Sense Relation Network sheets (SRNs) on which 'standard notion words are offered as prompts for the elicitation of dialectal variants' (Llamas 1999: 102). Based on Aitchison's (1994) idea that the mental lexicon is comprised of a series of interconnected 'word-webs' or 'semantic networks', the 'SRNs are built around domains of language, and in this regard are akin to the grouping of questions by subject matter in the SED questionnaire' (Upton and Llamas 1999: 299). When the SRNs have been completed, the responses provided by informants are then discussed in an interview. Llamas states:

In order to obtain the required informal speech style combined with data on lexical variation in the interview, the fieldworker 'leads' a conversation around semantic fields. To lessen the formality of the interview context, the interview is undertaken with socially paired informants, permitting interaction to be more like a conversation than an interview.

(Llamas 1999: 98)

In this way, the simultaneous collection of comparatively casual-style lexical, phonological and grammatical data is possible.

The original core Survey of Regional English methodology comprised the following:

- An untitled sheet for informants headed by a section asking for name, place of birth, and details of other places lived in and for how long. The sheet also provided instructions as to how to complete the Sense Relation Network sheets.
- A sheet entitled *Biographical Information* which asked informants for information such as sex, age, assessment of social class, and education.

- Three Sense Relation Networks (SRNs) entitled: *Feelings, Actions and States: The Outside World*; and *People*.

Copies of these sheets can be found in Appendix 1.

3.3 Refinement of the Original Survey of Regional English Methodology

I began work with Esther Asprey, University of Leeds, and Lourdes Burbano Elizondo, University of Sheffield, in May 2002 to evaluate and improve the existing SuRE data collection materials. Alterations were made to the methodology for a variety of reasons, one being to ensure clarity and ease of use for informants. The changes eradicate elements of the methodology deemed by Esther Asprey, Lourdes Burbano Elizondo and me to be unproductive or superfluous, and ensure that the SuRE project conforms to current legal requirements arising from the UK Data Protection Act 1998 (Information Commissioner 1998).

3.3.1 Alterations Made to the Untitled Instructions Sheet

3.3.1.1 Repetition between the Instructions Sheet and the *Biographical Information Sheet*

It was apparent that there was repetition between the *Biographical Information* sheet and the instructions sheet, informants being asked for their name and place of birth on both of these pages. As requesting this information twice was superfluous, it was decided that these details would be removed from the instructions sheet. It was also decided that the question *other places lived and for how long*, requested on the instructions sheet, would be more appropriately included on the *Biographical Information* sheet. In this way, there would be a clear, logical distinction between the two pages.

3.3.1.2 An Addition to the Instructions Sheet – *What is SuRE?*

Out of courtesy to participants giving up their time to help, the new instructions sheet includes an opening paragraph which provides information about the SuRE project.

This paragraph details the locations in which research has been and is currently being carried out, and describes to informants the aims of the Survey.

3.3.1.3 Changes to the Wording of the Instructions Sheet

Several changes were made to the wording of the original instructions sheet. A few of these changes were purely stylistic. Others were made for reasons of clarity or practicality, for ideological reasons, or to reflect changes to the core methodology as a whole.

The first directive to informants on the original instructions sheet read *please complete the [SRN] sheets with words you think are **dialect** words or are local to the area you are from* [original emphasis]. It was felt that to introduce the concept ‘dialect’ might result in informants recording only lexis believed to be traditional dialect in the layman’s sense of the term, that is the language used by older, especially rural, speakers. Since an aim of SuRE is to record all examples of lexical variation, this instruction was altered to read *please complete the sheets with words you think are local to the area you live in*. The change in wording from *local to the area you are from* to *local to the area in which you live* was made to ensure that informants who had moved to an area under investigation completed the SRNs with words from that area, rather than from the one in which they had been born or brought up. However, this is not an issue in the Southampton area survey for reasons discussed in section 3.6.

Since it had been decided to remove the ‘any others’ sections from the SRNs (see 3.3.3.3 below), the information concerning how to complete these was also

removed from the instructions sheet. The reference to the *Identification Questionnaire* used by Llamas (2001) in her Middlesbrough survey was also removed, since this questionnaire was not a core element of the SuRE methodology, but an add-on, designed for use in that particular area.

It was decided that the instruction to the informant on the original sheet to *put down more than one word, if you like* was not sufficiently forceful, and was reworded to read *give as many examples as you can*. To discourage informants from leaving blank spaces next to words for which they could not think of any alternatives, a rather dispiriting experience, the following request was added: *if the only word you can think of is the one given please write that down. This is important information in its own right*. This, it was hoped, would reassure informants that all responses were of value.

Lastly, a note to informants was included, explaining that a short discussion about the words and phrases they had thought of would follow the completion of the Sense Relation Network (SRN) sheets. Participants are directed to the *Confidentiality and Consent* sheet (discussed in section 3.3.4 below) for more information about this.

3.3.1.4 Aesthetic Alterations to the Instructions Sheet

A logo for the SuRE project was designed which now appears at the top of the instructions sheet, this being the first page of the revised core methodology. Clear sub-headings and a page border were also added to the sheet. Though only minor alterations, these changes help to ensure clarity.

3.3.2 Alterations Made to the *Biographical Information Sheet*

3.3.2.1 Title

The original title *Biographical Information* was deemed to be overly formal, given that access to the vernacular is the aim of the Survey, and so was changed to *About You*. It was felt that responses given, both when completing the sheets and later in the interview, might be adversely affected by the use of very formal language in the sheets to be completed by informants. The concern was that informants might feel the need to respond with formal language not representative of their usual speech. It was also hoped that by using relatively informal English at all times, the methodology would be accessible to a wide range of informants.

3.3.2.2 Name

It was decided that informants should not be asked to give their name on the *About You* sheet. Those taking part in the study would be asked to complete a *Confidentiality and Consent* form (see section 3.3.4 below), requiring them to sign and print their name, making it unnecessary to ask for this information on the *Biographical Information* sheet as well.

3.3.2.3 Sex

Since the Southampton area survey is concerned in part with the potential influence of the speaker variable of sex on language, it was decided that the revised sheet, like Llamas's original, should ask all informants for their sex. This question therefore remains unaltered.

3.3.2.4 Age (dob)

Like the speaker variable of sex, speaker age has been shown to influence speech significantly (Labov 1972b; Trudgill 1974; Macaulay 1977; Milroy 1987a). In light

of this, it was decided that data concerning informant age should continue to be collected. However, it was felt that overtly asking the age of informants, particularly older female informants, might be considered impolite and result in a degree of ill will in any subsequent interviews. Balancing this consideration against the need of fieldworkers for details of informants' ages so that speech obtained could be analysed for signs of age-grading or changes in progress, it was decided that asking informants indirectly for their age would be best. The *About You* sheet was therefore altered from *age (dob)* to read *date of birth*. The abbreviation *dob* employed by Llamas was not used since it might not have been immediately apparent to some informants what details were being requested. Clarity was again of paramount importance here.

3.3.2.5 Place of birth

This section of the original sheet was retained on the revised *About You* sheet, to facilitate the compilation of a comprehensive speaker biography. In addition to asking informants for their place of birth, it was also decided to ask for their *current place of residence*. The collection of these data was of particular importance in the Southampton area survey given the amalgamation of the city of Southampton and the neighbouring borough of Eastleigh to form a single sampling universe (section 1.2). It was deemed vital to record the precise location of informants, in order that possible differences between the speech of informants from Eastleigh Borough and those from the city of Southampton, perhaps relating to issues of social class, could be easily investigated.

The question *other places you have lived and for how long*, found on the original instructions sheet, was grouped with *place of birth*, and *other places you have lived and for how long*, as all are concerned with informants' geographical

details. Given the geographically mobile nature of contemporary society, it is to be expected that at least some speakers will have lived in more than one area. Data are requested on other places they have lived, in order that other possible influences on informants' speech might be ascertained.

3.3.2.6 Birthplace of mother and birthplace of father

These questions are retained in the revised sheet. Collecting this information was felt to be important, as a parent's place of birth has been shown in some cases to influence the speech of their children. One example of this is Payne's (1980) investigation in Philadelphia's King of Prussia suburb. The investigation concerns the extent to which children 'acquire the phonological system of a second dialect after moving from one dialect region to a new one' (Payne 1980: 143-144). Payne states that:

[u]nless a child's parents are locally born and raised, the possibility of his acquiring the short-*a* pattern is extremely slight even if he were to be born and raised in King of Prussia.

(Payne 1980: 174)

Similarly, Trudgill states:

[M]y current research in Norwich shows that even adults who have lived all their lives in Norwich and who otherwise have perfect Norwich accents may not, if their parents were from some other dialect area, have mastered successfully the Norwich distinction between /u:/ and /ʌu/ [...]:

/u:/	/ʌu/
moan	mown
nose	knows, etc.

Even after a lifetime's exposure to this distinction, they neither produce it nor imitate it correctly.

(Trudgill 1983: 13-14)

Data concerning the birthplace of informants' mothers and fathers could help to explain differences in the speech of informants from the same area.

3.3.2.7 Birthplace of grandmothers and birthplace of grandfathers

Original questions asking informants for the birthplace of their grandparents were omitted from the revised *About You* sheet. Dyer's (2002) research in Corby, a town in Northamptonshire which saw a large influx of inhabitants from Scotland between the 1930s and the 1970s, indicates that the collection of data on grandparents' place of birth might be superfluous. Dyer states:

[T]he working hypothesis that a family [in Corby] with first generation Scottish-English speakers would be more likely to have grandchildren who spoke a dialect with Scottish features, than a family with an English-born first generation, turned out to be false, since the ethnicity of the grandparents appeared to have little influence on the resulting dialect of the grandchildren.

(Dyer 2002: 102)

In light of this, it was felt that data collected on the birthplace of parents would be sufficient for any analysis of family as a linguistic influence.

3.3.2.8 Ethnic group

In cases such as those described by Labov et al. (1968) in New York, ethnicity has been shown to influence language use. Llamas's question asking for details of an informant's ethnicity was therefore retained on the revised *About You* sheet. The possibility of having boxes such as *white* or *black*, for example, for informants to tick was rejected as being restrictive. The way in which people choose to define their own ethnicity and how this might be reflected in their speech was felt to be potentially far more telling. Milroy (1987b: 104) argues that 'ethnicity is a culturally created category, in no sense objectively 'given' or verifiable [original quotation marks]'. She states:

Ethnicity, like social class, presents problems of definition, but can reasonably be described as an individual's sense of belonging to a distinctive group whose members share a common history and culture.

(Milroy 1987b: 103)

While not a speaker variable specifically under examination in the Southampton area survey, the collection of data concerning a person's ethnicity is justified in the interests of creating a useful database.

3.3.2.9 Occupation (current/usual)

Social class has been shown in many studies to influence language use (Labov 1966; Trudgill 1974; Macaulay 1977), and occupation has been employed in many such studies as an important indicator of social class (Labov 1966; Trudgill 1974; Macaulay 1977). Similarly, occupation is used in the National Statistics Socioeconomic Classification, and its predecessor the Registrar-General's Social Class, to determine a person's social class for official purposes (Office for National Statistics 2005: 2-3). For these reasons it was decided that the original question concerning occupation should be retained on the *About You* sheet. *Occupation (current/usual)* was, however, amended on the revised sheet to read *usual occupation(s)*. This was firstly to allow informants with more than one occupation, for example students with jobs, to give details of both or all of these. Like Llamas's wording, the revised question also allows informants doing a job different to their usual occupation at the time of the interview to give details of their normal employment, and retired informants to state their previous employment. At the risk of embarrassing or offending informants by asking overly personal questions, *current occupation* was removed so that those unemployed at the time of the interview would not be required to disclose this information.

The question on occupation was also expanded to ask informants for *mother's usual occupations(s)* and *father's usual occupation(s)*. On the matter of parents' occupation, Trudgill states:

Father's occupation is [...] clearly of some importance in assessing social class, since one is initially born into the particular social class group which is that of one's parents. This is particularly important from a linguistic point of view, since social mobility is known to have an effect on linguistic behaviour.

(Trudgill 1974: 37)

Trudgill's comments were made over thirty years ago, and it is not the case that all families take their position from the social status of the father, hence the decision to ask informants for the occupation(s) of both their father and mother. These additions had two purposes. In the case of younger informants who might still be in full-time education, information concerning social class could be gleaned from an assessment of the occupation of their parents. Secondly, these questions would allow the fieldworker to distinguish instances in which an informant might have changed social class from that which they had when they were a child.

3.3.2.10 Assessment of social class

Whilst occupation alone is used as a marker of social class both in linguistic surveys and for official purposes, social class is nonetheless a more complex topic than the successful use of this single indicator suggests. As discussed in section 2.3, Milroy (1987b: 101) argues that social class is a '*proxy variable* [original italics]' for a range of different factors, including wealth, prestige and life-style. In light of Milroy's comments, it was decided that a question should be included on the *About You* sheet which reflected the personal and subjective nature of social class. Llamas's original wording, *Assessment of social class*, was, however, felt to be ambiguous. On the revised sheet, this was amended to read *Can you say which social class you belong to? If so, which?* In this way, informants are given the option as to whether or not to respond. It is also left to them as to how exactly they define their own social class. As in the case of *Ethnic group*, the possibility of offering tick boxes was discussed and rejected. Though tick boxes are useful in some respects, in that they ensure

informants' responses are comparable, they are also restrictive since they force informants to assess their social class on a spectrum which might not correspond with their own. Far more telling is an informant's own assessment. Macaulay (1977: 65) states that the comments made on the class system by his Glasgow speakers 'indicate that the informants were aware of their own social position'. In instances where a speaker's use of language does not pattern as expected according to social class as defined by their occupation, it is possible to analyse the speaker's language in light of their own subjective assessment of their social class, as one possible means of explaining the findings.

3.3.2.11 Housing

The section *housing* included in Llamas's *Biographical Information* was deemed inappropriate for the revised sheet, and was therefore removed. Whilst housing was used by Trudgill (1974) in his original Norwich study as a marker of social class, it was felt by those undertaking the revision of the existing SuRE methodology that to ask informants directly for this kind of very personal information risked causing them offence, something which could well influence the subsequent interview.

3.3.2.12 Education

The section entitled *education* was felt to be a subject which had to be dealt with sensitively. Initially, the section was rephrased to read *qualifications (e.g. O-levels, A-levels, GCSEs)*, but this was felt to be inappropriate since it is not the qualifications a person has obtained which are important but rather the level to which they have been educated. By rephrasing this section to read *education to what age (e.g. 14, 16, 18, 21)*, informants are not forced to disclose what qualifications they have, but fieldworkers are able to make a reasonable assessment of the level to which informants have been educated.

3.3.3 Alterations Made to the Sense Relation Networks

Several alterations were made to the Sense Relation Networks. Some changes were implemented on the grounds of clarity, others for reasons of space. Notion words and phrases thought to be unproductive were removed, and were replaced with terms that it was believed would yield more data.

3.3.3.1 Renaming of the Sense Relation Network Sheets

The Sense Relation Network sheets originally entitled *Feelings, Actions and States* and *The Outside World* were renamed *Being, Saying and Doing* and *Everyday Life* respectively. There were several reasons for making these changes. The title *Feelings, Actions and States* was felt, like the title *Biographical Information* (discussed in 3.3.2.1), to be too formal, the concern being that this would perhaps prompt informants to complete the SRNs with equally formal language. The new title, *Being, Saying and Doing*, uses plain, informal English, thus helping to ensure accessibility and guard against the production of overly formal responses. The SRN originally entitled *The Outside World* deals with food, money, the law and houses, as well as nature and the weather, and so was renamed *Everyday Life* better to reflect its contents.

3.3.3.2 Alterations to the Sense Relation Network Subdivisions

Several of the subdivisions of each of the three Sense Relation Networks were renamed, and other new subdivisions were added. In the case of the *Being, Saying and Doing* SRN, the renaming of the subdivisions reflected the renaming of the SRN as a whole. The original SRN entitled *Feelings, Actions and States* had three subdivisions: *feelings and states*, *saying things*, and *doing things*. When this SRN was renamed *Being, Saying and Doing*, these three subdivisions were renamed *being*, *saying*, and *doing* respectively. Changes made to the *Everyday Life* (originally *The*

Outside World) and *People* SRNs were made in order that notion words and phrases were grouped under the most appropriate headings, particularly in the light of the removal of unproductive notion words and phrases (see section 3.3.3.4) and the addition of new ones in their place (see 3.3.3.6). The three subdivisions on *The Outside World* SRN were *buildings and jobs*, *nature and weather*, and *food and money*. On the renamed *Everyday Life* SRN, they are: *nature and weather*; *the home*; *crime and the law*; *money*; and *eating, drinking and smoking*. The original *People* SRN was subdivided into *personality*, *appearance*, *body*, and *ages and relationships*. The subdivisions on the revised *People* SRN are: *personality*; *looks, clothes and accessories*; *body*; and *relations and relationships*.

3.3.3.3 Removal of the *any others* Sections from the Sense Relation Network Sheets

It was decided that the sections entitled *any others* which appeared in each subdivision of all three SRNs, in total 10 times, should be removed on the grounds that they were somewhat confusing. The original instructions sheet contained the following request to informants pertaining to these sections:

Use the sections called 'any others' to note down any extra words or expressions you think of (yourself, or in discussion with others). If these are words for things not listed on the sheet, please put down what you think they mean, or what someone not necessarily from your area would understand by them.

This implied that the *any others* sections were to be used not only for words not listed on the SRNs that the informant wished to mention, but also as a kind of 'overflow' area for terms that related to existing notion words. It was felt, however, that while words offered by informants for things not listed on the sheet would be of interest, they would presumably vary greatly from person to person and would therefore be of limited use due to their lack of comparability. Similarly, provided that plenty of room was given for informant responses, the use of the *any others*

sections as ‘overflow’ areas was deemed to be unnecessary, given that informants’ words for things listed on the SRNs could be easily given in the spaces provided underneath the notion word. It was decided that space freed by the removal of the *any others* sections would be better used for additional notion words.

3.3.3.4 Removal of Notion Words Believed to be Unproductive

In total, seven notion words or phrases featured in the original SRNs were removed from the revised SRNs. These were: *eat quickly*; *thank*; *£1*; *£1,000*; *food (in general)*; *legs*; and *teeth*. The decision to remove these words was based on both existing research in the West Midlands using the original SuRE methodology (Asprey 2001) and on fieldworker intuition, and was taken on the grounds that the words were believed to be unproductive.

3.3.3.5 Alterations to Existing Notion Words

On the subject of word association experiments, Aitchison (1994: 83) states that ‘adults are likely to respond with a word of the same word class: a noun tends to elicit a noun, an adjective another adjective’. She goes on to argue that:

[t]here are indications that, on the one hand, words from the same word class are closely connected in the mind, and that, on the other, those from different word classes are more loosely attached.

(Aitchison 1994: 101)

Several notion words, however, belong to more than one word class. In light of Aitchison’s comments, it was decided that it would be sensible to focus on just one of these classes and avoid ambiguity as to which word class the methodology is attempting to elicit. When a word is both a noun and a verb and the desired response is a verb, the infinitive is used as the notion phrase to eliminate confusion. The notion word *mad*, which has a variety of meanings including foolish, angry, and mentally ill, was altered to *insane*, again to avoid ambiguity.

Alterations were also made to ensure that the notion words are not leading, and that the most neutral Standard English term possible is employed. The notion word *boss* was altered to *person in charge at work*. Similarly, the existing opposites *intelligent* and *stupid* were altered to *intelligent* and *unintelligent*, since *unintelligent* is a more neutral word which might yield the term *stupid*. The notion word *glasses* was replaced with *spectacles*, *glasses* having been considered too colloquial and ambiguous to be used as a notion word. *Not use right hand to write with* was altered to *left-handed*, since this is the Standard English term. The notion phrase *(very) small stream* was changed to *running water smaller than river*. In the case of some speakers, *stream* is their only word for a body of water smaller than a river and the use of *stream* in the notion phrase might have proved confusing for this group since they do not have an alternative. To avoid prejudicing informant responses to the notion word *television*, *main room of the house (with TV)* was altered to *main room of the house (with television)*, since *TV* is one of the synonyms *television* might yield.

In several cases on the original SRNs, *run away from (escape)*, *small walkway (path) between houses*, and *talk/chat (a lot)*, synonyms had been used in the notion phrases. These synonyms were removed, since they were clearly superfluous, and the notion phrases now read *run away from*, *small walkway between houses*, and *to talk a lot*. In the case of the notion phrase *pleased/proud*, it was decided that these two words were not synonymous, and so *proud* was removed, *pleased* being the notion word believed to be most productive. The notion phrase *throw away* was changed to *to throw*, and the phrase *men's facial hair (above lip and in front of ears)* now reads *men's facial hair*. Both of these alterations were made so that a broader range of terms might be collected. So as not to be discriminatory, *policeman* was

changed to *the police*. Lastly, minor alterations were made to the use of bracketing and examples used in the notion phrases, again to ensure clarity.

3.3.3.6 Addition of New Notion Words

Aitchison (1994: 84) argues that one of the means by which words are connected in the mind is by co-ordination. On the subject of word association experiments, she states:

The commonest response involves co-ordinates, words which cluster together on the same level of detail, such as *salt* and *pepper*. Opposites come into this category, as they are co-ordinates in a group consisting of only two members, as with *left* and *right*, or they are the two commonest members in a larger group, as with *hot*, *cold*, *warm*, *cool*.

(Aitchison 1994: 84)

Several co-ordinates were already present on the original SRNs, *hot* and *cold*, and *attractive* and *unattractive*, for example. However, whilst the word *tall* appeared, the word *short* did not. In light of Aitchison's comments, *short* was added as the opposite of the existing notion word *tall*, in the expectation that the pairing might prompt the elicitation of more variants for both words. Similarly, *rich* was added as an opposite of *not have any money left*, and the opposites *fat* and *thin* were added. As well as these co-ordinates, the following notion words and phrases were included on the revised SRNs as it was believed that they would also prove to be productive: *to be annoyed by someone*; *to kiss*; *to vomit*; *play truant*; *hello*; *how are you?*; *goodbye*; *long seat in main room of house*; *sweets*; *non-alcoholic drinks*; *cigarettes*; *stomach*; and *bottom*.

3.3.3.7 Alterations to the Presentation of the Sense Relation Network Sheets

Alterations to the layout of the SRNs were inevitable given the removal of unproductive notion words and phrases, and the addition of both new notion words and of new SRN subdivisions. With these new additions, space was at a premium, so

the bubbles containing the names of the SRNs and the SRN subdivisions were made as small as possible to allow informants plenty of room in which to write. Similarly, the subdivisions were rearranged to make the most effective use of space.

On the original SRNs, one or more lines were given under each notion word for informant responses. Though it was decided that the lines should not be removed completely, since they act as a guide as to where exactly informants should write, the different number of lines given were considered to be misleading, implying that different numbers of responses were expected for different notion words. It was decided that the SRNs should be standardised, with just one line under a notion word or phrase, leaving as large a space as possible between the word and the line.

Colour and font are also used to good effect on the revised sheets. While each of the original three SRNs was printed in a different colour, the revised SRNs use multiple colours on each page. The subdivision bubbles are printed in a variety of colours, and each of the three title bubbles is a different colour to ensure that the SRNs are distinct from one another. The aim is to make the sheets as 'visually pleasing as possible' (Upton and Llamas 1999: 299) to ensure that completing them is an enjoyable experience. The font was changed from Times New Roman on the original methodology sheets to Arial on the revised sheets to give them a more modern look. Space is included on each page of the methodology for an informant code so that if sheets from an informant's questionnaire become separated from the *Confidentiality and Consent* sheet, which features the respondent's name, it is clear to which informant they belong. Each sheet is marked as being copyright School of English, University of Leeds.

3.3.4 Confidentiality and Consent Form

Perhaps the most important refinement to the original core SuRE methodology is the addition of a form entitled *Confidentiality and Consent*. The UK Data Protection Act 1998 (Information Commissioner 1998) regulates the collection, storage, and processing of both written and electronic data, and the *Confidentiality and Consent* sheet was designed to ensure the SuRE project complies with this Act. In preparing this form, advice was taken from the University of Leeds lawyer and an independent solicitor. The *Confidentiality and Consent* sheet details the uses to which the collected data, both written and spoken, will be put. The form assures informants that these data will remain completely anonymous outside the SuRE project. They are also informed of their right, if they submit a request in writing, to receive an electronic copy of any information that the University of Leeds uses and stores. Informants are asked to sign the form to indicate their consent to the collection, use, storage and processing of the written and spoken data they have provided. Speakers under the age of 18 must obtain a parent's or guardian's signature in order to participate. A copy of the revised core SuRE sheets, including the *Confidentiality and Consent* sheet, can be seen in section 3.4 below.

3.4 The Revised Core SuRE Methodology Sheets

SuRE

What is SuRE?

SuRE, the Survey of Regional English, is a joint project from the Universities of Leeds and Sheffield. Its purpose is to collect information about the way English is used in the British Isles today. Research has recently been completed in Middlesbrough, and fieldworkers are now carrying out research in the Black Country, Southampton, and Sunderland. The ultimate aim of the Survey is to create a lasting record of British English as it is spoken at the turn of the 21st century.

Instructions

- Please complete the sheets with words you think are local to the area you live in.
- Write down whatever comes to mind – words that you use everyday when talking to friends, for example.
- Once you have done that, think about it for a while and note down any other examples of words local to your area which come to mind.
- Feel free to discuss the words with other people from the same area as you but try to keep a note of who you discuss them with, especially if you note down their suggestions.
- Feel free to use expressions as well as single words.
- Give as many examples as you can. If the only word you can think of is the one given please write that down. This is important information in its own right.
- When you have completed the sheets, a short discussion about the words and phrases you have thought of will follow. More information is provided about this on the *Confidentiality and Consent* sheet.

About You

Sex.....

Date of birth.....

Place of birth.....

Current place of residence.....

Other places you have lived and for how long.....

.....

.....

Birth place of mother.....

Birth place of father.....

Ethnic group.....

Usual occupation(s).....

Mother's usual occupation(s).....

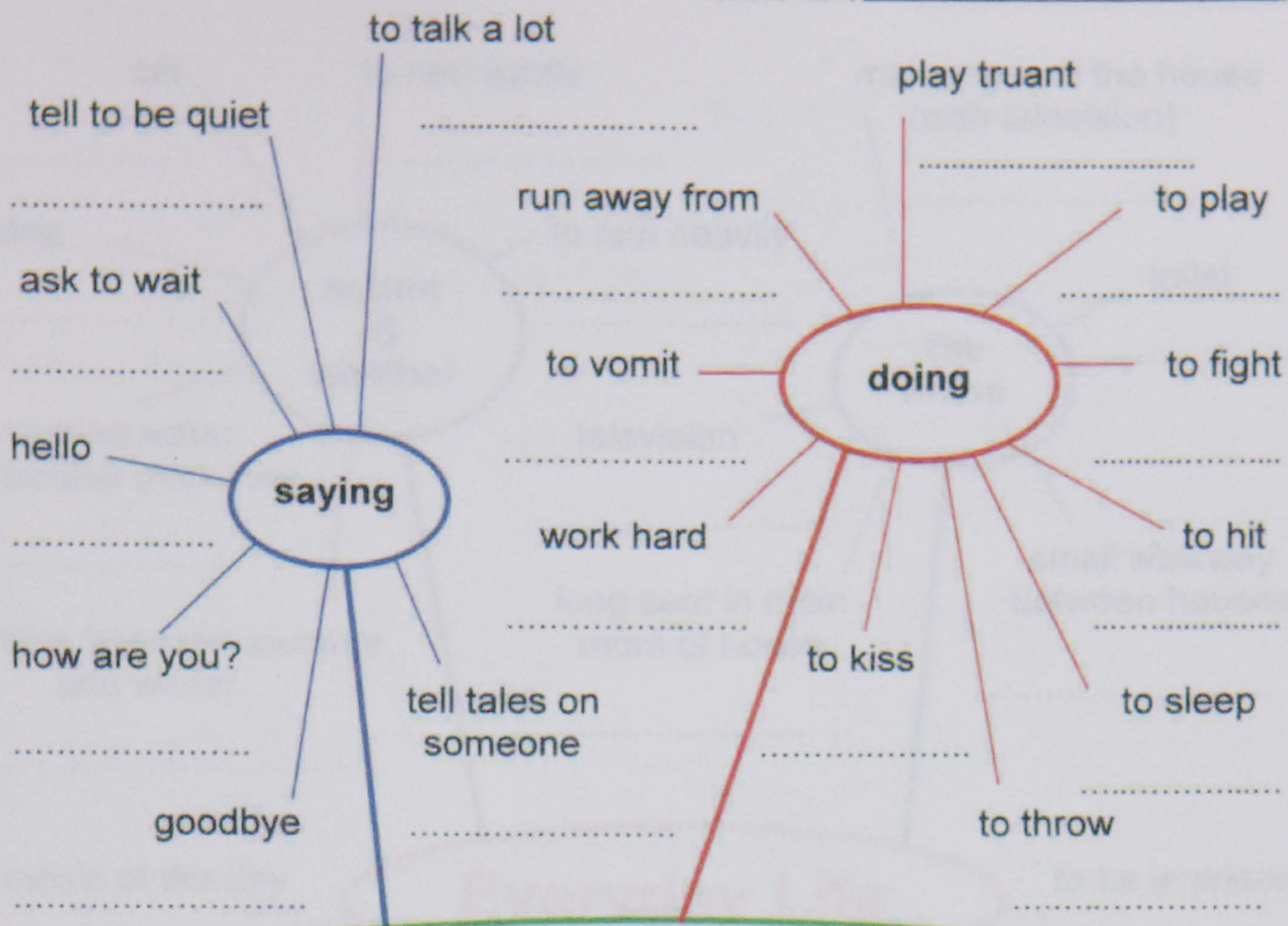
Father's usual occupation(s).....

Can you say which social class you belong to? If so, which?

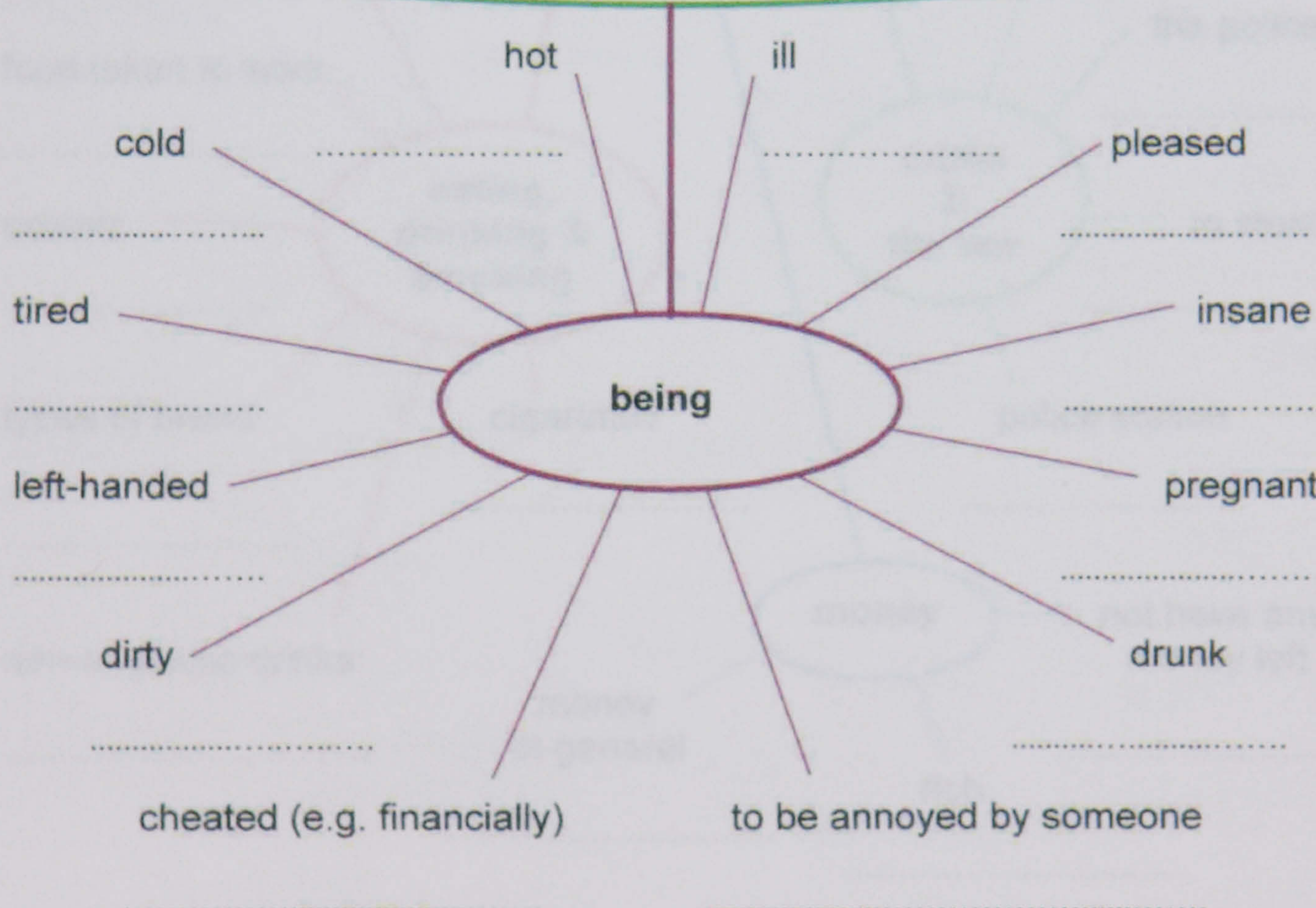
.....

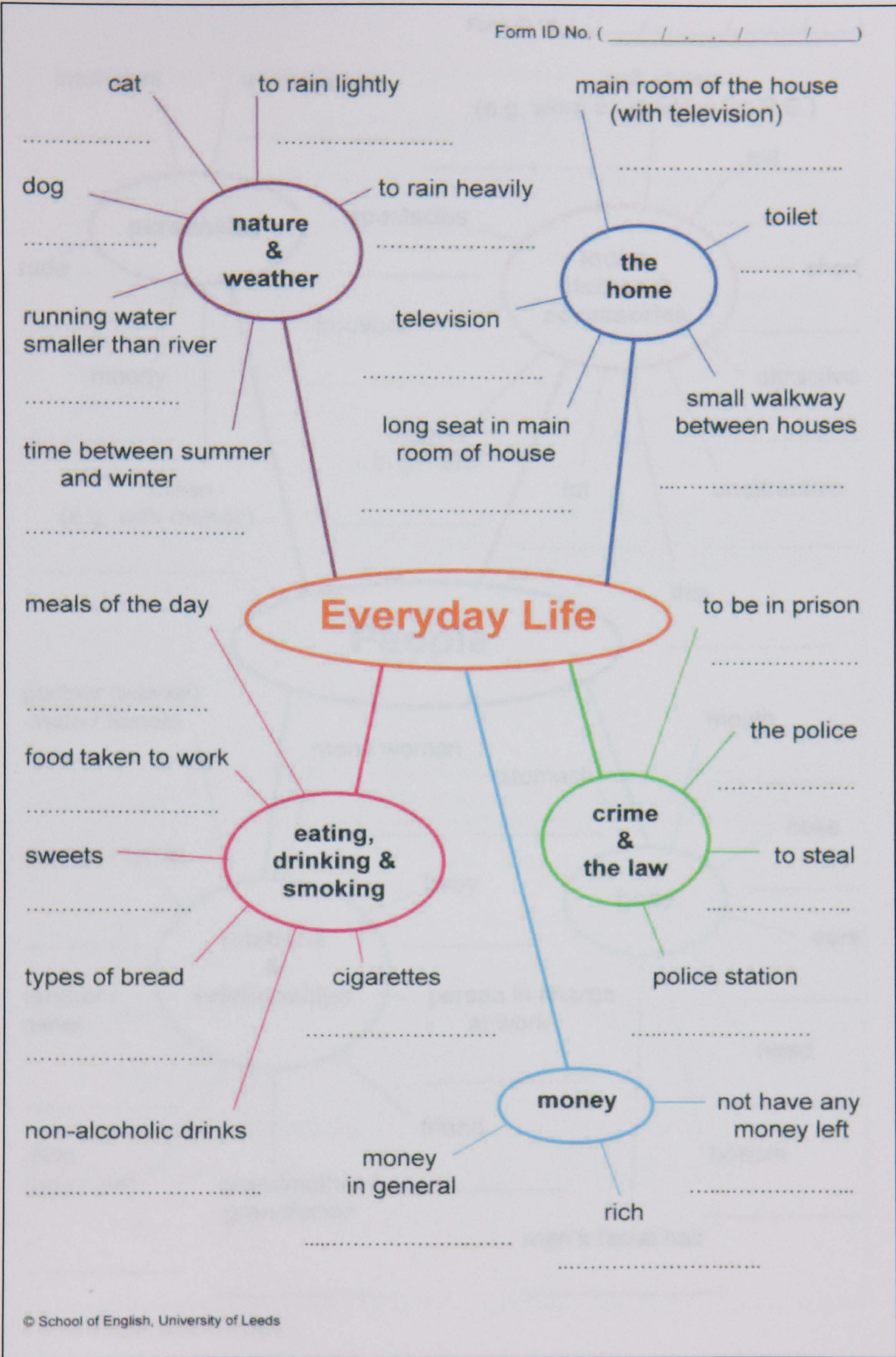
Education to what age (e.g. 14, 16, 18, 21)

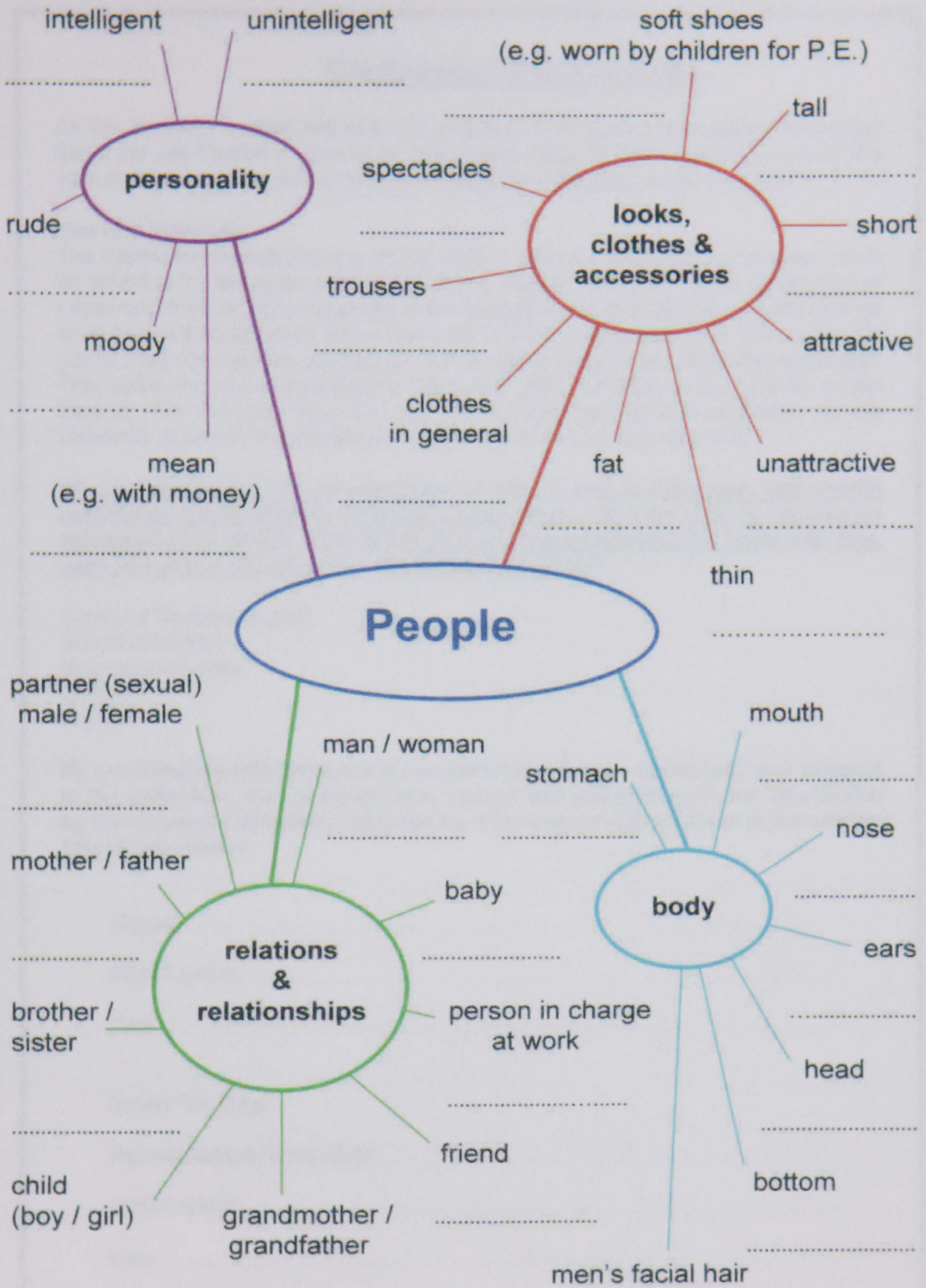
.....



Being, Saying & Doing







Confidentiality and Consent

As the *Instructions* sheet explains, the purpose of this survey is to gather information about the way English is used in the British Isles today. In order to do this, I would like your permission to record the discussion about the sheets you have completed.

Use of Information

The information provided by you will be used to describe language in your area. It will be added to the electronic database of British dialects being built by the Universities of Leeds and Sheffield (**this database is the project referred to below**), and will then go on to be used worldwide in future academic research and publications. The recordings you provide may also be used as an Internet resource by other universities worldwide. This would involve the publication of segments of the recording as a sound file on the Internet. The data you have provided will be held securely and indefinitely by the University of Leeds, in accordance with the Data Protection Act (UK) 1998.

All the information you provide, both in written and spoken form, will remain completely anonymous outside the project, and it is your right to receive an electronic copy of any information you have provided that the University then uses and stores, by contacting the following address:

Survey of Regional English
School of English
University of Leeds
LEEDS
LS2 9JT

By providing me with written and recorded information you indicate your consent to the collection, use, and electronic storage and processing of this information by the University of Leeds, solely for the purposes described above in the section *Use of Information*.

Signed.....

PRINT NAME.....

Date.....

(Under 18s only)

Parent/Guardian's signature.....

PRINT NAME.....

Date.....

3.5 Additions Made to the Core SuRE Methodology for the Southampton Area Study

As well as permitting the collection of lexical, grammatical, and phonological data, the SuRE methodology is also easily expanded to fit the requirements and interests of individual researchers. In the case of the Southampton area study, several additions have been made to the core method.

3.5.1 The Southampton Area Identity Questionnaire

Mendoza-Denton (2002: 478) states that there is doubt in some quarters as to whether quantitative research is sufficient to ‘conceptualize and investigate identity’. However, on the subject of qualitative research, Atkinson and Silverman (1997: 304) urge researchers who put ‘special faith in the interview as the prime means of data collection’ to be cautious. Their concern is the ‘elevation of the experiential as the authentic’ (Atkinson and Silverman 1997: 305). Atkinson and Silverman (1997: 315) argue that our society places great store by the interview as a means of accessing ‘the true identity’ of the informant. The authors urge us, however, not to ‘miss the significance of narrative and biographical work in inventing the self’ (Atkinson and Silverman 1997: 318). Atkinson and Silverman (1997: 322) are not attempting to dissuade researchers from the use of interviews to collect data on speaker identity, but rather reminding them that the personal narratives elicited in interviews are not any more ‘authentic or pure a reflection of the self than any other socially organized set of practices’.

With these arguments in mind, both qualitative and quantitative data on identity are collected in the Southampton area study, and an identity questionnaire based on that used by Underwood (1988) in his study of Texan identity is employed in order to obtain these data. A copy of the Southampton area identity questionnaire

can be seen in section 3.5.1.1. Informants are asked four questions, the first of these being if they feel any closer to local people than to people from elsewhere. The second question requires informants to imagine they are the manager of a company that must hire a scientist. Two people apply – one born and educated in Southampton/Eastleigh Borough and the other born and educated in another place. Presuming the applicants are equally qualified, informants are asked to say whether they would pick the local person or the applicant from elsewhere. The third question asks informants whether they would favour a local Member of Parliament who was locally born and raised or one from another area. Of the use of these questions in his Texan questionnaire, Underwood states:

Whereas the first question directly asks subjects to say whether they feel close to other Texans, the other two measure what Reed [...] calls 'in-group preference' in a situation that forces the subject to make a choice. By the tabulation of scores appropriate to a subject's responses to these questions, an individual's level of Texan identity can be scored on a scale from 0 to 6, from low identification to high.

(Underwood 1988: 410)

In addition to the three questions based on those used by Underwood in Texas, a fourth question on the Southampton area identity questionnaire asks informants whether they feel closer to people from the South East or the South West of England. For each of the four questions, informants can select a neutral response (*can't say, don't know, it depends, neither*) so that they are not forced to give an answer which does not correspond to their own opinion on a particular subject, something which could well skew results. Speakers' responses to all four questions are discussed in the interview, so as to elicit qualitative data as well as, in the case of the first three questions, quantitative data. Responses to question 4 are not quantified. The same scoring employed by Underwood is used in the Southampton area study. A positive answer (one favouring local people) has received a score of 2.

a neutral response (answering *don't know, can't say* or *it depends*) a score of 1. and a negative response (one favouring people from outside the area) a score of 0.

3.5.1.1 A Copy of the Southampton Area Identity Questionnaire

Form ID No (_____)

Please read the following questions and tick the appropriate boxes

1. Some people in Southampton and Eastleigh Borough feel they have a lot in common with other people from these areas. How about you? Would you say you feel close to other people from the local area, or that you don't feel much closer to them than you would to people from somewhere else?

- a. Feel closer to local people
- b. No closer to local people than to others
- c. Don't know, can't say

2. You are the manager of a company that must hire a scientist. Two people apply – one born and educated in Southampton/Eastleigh Borough and the other born and educated in another place. If they were equally qualified, which would you prefer, the person from Southampton/Eastleigh Borough or the person from somewhere else?

- a. Person from Southampton/Eastleigh Borough
- b. Person from another area
- c. It depends, don't know

3. Two people want to be your local Member of Parliament. One was born and raised locally, and the other was born and raised in another area. Which one would you favour, the person born and raised locally or the person born and raised in another area?

- a. Person from Southampton/Eastleigh Borough
- b. Person from another area
- c. It depends, don't know

4. Overall, would you say that you feel closer to people from the South East or the South West of England?

- a. South East
- b. South West
- c. Neither

3.5.2 The Southampton Area Language Questionnaire

Cheshire et al. (1989: 185) describe our knowledge of the phonology of British English dialects as being greater than our knowledge of their morphology and syntax. This is still the case, and there is no doubt that this is due in part to the often problematic and time-consuming nature of grammatical data collection in comparison with the collection of phonological material. Rickford et al. state that:

unlike phonological variables, which show up with high frequencies in [sociolinguistic] interviews, syntactic variables often involve special semantic and pragmatic circumstances which may occur rarely or unpredictably in interview settings.

(Rickford et al. 1995: 106)

It took Rickford et al. (1995: 106) eight years to collect 1200 tokens of the *as far as* construction. Around 500 of those tokens were taken from computer corpora. Though possible in phonological data collection, it is clearly unrealistic to expect to elicit a statistically meaningful number of tokens of a given grammatical variable by simply waiting for a speaker to produce that particular variable in an interview conducted in a limited time-span. Even if the informant were to produce sufficient tokens of a particular variable, it is unlikely that their utterances would be comparable with those of other informants, since the utterances would not have been elicited in a structured manner. This would thus limit the usefulness of the data from a social dialectological point of view.

As in Llamas's (1999) Middlesbrough study, a language questionnaire of the type used by Cheshire et al. (1989) in the Survey of British Dialect Grammar is employed in order to investigate nonstandard grammatical features in the speech of the Southampton area (a copy of the questionnaire can be seen in section 3.5.2.1). By using a questionnaire, it is hoped that at least some of the problems of grammatical data collection can be overcome. Speaker responses to the

questionnaire are well structured and easily comparable. Not only is it possible to compare one speaker's responses with those of another speaker, but it is also possible for an informant's responses to the questionnaire to be 'compared to and correlated with the informant's actual usage of nonstandard grammatical features in informal speech' (Llamas 1999: 106).

Cheshire et al. (1989: 187) include in the Survey of British Dialect Grammar most of the grammatical features whose distribution has been investigated by the Survey of English Dialects (Orton et al. 1962-1971) so that they 'could trace the extent to which the current distribution of these features reflects earlier boundaries'. In employing a language questionnaire similar to that used by Cheshire et al., the Southampton area study is comparable with both the Survey of British Dialect Grammar and the SED.

Sentences that exemplify nonstandard grammatical forms believed to be salient to the Southampton area were selected from the questionnaire used by Cheshire et al.. The grammatical variables under investigation in the Southampton area study are detailed in section 1.7.4. Several Standard English sentences are also included by way of a control to ensure that the questionnaire is being completed correctly by informants. As in the case of Llamas's (1999: 117) language questionnaire, three possible responses to each sentence are offered: that the informant would use the sentence themselves when talking to a friend; that they would not use the sentence themselves but hear it used by others in the local area; and/or that they would use it themselves when writing to a friend. Trudgill argues:

One important ability that speakers have with respect to their own dialect is that they are able to make grammaticality judgements about it. Drawing upon their native linguistic competence, they are usually able to state whether particular grammatical constructions are or are not possible in their variety.

(Trudgill 1983: 14)

Please read the following sentences and tick any appropriate boxes

	I'd use this when I was talking to a friend	I wouldn't use this but some people do use it in Southampton/ Eastleigh Borough	I'd use this when I was writing to a friend
That ain't a bird	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
That ain't working	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I ain't got a clue	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
He writes really quick	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
She laughs really loudly	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I'm going up my friend's house later	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I'm going down my friend's house later	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I'm going over my friend's house later	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I'm going round my friend's house later	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
He knocks his hat off of his head	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Look at them big spiders	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
This is the beautifullest house I've seen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I've never seen a beautifuller one	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
This is the most beautifullest house I've seen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I've never seen a more beautifuller one	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
But this is the worstest one I've seen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I've never seen a worser one	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
John likes doing that hisself too	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Yes, lots of people do it theirselves	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please read the following sentences and tick any appropriate boxes

I'd use this when I was talking to a friend

I wouldn't use this but some people do use it in Southampton/ Eastleigh Borough

I'd use this when I was writing to a friend

- The films what I like best are horror films
- The films as I like best are horror films
- The films at I like best are horror films
- He's the boy what ate all my sweets
- He's the boy as ate all my sweets
- He's the boy at ate all my sweets

- That town is nearly twenty mile away
- To make a big cake you need two pound of flour
- This string is three inch long
- I've got four dozen eggs

3.5.3 The Southampton Area Word List

The final addition to the core SuRE methodology in the Southampton area survey is a word list which can be seen in section 3.5.3.1. Though ‘the word list test is [...] an artificial one, in that informants are not likely to use their ‘normal everyday pronunciation’ when reading out a single list of words [original quotation marks]’ (Trudgill 1974: 48), it is an important tool in facilitating the collection of formal style phonological data for comparison with the more casual style data collected during interviews. Labov (1972b: 208) argues that ‘*styles can be ranged along a single dimension, measured by the amount of attention paid to speech* [original italics]’.

This dimension, from least to most formal, is, according to Labov (1966: 91-98) as follows: casual speech; careful speech; reading style; word list; minimal pairs.

Whilst acknowledging that Labov has been criticised for his analysis of style ‘which does not, for example, address the issues of how speakers use linguistic variation to mark out their psycho-social orientation to each other’, Milroy (1987b: 37) argues that it was not Labov’s intention to make a ‘general statement about the possible range of speech styles’. In light of that fact, Milroy states that:

there is no reason why researchers should not continue to use Labov’s original methods of sampling styles, despite their limitations; they are useful and practicable where the objective is to examine contrastively the characteristics of two or more different types of language used by the same speaker.

(Milroy 1987b: 37)

The addition of a word list to the core SuRE methodology is also useful given that:

a test of this form ensures that a large amount of necessary phonological information can be gathered which one could not otherwise be certain of eliciting during an entirely conversational interview.

(Trudgill 1974: 48)

Wells’s (1982: xviii) standard lexical set keywords form the basis of the word list. He describes these keywords as ‘intended to be unmistakable no matter what

accent one says them in' (Wells 1982: xviii), and Foulkes and Docherty (1999: 7) state that they 'have become a standard tool within descriptive dialectology'. In addition to the lexical set keywords, words which exemplify consonantal features under investigation in the Southampton area are also included. These are as follows:

T Glottalling

- *Stutter* – intervocalic
- *Rattle* – preconsonantal
- *Bought* – word-final

TH Fronting

- *Third* – θ word-initial
- *This* – ð word-initial
- *Brother* – ð intervocalic
- *Teeth* – θ word-final
- *Methane* – θ intervocalic
- *Bathe* – ð word-final

L Vocalization

- *Milk* – preconsonantal
- *Middle* – word-final
- *Rattle* – word-final

Rhoticity

- *Far* – word-final
- *Farm* – preconsonantal

(ng)

- *Laughing*

The word *opposite* is also included on the word list as I had heard some Southampton area speakers voicing /p/ in this word and wished to investigate this phenomenon further.

3.5.3.1 A Copy of the Southampton Area Word List

Kit
Stutter
Dress
Rattle
Trap
Lot
Bought
Strut
Third
Foot
This
Bath
Brother
Cloth
Laughing
Nurse
Far
Fleece
Farm
Face
Milk
Palm
Middle
Thought
Goat
Goose
Teeth
Price
Choice
Methane
Mouth
Near
Square
Start
North
Force
Cure
Bathe
Opposite

3.6 Speaker Selection

There were three criteria for the selection of speakers for the Southampton area study, the first of these being that the speaker should have acquired their language in the area. While some speakers were born outside the sampling universe, for example in nearby Winchester where there is a large maternity unit, all learnt to speak while resident in the Southampton area. Trudgill (1983: 13) remarks that 'it is a very rare adult that successfully masters the speaking of a new dialect in all its details', and Payne's (1980) work in King of Prussia shows that even children are variably successful in their acquisition of a new accent. The second criterion for speaker selection was that the person should have spent the majority of their life in the Southampton area. Speakers who have not might have linguistic behaviour uncharacteristic of the area. However, given the geographical mobility which characterises contemporary British society, it was deemed unrealistic and unrepresentative to attempt to single out for investigation only those who had always been resident in the Southampton area. While some speakers in the survey had never lived anywhere but the Southampton area, many informants had lived for varying lengths of time in other locations. The third, and most obvious, criterion was that the speaker should be resident in the Southampton area at the time of the interview.

Speakers were obtained for the Southampton area survey using a variety of methods. Contact was made with a local sixth form college and a Southampton church, and a letter was written to the *Southampton Daily Echo* asking for volunteers. The majority of informants, however, were obtained using the 'friend of a friend' technique (Milroy 1987b: 66). Of the use of this method in her Belfast study, Milroy states:

I introduced myself initially [...] not in my formal capacity as a researcher, but as a 'friend of a friend' [...] mentioning the name of a person categorized as an insider with whom I had previously made contact and who had given me the names of persons who might initially be approached. As a consequence of the reciprocal rights and obligation which members of closeknit groups contract with each other, the mention of the insider's name had the effect of guaranteeing my good faith.

(Milroy 1987b: 66)

Through friends and family, I was introduced to, and interviewed, people who fulfilled my speaker selection criteria. A degree of controversy, however, surrounds this particular sampling method. Trudgill states:

Informants selected solely because they are available and willing to be interviewed are simply a part of the population of the city, not a representative sample, and no valid statements concerning the language of the city as a whole can be based on evidence obtained from informants selected in this way.

(Trudgill 1974: 20-21)

This is a view echoed by Labov, who argues:

An accurate view of an urban community cannot be obtained by the study of a few individuals, or of small groups, nor even of extended social networks of 30 or 40 individuals. Most importantly, it cannot be obtained by any approach that begins with the personal connections of the investigators. A truly representative sample of the speech community must be based on a random sample in which each one of several million speakers has an equal chance of being selected.

(Labov 2001: 38)

In his original Norwich study, Trudgill (1974: 21-25) employs a quasi-random sampling method whereby informants are selected from the register of electors and then approached, firstly by letter and then in person, to participate in the survey. He does, however, highlight several problems associated with the use of this method which include the low success rate for securing interviews, the unsuitability of certain selected participants, and the absence of informants under the age of twenty-one, given the voting age at the time the study was conducted (Trudgill 1974: 25-27). In light of these problems, and the possible risk to a lone female fieldworker

approaching strangers, and given the success of Milroy's Belfast study, the use of the 'friend of a friend' technique was judged to be the best means of selecting the majority of speakers for the Southampton area survey.

3.7 The Interview

Milroy (1987b: 41) states that 'an interview in western society is a clearly defined [...] speech event to which a formal speech style is appropriate'. This is a view supported by Labov, who argues:

In the main body of an interview, where information is requested and supplied, we would not expect to find the vernacular used. No matter how casual or friendly the speaker may appear to us, we can always assume that he has a more casual speech, another style in which he jokes with his friends and argues with his wife.

(Labov 1972b: 209)

Milroy goes on to argue that:

individuals who are being questioned will seldom produce large volumes of speech in their replies. [...] sometimes they may perceive questions as a 'test' of some sort [original quotation marks].

(Milroy 1987b: 46)

Informants for the Southampton area survey were given their questionnaire packs at least five days in advance, giving them ample time to prepare. In this way, it was hoped to gain access to a style of speech as close to the vernacular as possible. By ensuring that informants know what will be discussed in advance, the SuRE methodology aims to lessen the test-like nature of the interview situation.

Interviews took place in a variety of locations: in the informant's home; at my own home if the informant did not want me to call at theirs; in a classroom at the college I had contacted; and, on one occasion, in a public house. Wherever possible, interviews took place with pairs of speakers. In some cases, however, this was not possible. When a speaker did not have a partner, either a friend or relative, to be

recorded with, they were interviewed alone. Though such speakers could have been paired with others in a similar situation, it was felt that to have done so would have made the interview situation more rather than less intimidating, given that they would probably not have known me, and would definitely not have known the person with whom they had been paired. Of course, there is a possibility that the speech of those interviewed alone might be different to those interviewed in pairs, and this is taken into consideration when analysing the collected data.

Recordings were made using a Sony MZ-N707 net minidisc recorder and a Sony ECM-717 stereo microphone. Recorded interviews were transferred to compact disc using the Sound Forge program for backup purposes.

Interviews were, in the main, conducted in the following order:

1. *Being, Saying and Doing* SRN
2. *Everyday Life* SRN
3. *People* SRN
4. Identity questionnaire
5. Grammar questionnaire
6. Discussion about the local area which normally covered speaker likes and dislikes, suggestions for improvement, thoughts on local community
7. Word list

It was felt that by leading with the sheets the informants had been asked to prepare in advance, informants would feel more relaxed when the time came to answer questions and read the word list, reading aloud being a task few people relish, particularly when being recorded. The interviews lasted between twenty minutes and one hour and ten minutes, depending on whether the interview was being conducted with a single informant or with a pair and on how talkative the speakers were.

3.8 Sample Size

A total of seventy-six informants were recorded but, given the time constraints of the research, the final sample on which the Southampton area study is based consists of sixty speakers, the same number as for Trudgill's (1974: 27) Norwich survey. Labov (2001: 38-39) argues that 'a reliable sample of a very large city can be achieved with comparatively few speakers: in most cases, less than a hundred'. He states:

[W]e find that the basic patterns of class stratification, for example, emerge from samples as small as 25 speakers. [...] regular arrays of stylistic and social stratification emerge even when our individual cells contain as few as five speakers and we have no more than five or ten instances of the given variable for each speaker.

(Labov 1972b: 204)

The speaker cells for the Southampton area survey are shown in Table 3.1.

Table 3.1 Speaker cells for the Southampton area survey

	Southampton City		Eastleigh Borough	
	Male	Female	Male	Female
<30	5	5	5	5
30-59	5	5	5	5
>59	5	5	5	5

Labov argues:

An understanding of age effects on language – the principles that govern distributions in apparent time – requires an understanding of the changes in social relations across speakers' life histories that bear upon their acquisition and use of linguistic norms and their ability to put them into practice. These include their changing affiliations with successive reference groups, their acquisition and employment of symbolic capital, and relaxation of the norms of the dominant society in old age.

Divisions of the age continuum into groups must be roughly consonant with life stages.

(Labov 2001: 101)

The examination of the effect of age on language in the Southampton area takes place in apparent time. The distribution of linguistic variables across age groups are analysed, and the results used to infer temporal developments. This methodology contrasts with real-time studies, 'whereby linguists make a series of observations of

similar populations over many years' (Chambers 2003: 212). As Labov (2001: 101) argues, it is important that the age groups under examination are consistent with life stages. The Southampton area speaker sample is divided into three groups: those under 30; those between 30 and 59; and those aged 60 and over. These groupings correspond roughly to adolescence and young adulthood, adulthood and middle-age, and later middle-age and old age, which have been shown to produce patterns of linguistic age differentiation in empirical research such as that discussed (Labov 1972b; Trudgill 1974; Macaulay 1977).

The final 60 speakers chosen for the survey were selected to ensure an even spread of ages within each age group and because they had provided sufficient data for analysis. Eight 17- to 19-year-olds were interviewed at the Southampton City college I visited. Clearly to have an under-30 speaker cell comprised predominantly of informants under the age of 20 would be unrepresentative, so more informants from Southampton under the age of 30, but over 20, were recorded to provide an even age spread, and only four of the college informants were used. In some instances, speakers simply did not speak for long enough to provide sufficient data for analysis, and their recordings were rejected. In a few cases, speakers said that they were from Southampton City but were actually from Eastleigh Borough according to administrative boundaries, and their recordings were rejected on the grounds that they did not fit clearly into either half of the sampling universe. While the matter of identity for these people is obviously complex and potentially of great interest, there was not a sufficient number of these speakers, nor was there sufficient time, to investigate the issue. However, it is further justification of the decision to amalgamate Southampton City and Eastleigh Borough to form a single sampling universe, since it shows that the two areas are linked in the minds of their inhabitants.

3.9 Social Stratification of Southampton Area Informants

Section 2.3.1 details some of the approaches taken by linguists to the matter of defining and determining social class. Each of these methods has its advantages and disadvantages, and all methods have been considered when deciding how to determine social class for the Southampton area study.

On the SuRE *About You* sheet, information was requested from informants, regarding education, parents' occupation, own occupation, and place of residence, meaning that a modified version of the multiple-item social class indices employed by Labov (1966) and Trudgill (1974) would be possible. However, there is no consensus as to which or how many indicators of social class should be employed in constructing such an index. Chambers (2003: 52) argues that 'multidimensional indices are much more elaborate than is necessary in sociolinguistics', and this means of determining social class has been rejected as being unnecessarily complex and subjective.

As discussed in section 1.3.1, the sampling universe for the current survey is not a socially homogeneous one, so the use of geographical area as determinant of social class, as employed by the Milroys (1978) in Belfast, is not viable. The social network approach adopted by Eckert (2000) has been rejected on the grounds that speakers in the Southampton area study have been examined in relative isolation, so it is not possible to ascertain the networks to which they belong.

Since occupation is agreed by many linguists to be 'the touchstone of social class membership' (Chambers 2003: 47) (see, for example, Labov 1966: 64, Trudgill 1974: 36), the Southampton area survey uses this as the criterion by which to stratify informants. Occupation as the sole measure of social class has been used by Macaulay (1977) in Glasgow, and his study shows it to be a relatively simple yet

effective means of stratifying a speaker sample. To ensure that stratification by occupation is as objective as possible, the National Statistics Socioeconomic Classification (NS-SEC) is employed to rank and group informants.

3.9.1 The National Statistics Socioeconomic Classification (NS-SEC) and its Use in the Southampton Area Study

The National Statistics Socioeconomic Classification (NS-SEC) replaces Social Class based on Occupation (formerly the Registrar-General's Social Class) and Socioeconomic Groups, the two socioeconomic classifications previously widely used in academic research and official statistics in the UK (Office for National Statistics 2005: 2). There exists a precedent in linguistic research for the use of such classification scales in the work of Trudgill (1974) and Macaulay (1977). The Office for National Statistics states:

NS-SEC has been constructed to measure employment relations and conditions of occupations. Conceptually, these are central to showing the structure of socio-economic positions in modern societies and helping to explain variations in social behaviour and other social phenomena.

(Office for National Statistics 2005: 3)

NS-SEC is comprised of seventeen main operational categories, some with sub-categories (Office for National Statistics 2005: 8). These are shown in Figure 3.1. Of these seventeen categories, fourteen, L1 to L14, are functional, and three, L15 to L17, are residual (Office for National Statistics 2005: 8). 'The functional categories represent a variety of labour market positions and employment statuses' (Office for National Statistics 2005: 8), and can be collapsed down into eight-, five- or three-class versions of NS-SEC (Office for National Statistics 2005: 15).

However, the residual categories, used to classify students and people who, for whatever reason, cannot be otherwise classified, are excluded when the operational

categories are collapsed down into analytic classes (Office for National Statistics 2005: 8).

Figure 3.1 Operational Categories and Sub-categories of the National Statistics Socioeconomic Classification (NS-SEC) (Office for National Statistics 2005: 9)

L1 Employers in large organisations	L10 Lower supervisory occupations
L2 Higher managerial occupations	L11 Lower technical occupations
L3 Higher professional occupations	L11.1 Lower technical craft
L3.1 'Traditional' employees	L11.2 Lower technical process operative
L3.2 'New' employees	L12 Semi-routine occupations
L3.3 'Traditional' self-employed	L12.1 Semi-routine sales
L3.4 'New' self-employed	L12.2 Semi-routine service
L4 Lower professional and higher technical occupations	L12.3 Semi-routine technical
L4.1 'Traditional' employees	L12.4 Semi-routine operative
L4.2 'New' employees	L12.5 Semi-routine agricultural
L4.3 'Traditional' self-employed	L12.6 Semi-routine clerical
L4.4 'New' self-employed	L12.7 Semi-routine childcare
L5 Lower managerial occupations	L13 Routine occupations
L6 Higher supervisory occupations	L13.1 Routine sales and service
L7 Intermediate occupations	L13.2 Routine production
L7.1 Intermediate clerical and administrative	L13.3 Routine technical
L7.2 Intermediate sales and service	L13.4 Routine operative
L7.3 Intermediate technical and auxiliary	L13.5 Routine agricultural
L7.4 Intermediate engineering	L14 Never worked and long-term unemployed
L8 Employers in small organisations	L14.1 Never worked
L8.1 Employers in small organisations (non-professional)	L14.2 Long-term unemployed
L8.2 Employers in small organisations (agriculture)	L15 Full-time students
L9 Own account workers	L16 Occupations not stated or inadequately described
L9.1 Own account workers (non-professional)	L17 Not classifiable for other reasons
L9.2 Own account workers (agriculture)	

There are three means of deriving the functional categories of NS-SEC, and these are dependent on the amount of information the researcher has collected about their informants' jobs. To use the full method, it is necessary to know a person's occupation, their employment status (whether they are an employer, an employee or self-employed, and whether or not they are a supervisor), and the size of the organisation for which they work (Office for National Statistics 2005: 24). To use the reduced method, information about occupation and employment status is needed, and to use the simplified version, information about occupation alone is required. The Office for National Statistics (2005: 21-22) recommends groups of specific questions be asked in order to derive this information. These questions are quite detailed, for example asking informants to describe fully what the company for which they work makes or does, and it was felt that to ask questions regarding employment in addition to those which already appear on the SuRE *About You* sheet would place a disproportionate amount of emphasis on occupation, and might appear intrusive. For this reason, the Office for National Statistics questions have not been employed in the current survey.

Instead, an adapted version of the simplified means of deriving operational categories has been selected. Using the job titles or job descriptions provided by informants in response to the request on the *About You* sheet for *usual occupation(s)*, it is possible to allocate each speaker an operational category. For example, using the Standard Occupational Classification 2000 on which NS-SEC is based, a chartered accountant falls into the operational category 3.1 (Office for National Statistics 2005: 41). Once an operational category has been allocated, it is possible to allocate informants to an analytic class. Since the current survey is relatively small when compared with some of the research for which NS-SEC is used, the

decision has been made to employ the three-class version of the NS-SEC for analytic purposes. To subdivide the sample into the five- or eight-class version would be to risk producing groups which are so small that it would not be possible to draw meaningful conclusions regarding their language use. In the three-class version, occupational categories 1 to 6 equate to analytic class 1, occupational categories 7 to 9 equate to analytic class 2, and occupational categories 10 to 13 equate to analytic class 3 (Office for National Statistics 2005: 9, 15). Table 3.2 shows the operational categories and analytic classes, where possible, of the sixty Southampton area informants.

Table 3.2 National Statistics Socioeconomic Classification Operational Categories and Analytic Classes of Southampton Area Informants

Speaker number	Occupation	Operational category	Analytic class
1	Hairdressing salon owner	8.1	2
2	Housewife	16	N/A
3	Secondary school teacher	4.1	1
4	A-level student	15	N/A
5	A-level student	15	N/A
6	A-level student	15	N/A
7	A-level student	15	N/A
8	Undergraduate student	15	N/A
9	Administrator	7.1	2
10	Accounts clerk	7.1	2
11	Police officer	7.2	2
12	Hairdresser (employed)	13.1	3
13	Railway engineer	11.1	3
14	Cleaner	13.4	3
15	Instrument technician	11.1	3
16	Retail manager	5	1
17	Childminder	9.1	2
18	A-level student	15	N/A
19	Printer	11.1	3
20	Welder	13.3	3
21	Bank clerk	7.1	2
22	Secondary school teacher	4.1	1
23	Postgraduate student	15	N/A
24	Local government committee clerk	7.1	2
25	Retail assistant	12.1	3

Table 3.2 National Statistics Socioeconomic Classification Operational Categories and Analytic Classes of Southampton Area Informants (continued)

Speaker number	Occupation	Operational category	Analytic class
26	Bakery assistant	12.1	3
27	I.T. manager	3.2	1
28	Plumber	11.1	3
29	Shelf stacker	12.1	3
30	Undergraduate student	15	N/A
31	Car production worker	12.4	3
32	Accounts clerk	7.1	2
33	Actor	4.3	1
34	Business consultant	4.1	1
35	Welder	13.3	3
36	Housewife	16	N/A
37	Production supervisor	10	3
38	Plumber	11.1	3
39	Chartered accountant	3.1	1
40	Business consultant	4.1	1
41	Clerical worker	7.1	2
42	Housewife	16	N/A
43	Secondary school teacher	4.1	1
44	Secondary school teacher	4.1	1
45	Undergraduate student	15	N/A
46	Head teacher	4.1	1
47	Sales representative	4.2	1
48	Housewife	16	N/A
49	Fire fighter	7.2	2
50	Chartered accountant	3.1	1
51	Head teacher	4.1	1
52	No response	N/A	N/A
53	Carpenter	9.1	2
54	No response	N/A	N/A
55	Office worker	7.1	2
56	Secondary school teacher	4.1	1
57	Production operator	13.2	3
58	Sheriff's officer	7.2	2
59	I.T. administrator	7.1	2
60	Local government officer	7.1	2

3.9.2 The NS-SEC and Full-Time Students

A substantial proportion of Southampton area informants are full-time students.

Nine of the current study's sixty speakers state that they are undertaking A-level.

undergraduate or postgraduate study. As full-time students, they are categorised by

the NS-SEC as L15, a residual category which is omitted when occupational categories are collapsed into analytic classes. To exclude 15% of informants from the analytic classes would be to lose a significant proportion of speakers from analysis by social class. As such, an alternative means of classifying them has been adopted for the current study, based on Office for National Statistics guidelines.

which state:

Full-time students are recognised as a category in the full classification for reasons of completeness. Since many students will have had or still have paid occupations, you could classify them by current or last main job, although we would not usually expect them to be classified this way. Conventionally, where full-time students are included in analyses [...], they are normally allocated a position through their family household.

(Office for National Statistics 2005: 13-14)

All informants are asked to provide details of both their mother's and their father's occupation, so it is possible to allocate the students in the study to an occupational category, and by extension an analytic class, using this information. Since it is not necessarily the father who is the main breadwinner in a family, the students are allocated to the operational category and analytic class of the parent whose occupation ranks higher on the Standard Occupational Classification 2000 scale. Though this might be argued by some to be problematic (see section 2.3.1.5 for a discussion of Eckert's concerns regarding classifying adolescents according to their parents' occupations), the decision to allocate full-time students to analytic classes in this way is a pragmatic one, based on advice given by the Office for National Statistics, and a desire to examine the possible effects of social class on the language use of as many of the Southampton area informants as possible. Table 3.3 shows the allocation of students in the Southampton area study to analytic classes based on the occupation of their higher ranking parent.

Table 3.3 Allocation of Students to Analytic Classes Based on Occupation of Higher Ranking Parent

Speaker number	Occupation	Father's occupation and corresponding operational category	Mother's occupation and corresponding operational category	Student's analytic class based on operational category of higher ranking parent
4	A-level student	No response	Production line worker (12.4)	3
5	A-level student	Nurse (4.1)	Car mechanic (11.1)	1
6	A-level student	Teaching assistant (12.7)	Electrical engineer (3.1)	1
7	A-level student	Secretary (7.1)	Jazz musician (4.3)	1
8	Undergraduate student	Secondary school teacher (4.1)	Primary school teacher (4.1)	1
18	A-level student	Childminder (9.1)	Instrument technician (11.1)	2
23	Postgraduate student	Personal assistant (7.1)	Electrician (11.1)	2
30	Undergraduate student	Bank clerk (7.1)	Car production worker (12.4)	2
45	Undergraduate student	Nurse (4.1)	Legal practice manager (5)	1

3.9.3 Unclassified Informants

It has not been possible to allocate six of the sixty Southampton area informants to an analytic class. Two of these informants declined to give any information regarding their usual occupation. To press them for this information would have been to risk making them feel uncomfortable. This in turn might have affected the subsequent interview, and so no further questions regarding occupation were asked. The remaining four informants gave their occupation as 'housewife'. This falls into I.16, a residual operational category which is omitted when categories are collapsed into analytic classes. It is not, therefore, possible to allocate a class to informants falling

into this category. However, 90% of informants in the current study have been allocated to an analytic class and their language use can be analysed according to social class.

3.9.4 Analytic Class Sizes

The analytic classes in the current study do not contain exactly the same number of speakers as the classification of speakers according to social class is a retrospective one, informants having been selected to fulfil age, sex and locality quotas (see section 3.8). However, the numeric composition of the three groups is very similar, as Table 3.4 shows.

Table 3.4 Social Class Composition of Southampton Area Study

Analytic class	Number of informants
1 – Managerial and professional occupations	20
2 – Intermediate occupations	18
3 – Routine and manual occupations	16
Not ranked	6

3.10 Data Analysis

The following sections detail the different ways in which the lexical, grammatical and phonological data collected are analysed in subsequent chapters.

3.10.1 Data Analysis - Lexis

Several factors have been taken into account when deciding how best to order and analyse the lexical data collected. It was decided that only those words written by informants on the SRN sheets should be analysed according to the speaker variables of age, sex, social class and identity. It was felt that to include words elicited in the

subsequent interview, either as a result of the suggestions of the fieldworker or during discussion with other informants, might skew the frequency counts of the variants. Those words and phrases offered by informants on the SRNs represent most accurately their own lexical use, or at least their use as they are prepared to report it, uninfluenced by the suggestions and use of others in an interview environment. As the lexical data collected after the completion of the SRNs are not referred to, however, a good deal of useful material offered by informants remains unexamined. It is important, though, to be realistic about the level of detail that can be achieved in a thesis that deals not only with lexis but grammar and phonology as well.

In many cases, informants had employed nonstandard orthography in their SRNs. Examples of this include *settee* spelt *seatee* and *setee*, and *minging* (unattractive) spelt *mingin*. Though of great interest in its own right, not least because it might serve as an indicator of varying pronunciation, investigation of nonstandard orthography is beyond the scope of this study. The spellings have been standardised in order that data quantification should not be affected by variant spellings. It is these standardised spellings which are shown in Appendix 2, which lists, in decreasing order of frequency, all variants recorded by informants on the SuRE SRN sheets. A table of the standardised variants and their nonstandard counterparts is given in Appendix 3. There were also several instances in which two standard spellings of a variant were offered. Examples of this include *granny* and *grannie*, *grand-dad* and *grandad*, *plimsolls* and *plimsoles*. Rather than have two separate frequency counts for what is, for the purpose of this study, one variant, the preferred spelling in the *OED* has been taken as the standardised spelling, and the

other standard spelling subsumed under this. Appendix 3 includes these subsumed standard spellings.

Where the *to*-infinitive form of a verb, e.g. *to play*, has been offered in addition to the bare infinitive form, e.g. *play*, *to* is retained in parentheses in Appendix 2, and the two forms are treated as one variant for which a single count is given. For example, *(to) play* (30) indicates that both the *to*-infinitive form and the bare infinitive form of the verb *play* were collected, and that thirty tokens of the verb were collected in total. As with the nonstandard spellings, this approach was taken in order that lexical data counts are unaffected by variant forms. Unlike the inclusion of multiple variant spellings, however, the retention of *to* and *the* does not reduce the overall accessibility of the data. Where the same noun phrase has been offered both with and without an accompanying verb, for example, *got a bun in the oven* and *bun in the oven*, these two forms have been grouped, *(got a) bun in the oven*. This has only been done, however, when speakers have confirmed that this is the verb with which they would use the noun phrase. The abbreviation *N/R* in the table indicates that no response was offered by the informant.

As Johnson (1996: 80) states, 'lexical items are perceived as behaving in different ways from other levels of language, and as being less amenable to structural analysis'. When attempting to identify patterns in the large number of lexical variants produced by informants in the Southampton area, the difficulties of analysis become apparent. Only certain notion words and phrases, for example, *television* and *main room of the house (with television)*, lend themselves to straightforward quantitative analysis, since these elicit a restricted number of variants, the majority of which are found in high frequencies. In other cases, there are numerous variants but in low frequencies, sometimes only one token, making quantification of these data by

extralinguistic variable problematic. Though these are problems also found in grammatical and phonological data analysis, one suspects that they are particularly pronounced in the case of lexis, where there is so much idiolectal variation. This study takes a two-pronged approach to the analysis of lexical data, examining first the most frequently-occurring variants, which permit straightforward quantitative analysis, and then investigating the remaining tail-end of the data, i.e. those variants, of which there are often many, which occur in very low frequencies.

In some cases, speakers offered more than one word per variable when completing the SuRE SRNs. When analysing these data, all responses have been taken into consideration. In many cases, this results in a difference in the total number of tokens offered by different speaker groups. To display results in tokens could, therefore, be argued to be misleading. Five tokens of a variant out of a total of 30 tokens offered by a particular group of speakers is 16.67% of the total number of responses, whereas 5 tokens of that same variant out of a total of 25 tokens is 20%. If the data are displayed as tokens rather than as a proportion of use by a particular speaker group, the difference in use across groups is not necessarily apparent.

In order to overcome this problem, results are shown as a percentage of the total number of responses given by a particular speaker group, for example, male or female. This percentage is calculated by dividing the number of tokens of a particular variant by the total number of tokens offered and multiplying by 100. For example, eleven different variants and a total of thirty-four tokens were elicited for the notion word *pregnant* from male informants. Their female counterparts offered ten different variants and thirty tokens in total. One of the variants offered by both male and female informants was *preggers*, which was reported by four males and

four females. The percentage use according to the speaker variable of sex for this variant is calculated as follows:

$$\left[\frac{\text{tokens of } \textit{preggers} \text{ reported by particular speaker group}}{\text{total number of tokens offered for notion word } \textit{pregnant} \text{ by the same speaker group}} \right] \times 100$$

In the case of male informants, use of the variant *preggers* is 11.76% ((4/34) x 100), whereas for female informants it is 13.33% ((4/30) x 100), despite the two groups having reported the same number of tokens of the variant.

3.10.2 Data Analysis - Grammar

The Southampton area language questionnaire features forty-eight sentences. Of these, forty-five exemplify nonstandard features and three are control sentences. Speakers were provided with three possible responses to each sentence: that the informant would use the sentence themselves when talking to a friend; that they would not use the sentence themselves but have heard it used by others in the local area; and/or that they would use it themselves when writing to a friend. Appendix 4 shows all the responses given to the Southampton area language questionnaire. However, in chapters 5 and 7, in which a range of grammatical features are analysed, the focus is on the first of the three possible responses; reported informant use of a grammatical feature when talking to a friend. So-called 'double responses', where informants have selected in response to a sentence both option one, *I'd use this when I was talking to a friend*, and option two, *I wouldn't use this but some people do use it in Southampton/Eastleigh Borough*, are discounted, both from Appendix 4 and from analysis by extralinguistic variable.

Many of the grammatical features investigated by the language questionnaire were reported as being used when talking to a friend by informants so infrequently

that there would be no merit in analysing them according to extralinguistic variables. That is not to say that there is no merit at all in examining this collected data. It is certainly worthwhile examining the data as a whole before individually analysing more widely-reported grammatical features to look for telling general patterns. In order to analyse the data as a whole, each speaker is first allocated an individual grammar score to measure their overall reported use of the nonstandard grammatical features under investigation. Informants are given a score of 1 if they claim to employ a particular nonstandard feature when talking to a friend and a score of 0 if they do not. They are then given an overall score out of 45 (the total number of grammatical variables under examination, excluding control sentences). The higher the score, the more nonstandard grammatical forms they claim to use. These scores, when analysed according to speaker group, provide a range of scores for that group which can be compared with those of other groups. Ranges, however, are particularly affected by outlying scores. Overall variation in reported use of nonstandard grammatical variables according to extralinguistic variable is therefore best shown using a mean average for each speaker group. A mean grammar score is calculated by adding together the grammar scores of all informants who fall into a particular speaker group and dividing this number by the total number of informants in that category. A high mean grammar score corresponds to high reported use of nonstandard grammatical variables, while a low mean grammar score corresponds to low reported use of nonstandard grammatical forms which, by extension, equates to a claim to use standard grammatical forms.

Though of great interest when considered en masse, individual grammatical variables with low levels of reported use are unsuitable for examination by extralinguistic variable, since there are too few tokens of their reported use to

subdivide by speaker group. Meaningful conclusions cannot be drawn from a comparison of two reports of a feature from one speaker group with one token of the same feature from another group. However, given the stigmatised character of many of the features under examination, very few have high levels of reported use. Only three variables, for example, were reported as being used when talking to a friend by more than thirty speakers. Very high reported use was therefore deemed to be unrealistic grounds on which to select features for potential analysis, because the nonstandard nature of the features means that many people were unwilling to report their use of them.

Taking into consideration the stigmatised nature of the features under examination and the need for a sufficient number of tokens to produce meaningful analysis, the pragmatic decision was made to consider for examination by extralinguistic variable only those features reported by more than ten speakers as being used when talking to a friend. The application of this cut-off point provides both a reasonable breadth of variables from which to select the most salient for extralinguistic analysis, and enough tokens of these variables, once selected, to provide confidence in conclusions drawn from them. In total, seventeen of the forty-five nonstandard grammatical features examined by means of the Southampton area language questionnaire were reported by more than ten speakers. Features for analysis by extralinguistic variables have been selected from these seventeen, and are detailed in section 3.13.

Where grammatical data are analysed by sex and/or by age, the data are presented as tokens of reported use. One token is equivalent to one speaker's reported use of a particular feature. Where grammatical data are analysed by identity and/or by social class, however, reported use of a grammatical form by a particular

identity or social class group is presented as a percentage of the total number of informants having that particular level of identity/belonging to that particular social class. This is because the social classes and identity groups in the present study do not contain comparable numbers of speakers. For example, *that ain't a bird* is reported by 5 speakers with medium identity and 6 with high identity. It is not reported by anyone with low identity. The percentage use according to speaker identity for this variable is calculated as follows:

$$\left(\frac{\text{number of speakers reporting } \textit{that ain't a bird} \text{ in a particular identity group}}{\text{total number of speakers in that same identity group}} \right) \times 100$$

In the case of speakers with medium identity, use of *that ain't a bird* is 21.74% ((5/23) x 100), whereas for those with high identity, it is 19.35% ((6/31) x 100), despite more speakers from the latter group reporting their use of the variable. (*Ain't* and speaker identity is discussed at greater length in section 7.4.1). Presented by number of informants reporting the form, it appears that *that ain't a bird* is more likely to be used by a speaker with high identity than one with medium identity as it is reported by 5 speakers with medium identity and 6 with high identity. However, it could be the case that *that ain't a bird* is reported by more speakers with high identity than those with medium identity simply because there are more speakers in the former group than in the latter (31 compared to 23) rather than because of the degree of local identity felt by the speakers.

In chapter 5, in addition to analysis of reported use of four nonstandard grammatical variables according to the extralinguistic variables of sex, age and social class, a comparison of reported use with actual speaker use of a further two grammatical variables is also undertaken.

3.10.3 Data Analysis - Phonology

Milroy states:

Guy suggests that 30 tokens per variable is a reasonable goal to aim for. As he points out, N=30 is an important dividing line in statistics generally between large and small samples. [...] In fact, the data presented by Guy seem to conform to general statistical laws: if the number of tokens is lower than 10, there is a strong likelihood of random fluctuation, while a figure higher than 10 moves towards 90 per cent conformity with the predicted norm, rising to 100 per cent with 35 tokens.

(Milroy 1987b: 135)

With this in mind, fifteen tokens per speaker of the phonological features to be analysed by extralinguistic variables have been collected. The exception to this rule is the BATH vowel, for which as close to 15 tokens per speaker as possible have been collected. Further information on data collection and the BATH vowel can be found in section 6.0. Milroy (1987b: 135) argues that 'a figure as much as possible in excess of 10 is a sensible goal'. Again in line with Milroy's (1987b: 135) advice, these fifteen tokens are not subdivided to examine linguistic environment. Rather, tokens of a feature in the respective linguistic environments detailed in section 1.7.3 are considered as one. Results are shown as a percentage of the total possible number of tokens of a particular variable according to speaker group, for example male or female. This percentage is calculated by dividing the number of tokens of a particular variant offered by the total possible number of tokens, and multiplying by 100. For example, male informants offer a total of 313 tokens of L Vocalization out of a possible 450. Female informants offer 287 out of 450. The percentage use according to the speaker variable of sex for this variable is calculated as follows:

$$\left(\frac{\text{tokens of L Vocalization reported by particular speaker group}}{\text{total possible number of tokens of L Vocalization from the same speaker group}} \right) \times 100$$

In the case of male informants, L Vocalization is 69.56% $((313/450) \times 100)$ whereas for female informants, it is 63.78% $((287/450) \times 100)$. This means of data analysis offers ease of comparability. 69.56% and 63.78% are easier to compare than 313 and 287 as readers do not have to constantly be reminded of the overall scale. In the case of the extralinguistic variables of social class and identity, this method of data analysis also allows the use of a phonological feature to be compared across speaker groups of unequal sizes, something which is not possible if tokens are used.

3.11 Analysis of Data by Multiple Extralinguistic Variables

In some instances, data are analysed according to more than one extralinguistic variable simultaneously, for example the BATH vowel according to both speaker sex and age (see section 6.3.3). Social variables can interact in complex ways, and such analysis sometimes reveals patterns that are not apparent when analysis is conducted by one speaker variable at a time. However, some data are more amenable to this type of analysis than others are. Where the number of tokens of a variant is limited, to attempt to subdivide those tokens by multiple extralinguistic variables means that the number of tokens in each resultant speaker cell is very small. It is therefore difficult to draw meaningful conclusions regarding the influence of these extralinguistic variables. In the present study, this is a particular issue with the lexical and grammatical data collected, since the overall number of tokens is, in the main, small to begin with. It is far more telling to analyse the phonological data by multiple extralinguistic variables simultaneously. The large number of tokens is better suited to this type of analysis, as the resultant speaker cells contain more tokens, making the conclusions drawn from them more reliable.

3.12 Statistical Testing

In instances where differences in the use of a particular linguistic variable between speaker groups are small, statistical testing is undertaken. Advice has been sought from Dr. Paul D. Baxter, Lecturer in Statistics in the Department of Statistics,

University of Leeds, as to which test would be most appropriate for the Southampton area data. Dr. Baxter recommended the chi-squared test of independence be used to ascertain the statistical significance of this data. The formula for this test is:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

where O_i is observed frequency, E_i is expected frequency, and k is the number of categories being compared.

The Yates correction prevents the overestimation of statistical significance for small amounts of data. Though ‘a bone of contention’ (Upton and Cook 1996: 500), use of this correction is often recommended, and it has been applied in the present study to 2 x 2 contingency tables and in cases where expected frequencies are lower than 5. The formula for the Yates correction is:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i - 0.5)^2}{E_i}$$

Significance testing is carried out in most instances at the 95% level, as this is ‘the minimum generally used in the social sciences’ (Davis 1983: 76). Where the 95% level is not used, it is clearly stated at what level significance testing is being conducted.

3.13 Selection of Linguistic Variables for Examination

It is impracticable in a study of limited size to analyse all linguistic variables found in rich data according to the speaker variables of sex, age, social class and identity. Consequently, a selection of lexical, grammatical and phonological features is presented in chapters 4 to 7. Many features are analysed according to more than one extralinguistic variable, and according to multiple extralinguistic variables simultaneously, in order to ascertain the relative importance of these speaker variables on the use of the linguistic features.

The linguistic variables examined in these chapters have been selected in a variety of ways. In many cases, patterns became immediately apparent during the course of the interview, when it was clear that a particular variant was being used, or was reported as being used, more by one group of speakers, be it age-, sex-, class- or identity-based, than by another. Also, the comments made by informants have proved to be invaluable in selecting linguistic features on which to focus. In the case of lexis, informants remarked on particular words being 'old-fashioned' or 'only used by young people'. Similarly, when completing the language questionnaire, respondents reported certain grammatical constructions as being used by particular age groups, mainly the young, or by males rather than females. Speakers also commented on certain phonological features when asked to read the word list. These comments highlighted the salience of particular features in the sampling universe and provided an indication that they are perhaps worthy of further attention. Native speaker intuition, used by Trudgill (1974: 80) to select phonological variables in his original Norwich survey, has also been employed in the Southampton area when selecting linguistic variables for examination. An attempt has been made to select

linguistic variables for analysis which could be argued to be representative of other features on which data has been collected.

With these factors in mind, the following range of lexical, grammatical and phonological variables have been selected for analysis by extralinguistic variable. In the case of lexis, the variants *pregnant*, *toilet*, *grandmother* and *grandfather*, *long seat in main room of house* and *left-handed* have been chosen. The notion word *pregnant* was selected as, pregnancy being a female condition, it was thought that the notion word might exhibit signs of sex-based variation. It can be grouped with the notion phrase *men's facial hair* and argued to be representative of sex-specific terms. *Pregnant* was also a notion word which provoked a good deal of discussion among informants. *Toilet* was chosen on the grounds that there exist a number of stigmatised or taboo forms associated with this word, use of which might vary according to the speaker variables of age, sex and social class. Other notion words which have stigmatised or taboo forms associated with them include *to vomit* and *drunk*. The notion words *grandmother* and *grandfather* were selected as it was thought that they might exhibit signs of sex-based variation, being terms specifically for women and men respectively. Similar sex-specific notion words and phrases are *mother* and *father*, *brother* and *sister*, *man* and *woman*, *child (boy/girl)* and *partner (sexual) male female*. *Long seat in main room of house* was chosen as a neutral term, that is as a term which is not stigmatised or specific to either men or women, to see how it might correlate with the extralinguistic variables under examination. *Left-handed* was selected as it was hoped it would exhibit signs of age-based variation, left-handedness perhaps having been considered as more worthy of note in the past than it is now.

The grammatical features chosen for analysis by extralinguistic variables are: *ain't*; *there was* with a plural notional subject; nouns of measurement with zero plurals; nonstandard prepositions; adverbial forms; and present tense verb morphology. *Ain't*, like multiple negation, was frequently commented on unfavourably by speakers, and was chosen to see how, as an apparently stigmatised form, its reported use correlates with extralinguistic variables. By contrast, *there was* with a plural notional subject is not apparently subject to such overt stigmatisation, and is analysed as a seemingly neutral form. Nouns of measurement with zero plurals, for example *that town is nearly twenty mile away*, are a traditional dialect form (Wright 1905: 263; Upton and Widdowson 1996: 73) and were chosen to see how they correlate with speaker age. The nonstandard prepositions chosen for examination fall into two categories: simple prepositions where Standard English has complex prepositions, *I'm going up/down/over/round my friend's house later*; and a complex preposition where Standard English has a simple preposition, *he knocked his hat off of his head*. They were chosen to see if there is a difference in use of the simple and complex forms according to the speaker variables of sex, age and social class, *off of* apparently being subject to more negative comments among Southampton area speakers than its simple counterparts. Adverbs and present tense verb morphology have been selected as they occur in high frequencies in speaker interviews, thus permitting a comparison of speaker reported and actual use of these forms.

L Vocalization, rhoticity, and the BATH and PRICE vowels have been chosen as the phonological variables for examination by sex, age, social class and identity. Though L Vocalization, rhoticity and the BATH vowel have been selected primarily as

features which might inform the east-west identity debate, it was also thought that all four phonological features would vary with speaker age, sex and social class.

Chapters 4 to 7 thus provide an insight into the ways in which a selection of lexical, grammatical and phonological variables pattern with the extralinguistic variables of age, sex, social class and identity.

Chapter 4

4.0 Lexis

In the analysis of the potential relationships between lexical use and the extralinguistic variables of sex, age and social class, the two-pronged approach discussed in section 3.10.1 is used. The most frequently occurring variants which permit straightforward quantitative analysis are analysed first, and then the remaining tail of the data, those variants which occur in very low frequencies, is investigated. Chi-squared testing (see section 3.12) has been undertaken when differences in reported use of a variant between speaker groups is small.

As discussed in section 3.8, the Southampton area study has 30 male and 30 female informants, and 20 speakers in each of the three age groups (under 30, 30 to 59, and 60 and over). The speaker sample is comprised of 3 social class groups, class 1 being the highest, and class 3 the lowest. Class 1 has 20 informants, class 2 has 18 informants, and class 3 has 16 informants (see section 3.9 for a discussion of the social stratification of Southampton area speakers).

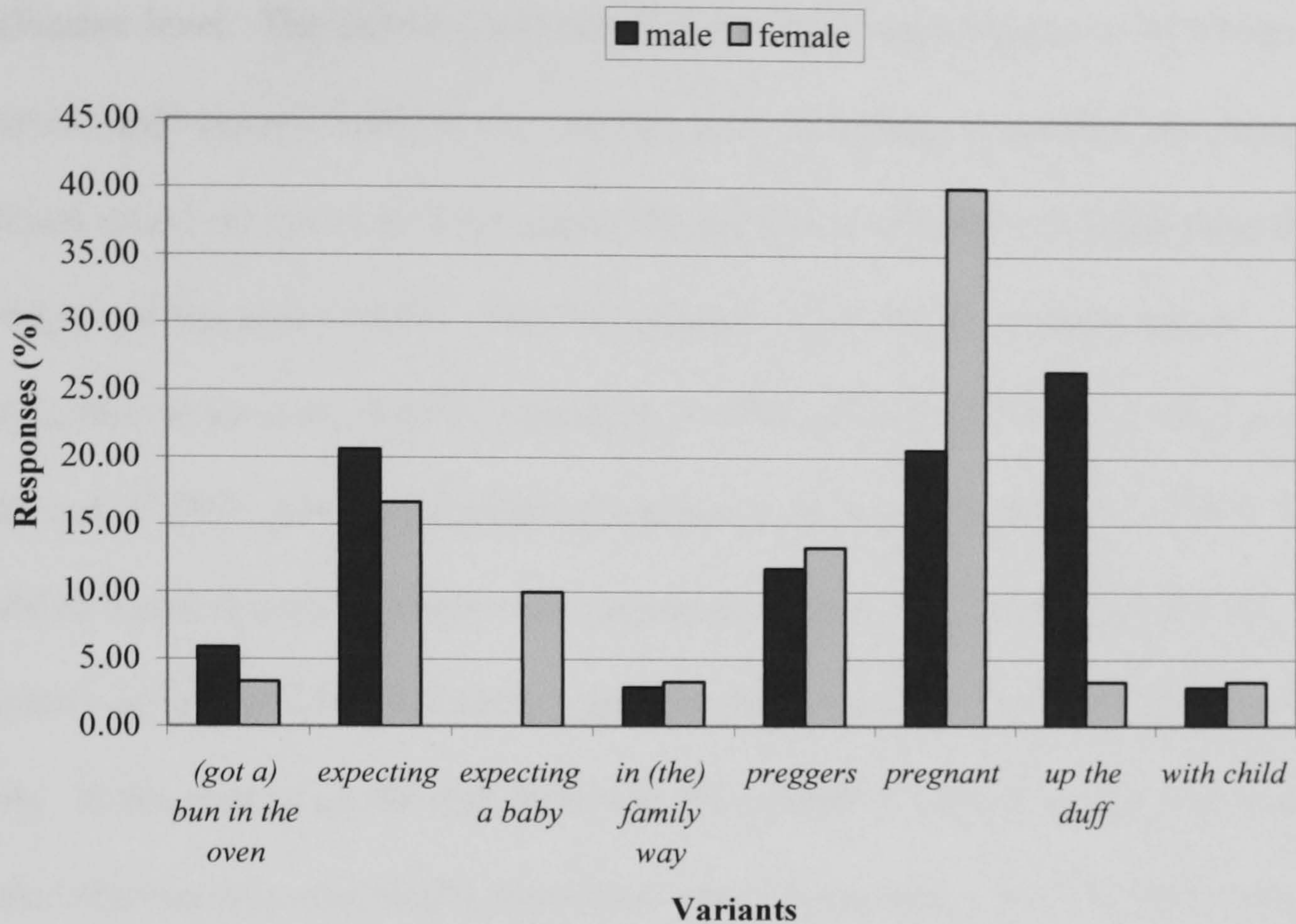
4.1 Pregnant

Thirteen different terms were elicited for the notion word *pregnant*, of which eight occurred more than once. In the following sections, *pregnant* is analysed according to the speaker variables of sex and social class.

4.1.1 Pregnant and Speaker Sex

Figure 4.1 shows the most frequently occurring variants for the notion word *pregnant*, according to the extralinguistic variable of speaker sex.

Figure 4.1 Most frequently reported variants for the notion word *pregnant* according to the speaker variable of sex



Raw data

variant	male	female
<i>(got a) bun in the oven</i>	5.88% [n = 2/34]	3.33% [n = 1/30]
<i>expecting</i>	20.59% [n = 7/34]	16.67% [n = 5/30]
<i>expecting a baby</i>	0.00% [n = 0/34]	10.00% [n = 3/30]
<i>in (the) family way</i>	2.94% [n = 1/34]	3.33% [n = 1/30]
<i>preggers</i>	11.76% [n = 4/34]	13.33% [n = 4/30]
<i>pregnant</i>	20.59% [n = 7/34]	40.00% [n = 12/30]
<i>up the duff</i>	26.47% [n = 9/34]	3.33% [n = 1/30]
<i>with child</i>	2.94% [n = 1/34]	3.33% [n = 1/30]

Pregnant is the preferred term for female informants, while *up the duff* is favoured by male informants. However, with the exception of *up the duff*, differences in use between males and females of the most frequently occurring

variants for the notion word *pregnant* are small. All variants have been tested for statistical significance using the chi-squared test of independence at the 95% significance level. The difference in use between males and females of all variants bar *up the duff* was not found to vary significantly according to speaker sex, since the observed values do not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). The chi-squared values for the variants are as follows: *bun in the oven* (0.012); *expecting* (0.006); *expecting a baby* (1.680); *in the family way* (0.397); *preggers* (0.036); *pregnant* (2.022); and *with child* (0.397). It should be noted that the variants *expecting* and *expecting a baby* are considered separately as they differ stylistically, *expecting* being more colloquial than *expecting a baby*. In the case of *up the duff*, however, the difference in use between male and female informants is statistically significant (the chi-squared value is 4.835), with men being more likely to use this form than women.

The remaining tail of the data comprises five variants, each of which is reported only once. Three of these are reported by male informants and two by females. The female informants report the variants *fall pregnant* and *having baby*, while the males report *bundle of joy*, *in the club* and *up the gut*. Of the male responses, *bundle of joy* is arguably an anomalous variant, being used usually as a term for a baby rather than for pregnancy.

Though only use of *up the duff* differs significantly according to speaker sex, *pregnant* was a notion word which produced a wide range of stylistically-marked variants, and which sparked a good deal of discussion among informants. In many cases, speakers made it very clear that they would only use their reported terms in certain domains. Use was dependent on the speaker's relationship with the pregnant

woman and with the addressee. The only female informant to report the term *up the duff* gave the following explanation of its use (KW is the fieldworker Kate Wallace):

KW: [What words have you got for] pregnant?

Jl: Up the duff, preppers, bun in the oven.

KW: And what sort of situations would you use those in? Would you use them to the person who was preppers or whatever they are? Or are they...

Jl: It depends how well we know them. If it was a mate, I'd be like... If, you know, you're messing about with them you'd be, yeah, "You're preppers" or "She's up the duff" but if it was somebody you'd just met or you didn't know very well, you'd probably just use the word pregnant.

In the case of this respondent, a 23-year-old female from Southampton, *up the duff* is used jokingly with friends and would not be used in a more formal situation with someone who was less well known to the respondent. It does not necessarily have negative connotations, though it is clearly informal.

The effect on language use of the speaker's addressee is also raised by a 56-year-old male from the Southampton area:

KW: And [what are your words for] pregnant?

PW: This all depends on what company you are [sic].

KW: OK. Tell me how it differs.

PW: Because some of them are rather crude in a sense. In the club is, you know, is a nice one. I think. The one I've heard is up the gut.

KW: Oh right.

PW: That's not very nice but that's what I've heard and that's what I have used, you know. [...] Bun in the oven, that's a sort of a friendly one. Sometimes if you're talking to someone, "Are you in the club?" But if you're talking about someone, "She's up the gut."

Like the 23-year-old female, this respondent's lexical choice differs according to the company in which he finds himself. His use of *up the gut* in particular is reserved for a certain audience. This speaker's positive evaluation of the variant *bun in the oven* is not shared, however, by a 78-year-old woman from Eastleigh Borough:

KW: What about words for pregnant, do you think?

LF: Well, the crude way is bun in the oven, isn't it?

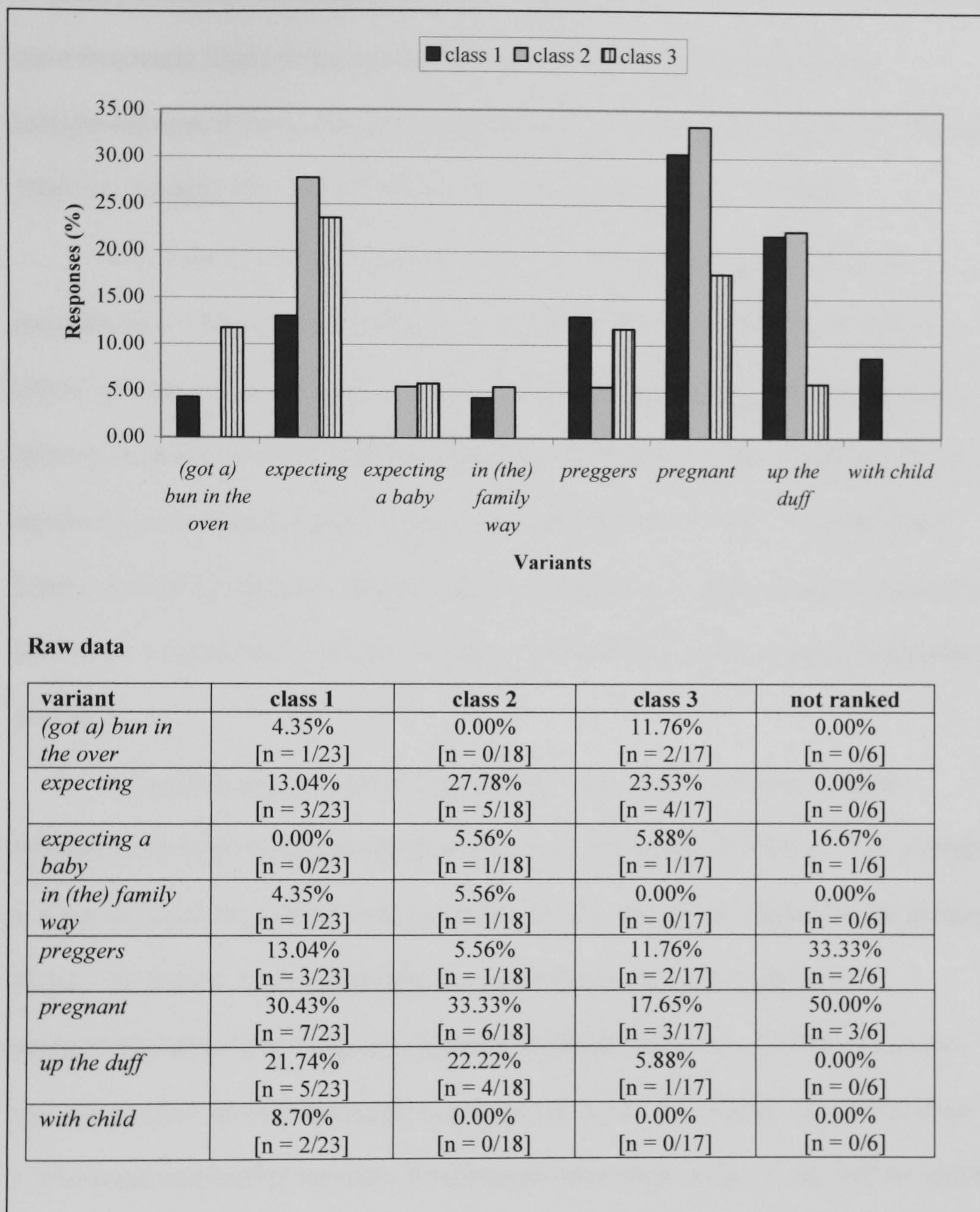
Though *up the duff* is the only variant for which there is a statistically significant difference between male and female reported use, it nonetheless appears that terms for *pregnant* do differ according to speaker sex. Thirteen different variants are offered in total by informants for this notion word and both male and female informants employ the same number of variants, ten each. However, three are used exclusively by men (*bundle of joy, in the club, up the gut*) and three solely by women (*expecting a baby, fall pregnant, having baby*). Males most frequently report using the highly colloquial *up the duff* while females most frequently report the more formal *pregnant*. Though both male and female informants report the effect of audience on their lexical choice, their judgements as to what are 'nice' or 'crude' terms for *pregnant* do not necessarily correspond, as in the example of *bun in the oven*.

4.1.2 Pregnant and Speaker Social Class

Figure 4.2 shows the variants reported more than once in response to the notion word *pregnant*, according to the extralinguistic variable of social class.

The differences in reported use of variants for the notion word *pregnant* between social class groups were not found to be statistically significant when chi-squared testing was undertaken, since the observed values do not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). The chi-squared values are as follows: *bun in the oven* (0.758); *expecting* (0.624); *expecting a baby* (0.150); *in the family way* (0.150); *preggers* (0.128); *pregnant* (0.501); *up the duff* (0.993); and *with child* (0.690).

Figure 4.2 Most frequently reported variants for the notion word *pregnant* according to the speaker variable of social class



Though not statistically significant, the reported use of different variants by the three social classes is nonetheless of interest. *Pregnant* is the variant preferred by classes 1 and 2, while *expecting*, described by the *OED* as colloquial, is the term most frequently reported by those in class 3. As *pregnant* is a formal term for having

offspring developing in the uterus, it is perhaps unsurprising that it is found more frequently in classes 1 and 2 than in class 3, given that standard or prestige forms are more frequently found in the speech of people from a higher socioeconomic background than in the speech of those from a lower socioeconomic class (cf. Labov 1966: 73, Trudgill 1974: 61-63, Wells 1982: 14, Chambers 2003: 57-58).

(Got a) bun in the oven is most frequently reported by class 3 speakers. It accounts for 4.35% of the responses from class 1 speakers, and is not reported by class 2 speakers. Reported use of both *expecting* and *expecting a baby* is higher for informants in classes 2 and 3 than it is for class 1 speakers. *In the family way* is not reported by any class 3 speakers, and accounts for 4.35% of class 1 responses and 5.56% of class 2 responses. Reported use of *preggers* is highest amongst informants from class 1, followed by those from class 3. *With child* is only reported by speakers in class 1.

It is noteworthy that *up the duff*, a slang term (*OED*) reported by some informants as only being appropriate with a restricted audience, namely close friends, is reported nearly four times more frequently by the middle and highest social groups than by the lowest. One explanation for this is that lexis is less subject to stigmatisation than other linguistic features (Johnson 1996: 81). It is likely too that age and sex have an important part to play in the use of this variant. All of the class 1 informants and half of the class 2 informants who report using *up the duff* are under 30, an age group for whom '[i]nfluence from the standard language is relatively weak' (Chambers and Trudgill 1998: 79). All informants in class 2, and all bar one in class 1 are also male, and males have been found to use fewer prestige or standard forms than their female counterparts (Chambers 2003: 116).

Five variants were reported only once in response to the notion word *pregnant*. A speaker in class 1 reported *fall pregnant*, while speakers in class 3 reported *bundle of joy* (presumed to be an anomalous response), *having baby*, *in the club* and *up the gut*. In total, informants in class 1 reported 8 variants, speakers in class 2 reported 6 variants, and speakers in class three reported 9, discounting *bundle of joy*.

The variety of terms reported by informants indicates that there is rich lexical variation in words and phrases for the notion word *pregnant* in the Southampton area. All three social class groups report using a range of terms which differ stylistically from the formal (*pregnant*, *with child*, for example) to the colloquial (*preggers*, *up the duff*, *up the gut*, for example). The latter group includes terms which speakers themselves acknowledge to be potentially offensive, or appropriate only in restricted contexts (see section 4.1.1). That such a broad range of terms has been collected is testament to the efficacy of the SuRE methodology, which, by lessening the formality of the interview as much as possible, encourages speakers to report all examples of lexical variation, including those which might be negatively perceived in certain situations. As discussed, some findings, such as the higher reported use of *up the duff* by those from classes 1 and 2 than by class 3, are surprising, and may be the result of influence from other extralinguistic variables such as speaker age and sex. Similarly, it might be that I, as a young, middle class female interviewer, had an influence on the variants that informants saw fit to report. It is clear, however, that all informants, regardless of social class have at their disposal a number of variants suitable for a variety of audiences and contexts.

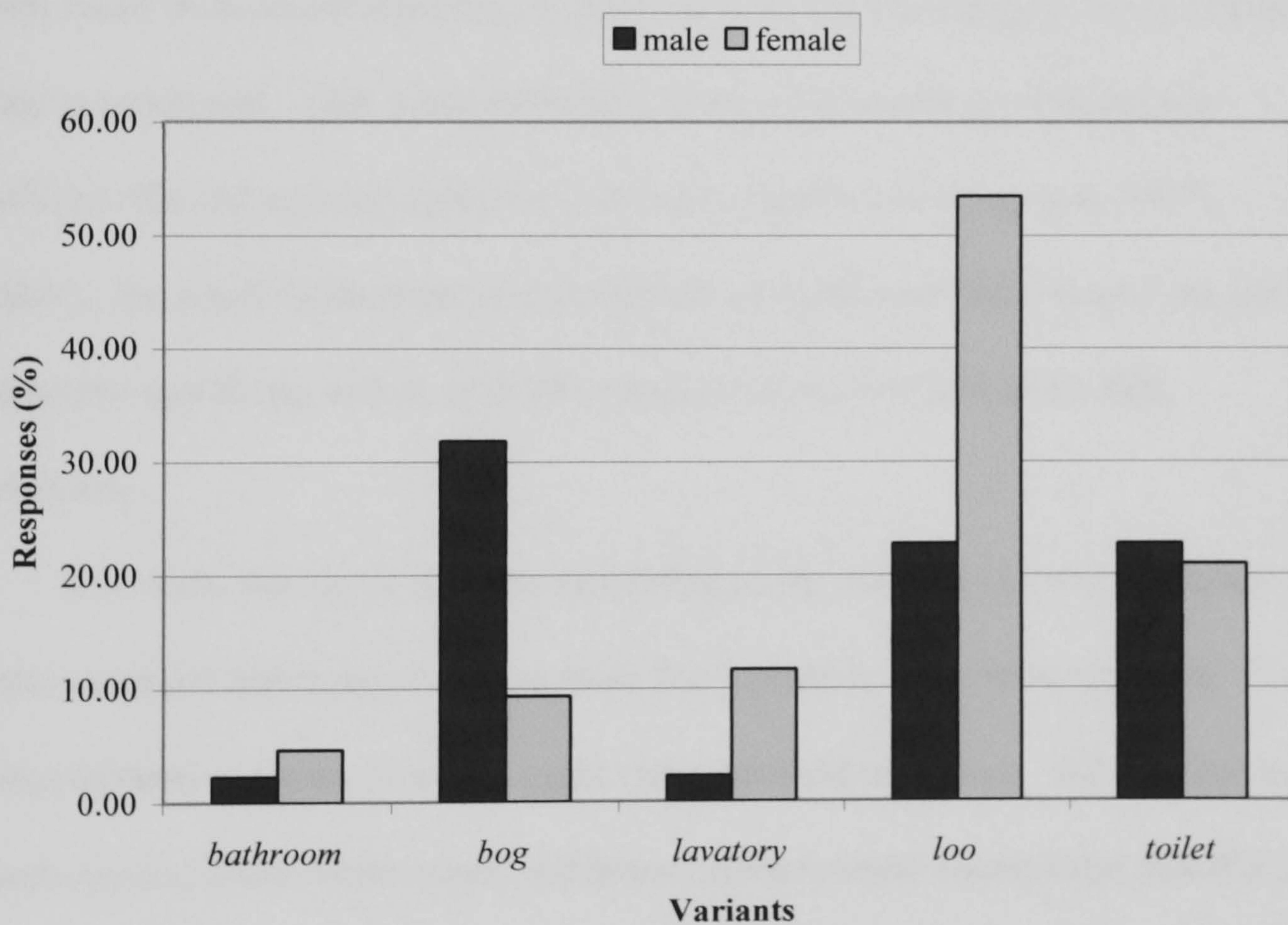
4.2 Toilet

Thirteen different terms were elicited for the notion word *toilet*. Of these, five occurred more than once. In the following sections, *toilet* is analysed according to the speaker variables of sex, age and social class.

4.2.1 Toilet and Speaker Sex

The most frequently occurring variants for the notion word *toilet* are shown in Figure 4.3, according to the extralinguistic variable of speaker sex.

Figure 4.3 Most frequently reported variants for the notion word *toilet* according to the speaker variable of sex



Raw data

variant	male	female
<i>bathroom</i>	2.27% [n = 1/44]	4.65% [n = 2/43]
<i>bog</i>	31.82% [n = 14/44]	9.30% [n = 4/43]
<i>lavatory</i>	2.27% [n = 1/44]	11.63% [n = 5/43]
<i>loo</i>	22.73% [n = 10/44]	53.49% [n = 23/43]
<i>toilet</i>	22.73% [n = 10/44]	20.93% [n = 9/43]

It is apparent from Figure 4.3 that the preferred term for *toilet* for female informants is *loo* and for males, *bog*. The difference in reported use between males and females of *bog* is significant at the 95% level, as the chi-squared value is 5.417, which exceeds the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). The differing reported use by males and females of *loo* is also statistically significant, with a chi-squared value of 7.483. The differing use of *bog*, a term described by the *OED* as slang, by males and females appears to support the assertion that women employ fewer nonstandard or stigmatised forms than do their male counterparts (Chambers 2003: 116).

By contrast, there is little sex-based variation apparent in the use of the variant *toilet*, with males reporting employing *toilet* only marginally more than their female counterparts. This small difference in reported use is not statistically significant, the chi-squared value for *toilet* and speaker sex being only 0.003. Similarly, the small differences in reported use of *bathroom* and *lavatory* are not statistically significant either, with chi-squared values of 0.000 and 1.686, respectively.

It is when the tail of the data pertaining to the variable *toilet* is examined that the most marked difference between male and female lexical use is apparent. *Toilet* elicited thirteen variants, of which eight were reported only once. All of these eight variants (*gents*; *khazi*; *little room*; *shit house*; *shitter*; *small room*; *trap*; and *W.C.*) were reported by men. The first point worthy of note is that men reported nearly three times as many words for *toilet* as their female counterparts, thirteen compared to five. The second point concerns the nature of some of the variants reported. That *gents* was reported only by a male informant is easily explicable, since it refers to a toilet for men. It is worthy of note, however, that the terms *shit house* and *shitter*

were reported only by male informants. The word *shit* is described by the *OED* as ‘not now in decent use’. That no female informants offered variants for *toilet* derived from the word *shit* could either mean that they do not use such terms or that they did not consider these terms appropriate for SuRE, despite the fact that the Survey places no restrictions on the types of words it asks speakers to record and, through its methodology, seeks to lessen the formality of the interview context.

If it is the case that some female informants deemed certain lexis unsuitable for SuRE, it could equally be the case that more male informants than the two who reported *shit house* and *shitter* use such variants for *toilet* outside the SuRE interview context. As in the case of the notion word *pregnant*, many speakers were very clear about the domains in which they would use their reported terms. When asked for his words for *toilet*, RS, a 21-year-old male from Eastleigh Borough replied, ‘It’s the bog or... The shitter, I use, but only with my friends’. AS, a 50-year-old male from Southampton, and his 40-year-old wife, SS, also reported the importance of context on lexical choice:

AS: Well I either call it the loo, the bog or the little room.

SS: [...] I normally say, “I’m going to the loo”.

Occasionally, I say, you know, “I’m going to the toilet”. I think it sometimes depends...

AS: Who’s here and again who you’re talking to. With grandmother, “I’m going to the toilet” but otherwise, “I’m just going to the loo”.

In addition to the possible influence of the interview context on reported lexical use, the effect of the fieldworker on responses given must also be considered. It is possible that, as a young, middle class female, I might have influenced the lexical variants offered by informants. Silverman states:

Your gender in relation to the gender of the people you are studying may turn out to be very important in relation to how you are defined and, therefore, what you find out.

(Silverman 2001: 59)

He goes on to suggest that variables like the age and social class of the researcher might also be significant when conducting fieldwork (Silverman 2001: 60). It was my impression that a minority of the younger males interviewed were trying to shock me with the lexical variants they reported. In contrast, it might have been the case that some older male respondents were concerned that certain lexical variants might cause me offence, and so restricted the terms which they reported. I had certainly not expected to record variants such as *little room*, *small room* and *W.C.* from male informants but not from their female counterparts.

Whether entirely representative of everyday usage or not, the data collected by the SuRE SRNs are still very telling with regard to the differences in lexical use between the sexes. Female respondents reported a relatively limited set of variants for *toilet* compared to their male counterparts, thirteen compared to five. The data show a difference in the way in which male and female informants claim to use lexis in an interview. Whether it is actually the case that female informants do not have such a wide range of variants for the notion word *toilet* as men, or that they do not employ terms such as *shitter* and *shit house*, is unclear from the SuRE data, as they are concerned with reported rather than actual usage. This, however, is not the most important issue.

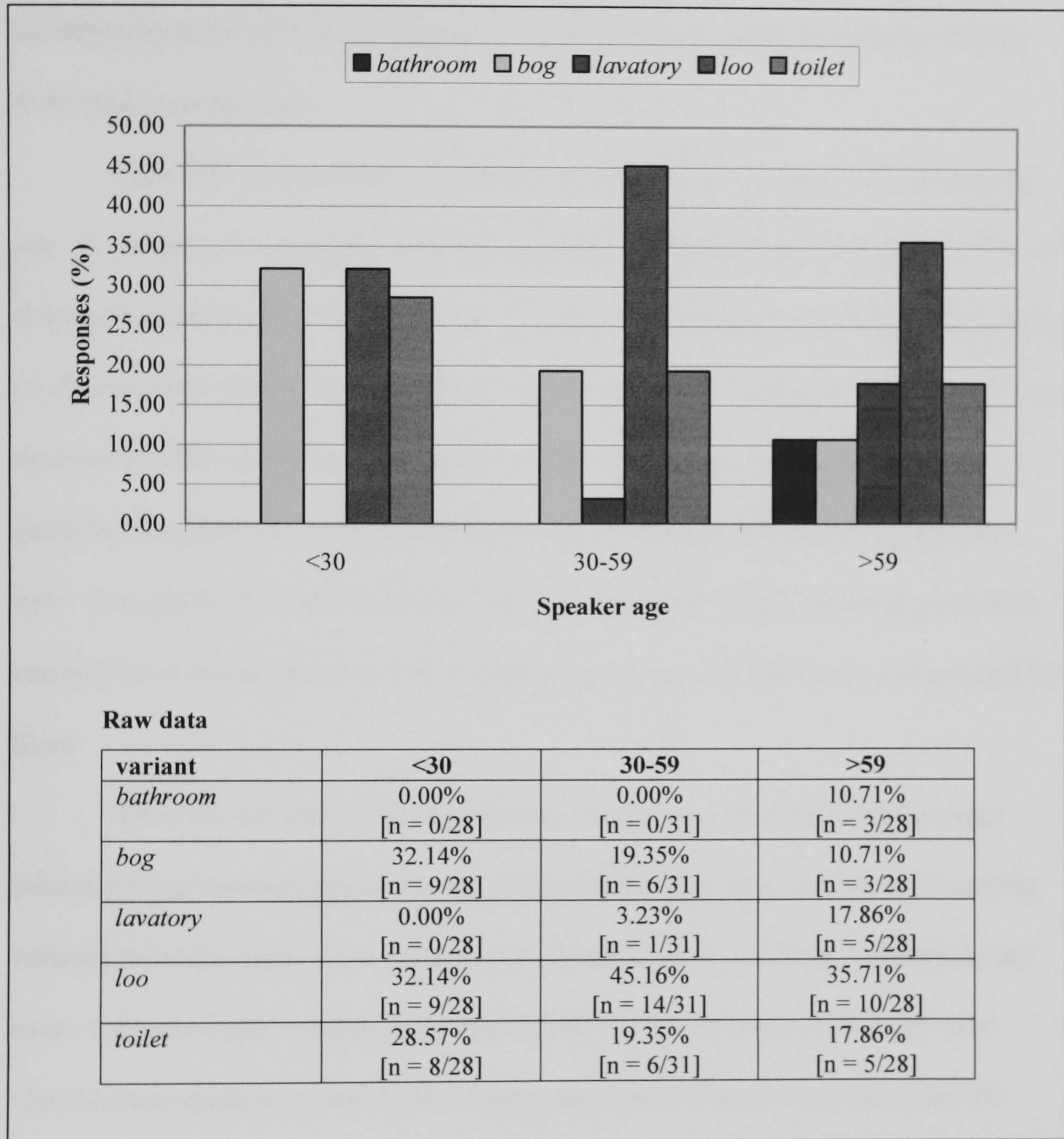
4.2.2 Toilet and Speaker Age

Figure 4.4 shows the most frequently occurring variants for the notion word *toilet* according to the extralinguistic variable of speaker age.

The differences in reported use of the most frequently occurring variants for *toilet* between speakers of different ages are not statistically significant at the 95% level, as none of the chi-squared values exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). The chi-squared values for each of

the notion words are as follows: *bathroom* (3.072); *bog* (3.970); *lavatory* (5.014); *loo* (1.145); and *toilet* (1.116).

Figure 4.4 Most frequently reported variants for the notion word *toilet* according to the speaker variable of age



Though not statistically significant, differences in reported use of variants across age groups are nonetheless noteworthy, since it is not perhaps how many tokens of a variant that are elicited, but rather the fact that variants are elicited at all from particular age groups that is of primary importance. It is apparent from Figure 4.4 that the most frequently reported variant for *toilet* for speakers in both the 30 to

59 and 60 and over age groups is *loo*. In the under 30 age group, *loo* and *bog* are the most frequently reported variants. As discussed in section 4.2.1, *bog* is described by the *OED* as slang. Since use of nonstandard or stigmatised features is often associated with the young (Labov 1972a: 257; Chambers and Trudgill 1998: 79), it is unsurprising that *bog* is more frequently reported by the youngest speakers than by their older counterparts.

Use of the terms *lavatory* and *bathroom*, however, increase with increasing age. *Lavatory* is not reported at all by the under 30s and accounts for only 3.23% of responses given by 30 to 59-year-olds. However, it makes up 17.86%, nearly a fifth, of all responses offered by informants aged 60 and over. Similarly, *bathroom* is not reported at all by speakers in the under 30 and 30 to 59 age groups in the current study but accounts for 10.71% of all responses offered by informants aged 60 or over. It might be the case that these terms are considered old-fashioned, given their use by older informants, or overly formal by younger speakers and so are avoided by them.

When the tail-end of the data for this variable is examined, a further age-related pattern becomes apparent. In addition to the five most frequently occurring variants for *toilet*, eight more variants were elicited, each of which is reported only once. Of these eight variants, four (*khazi*, *little room*, *shit house* and *trap*) were elicited from speakers in the 30 to 59 age group and two each from the under 30 group (*gents* and *shitter*) and those aged 60 and over (*small room* and *W.C.*). When these results are analysed alongside the more frequently occurring variants, it is evident that it is the 30 to 59 age group from which the largest number of different variants is elicited. In total, they claim to employ eight terms, compared to the seven used by those aged 60 and over and the five used by the under 30s.

A further point arising from an examination of the tail-end of the data elicited by *toilet* relates to the nature of the variants recorded. Discussed in 4.2.1 in relation to the notion word *toilet* and the speaker variable of sex are those variants which derive from the word *shit*. In 4.2.1, it was established that *shit house* and *shitter* are reported by male informants. When the age profiles of these two informants are examined, it is apparent that they are from the younger two of the three age groups. *Shit house* is reported by an informant in the 30 to 59 age group while *shitter* is reported by a respondent in the under 30 group. Though it might be the case that these variants are used by informants in the 60 and over age group, none report using them.

It is likely, as discussed in section 4.2.1, that I, as a young, middle class female, influenced the lexical variants offered by informants (Milroy 1987b: 50; Silverman 2001: 59). It is perhaps the case that some older male informants did not wish to cause me embarrassment and so restricted their reporting of terms which they perceived might offend me. This certainly appears to be a viable suggestion when the age profiles of users of the terms *shit house* and *shitter* are compared to those of the informants employing the variants *little room* (reported by a male in the 30 to 59 age group), *small room* and *W.C.* (both reported by males in the 60 and over age group). The terms derived from the word *shit* (a noun described by the *OED* as 'not now in decent use') are used by speakers from the youngest and middle age groups while the inoffensive *little room*, *small room* and *W.C.* are employed by speakers from the middle and oldest age groups. Fieldworker influence need not only affect older male speakers, however. The smaller number of variants for *toilet* offered by younger informants might be explained by an unwillingness on their part to report the terms which they usually employ, for fear of causing me offence.

The question has already been raised in section 4.2.1 as to how accurately the SuRE lexical findings might reflect actual use. It should be reiterated that the matter of strict representativeness of reporting is not the most important issue. What is of key importance is that data on *toilet* collected in the context of the SuRE interview show a marked difference in the way in which speakers of different ages claim to use lexis, not only in the case of the tail-end of the data but in the most frequently occurring variants as well.

4.2.3 Toilet and Speaker Social Class

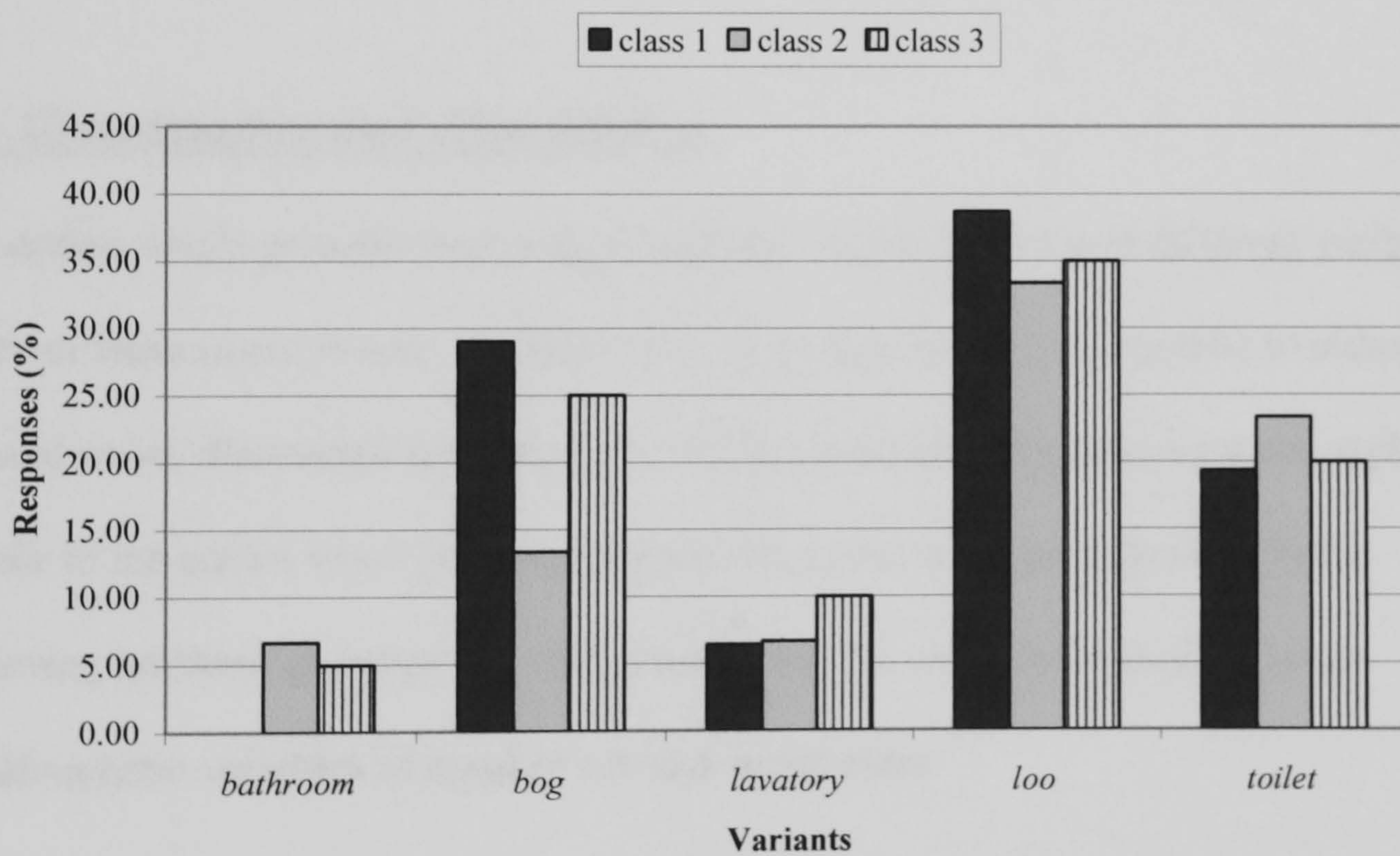
Figure 4.5 shows the variants reported more than once in response to the notion word *toilet*, according to the speaker variable of social class.

The variant *loo* is the preferred term for all three social classes. It is most frequently reported by class 1, followed by class 3. Differences in reported use between the three groups are small, however, and not statistically significant at the 95% level, since the observed chi-squared value (0.199) does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Bathroom is not reported by speakers from class 1, and accounts for 6.67% of reports from class 2 speakers, and 5.00% of reports from informants in class 3. These differences in reported use are not statistically significant, however, as the chi-squared value is only 0.603. *Bog* is most frequently reported by speakers from class 1. It accounts for 29.03% of responses from this group, compared with 25.00% of responses from class 3 speakers, and 13.33% of responses from those in class 2. Again, this difference in reported use between social classes is not statistically significant (the chi-squared value is 1.391). Reported use of the variant *lavatory* is almost identical amongst classes 1 and 2, increasing slightly among those in class 3. Differences between groups are small, however, and with a chi-squared value for

lavatory and social class of 0.057, they are not statistically significant. This is noteworthy in its own right, however, since one might have expected use of this variant, described by Ross (2007: 139) as an upper class term, to have been reported more frequently by those in class 1 than in the other two social classes. With a chi-squared value of 0.036, the small differences in reported use of *toilet* between social classes are not statistically significant either. Again, this is noteworthy, since one might have expected reported use of this variant, a term which Ross (2007: 139) describes as non-upper class, to have been higher among class 3 informants than those in classes 1 and 2.

Figure 4.5 Most frequently reported variants for the notion word *toilet* according to the speaker variable of social class



Raw data

variant	class 1	class 2	class 3	not ranked
<i>bathroom</i>	0.00% [n = 0/31]	6.67% [n = 2/30]	5.00% [n = 1/20]	0.00% [n = 0/6]
<i>bog</i>	29.03% [n = 9/31]	13.33% [n = 4/30]	25.00% [n = 5/20]	0.00% [n = 0/6]
<i>lavatory</i>	6.45% [n = 2/31]	6.67% [n = 2/30]	10.00% [n = 2/20]	0.00% [n = 0/6]
<i>loo</i>	38.71% [n = 12/31]	33.33% [n = 10/30]	35.00% [n = 7/20]	66.67% [n = 4/6]
<i>toilet</i>	19.35% [n = 6/31]	23.33% [n = 7/30]	20.00% [n = 4/20]	33.33% [n = 2/6]

Eight variants were reported only once in response to the notion word *toilet*. Class 1 informants reported using *gents* and *shitter*. Class 2 informants reported *khazi*, *little room*, *shit house*, *small room*, and *W.C.*. A class 3 informant reported *trap*. In total, informants from classes 1 and 3 reported using 6 different variants, while class 2 speakers reported a total of 10.

The breadth of variants reported by speakers from all three social classes shows that, regardless of class, Southampton area speakers have in their lexicon a number of terms for the notion word *toilet*, ranging from those that are arguably marked for social class (*lavatory* and *toilet*, according to Ross), to those which speakers themselves acknowledge to be potentially offensive and/or which are suitable for use with only specific audiences (see section 4.2.1).

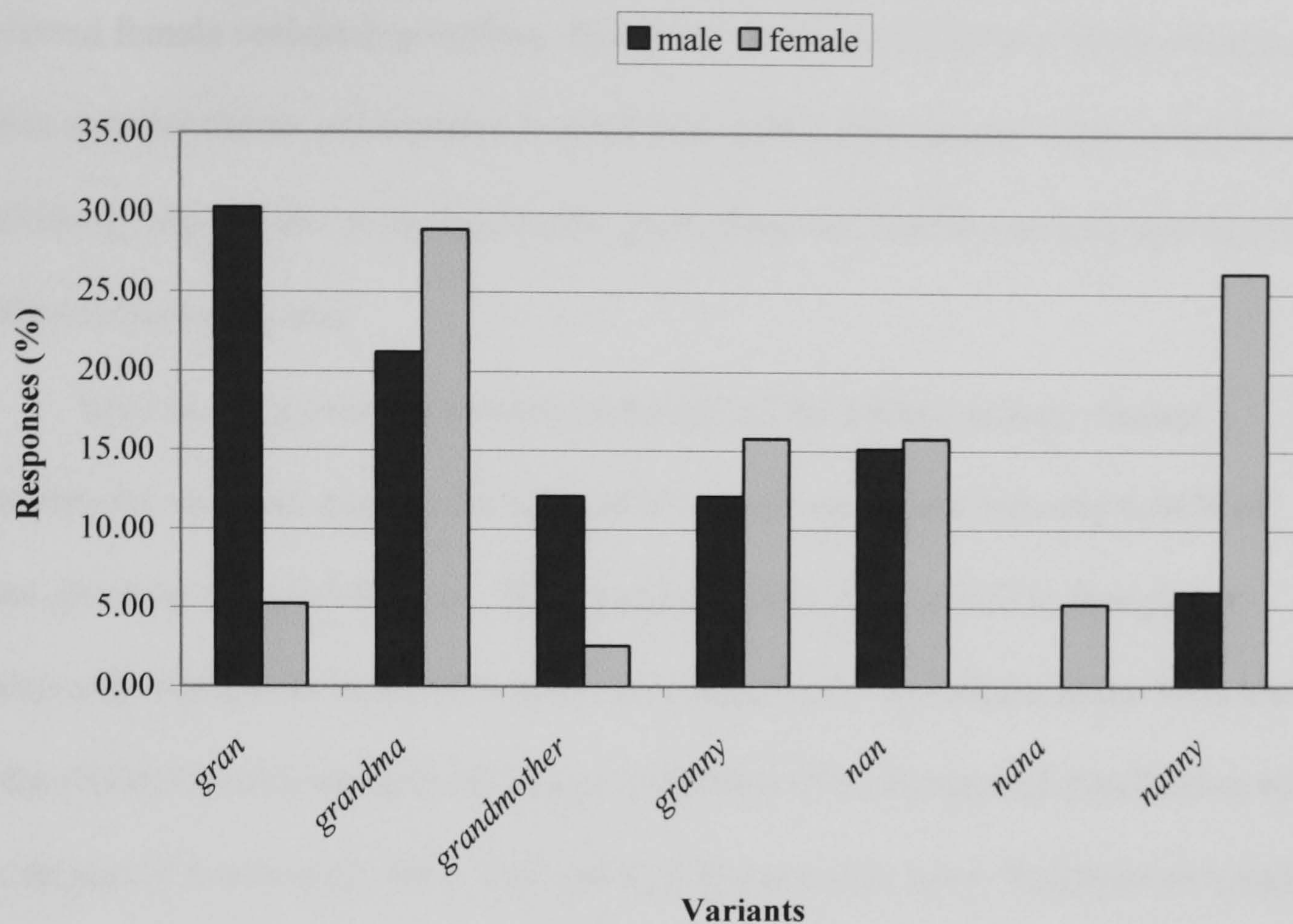
4.3 Grandmother and Grandfather

The notion words *grandmother* and *grandfather* elicited seventeen different variants. Three of these (*aged people*, *grandparents*, *grandparent*) are not specific to either sex and so are discounted from analysis. Of the remaining fourteen variants, eight pertain to the notion word *grandmother* and six to the word *grandfather*. In the following sections, *grandmother* and *grandfather* are analysed according to the extralinguistic variables of speaker sex and social class.

4.3.1 Grandmother and Speaker Sex

In the case of *grandmother*, seven of the eight variants elicited were reported by more than one informant, and it is these seven which are presented in Figure 4.6, according to the speaker variable of sex.

Figure 4.6 Most frequently reported variants for the notion word *grandmother* according to the speaker variable of sex



Raw data

variant	male	female
<i>gran</i>	30.30% [n = 10/33]	5.26% [n = 2/38]
<i>grandma</i>	21.21% [n = 7/33]	28.95% [n = 11/38]
<i>grandmother</i>	12.12% [n = 4/33]	2.63% [n = 1/38]
<i>granny</i>	12.12% [n = 4/33]	15.79% [n = 6/38]
<i>nan</i>	15.15% [n = 5/33]	15.79% [n = 6/38]
<i>nana</i>	0.00% [n = 0/33]	5.26% [n = 2/38]
<i>nanny</i>	6.06% [n = 2/33]	26.32% [n = 10/38]

Sex-based variation is seen clearly in the use of *gran*, the variant for *grandmother* most frequently reported by male informants. This term accounts for 30.30% of all male responses but only 5.26% of female responses. With a chi-squared value of 6.203, this difference in use between male and female informants is

statistically significant at the 95% level, as the observed value exceeds the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). The preferred female variant is *grandma*. However, since the difference in reported use of this term by males and females is small and, with a chi-squared value of 0.224, not statistically significant, *grandma*, unlike *gran*, does not function as a marker of sex in the Southampton area.

Sex-based variation is evident in the use of the variant *nanny*. *Nanny* accounts for over one quarter, 26.32%, of all female responses but only 6.06% of those given by male informants. With a chi-squared value of 3.818, though not statistically significant at the 95% level, it is statistically significant at the 90% level, as the observed value exceeds the upper 10% point of a chi-squared distribution with one degree of freedom (2.706). Like *gran*, it appears that *nanny* functions as a marker of sex-based variation in the Southampton area study, being used predominantly among women.

The variant *nana* is used exclusively by female informants. However, it accounts for only 5.26% of the total responses given by women, and the difference in male and female reported use of this variant is not statistically significant (the chi-squared value is 0.382). Similarly, the differences in reported use between male and female informants of *grandmother* (chi-squared value 1.196), *granny* (chi-squared value of 0.010) and *nan* (chi-squared value of 0.065) are small and not statistically significant. None of these variants function as markers of sex differentiation in the Southampton area.

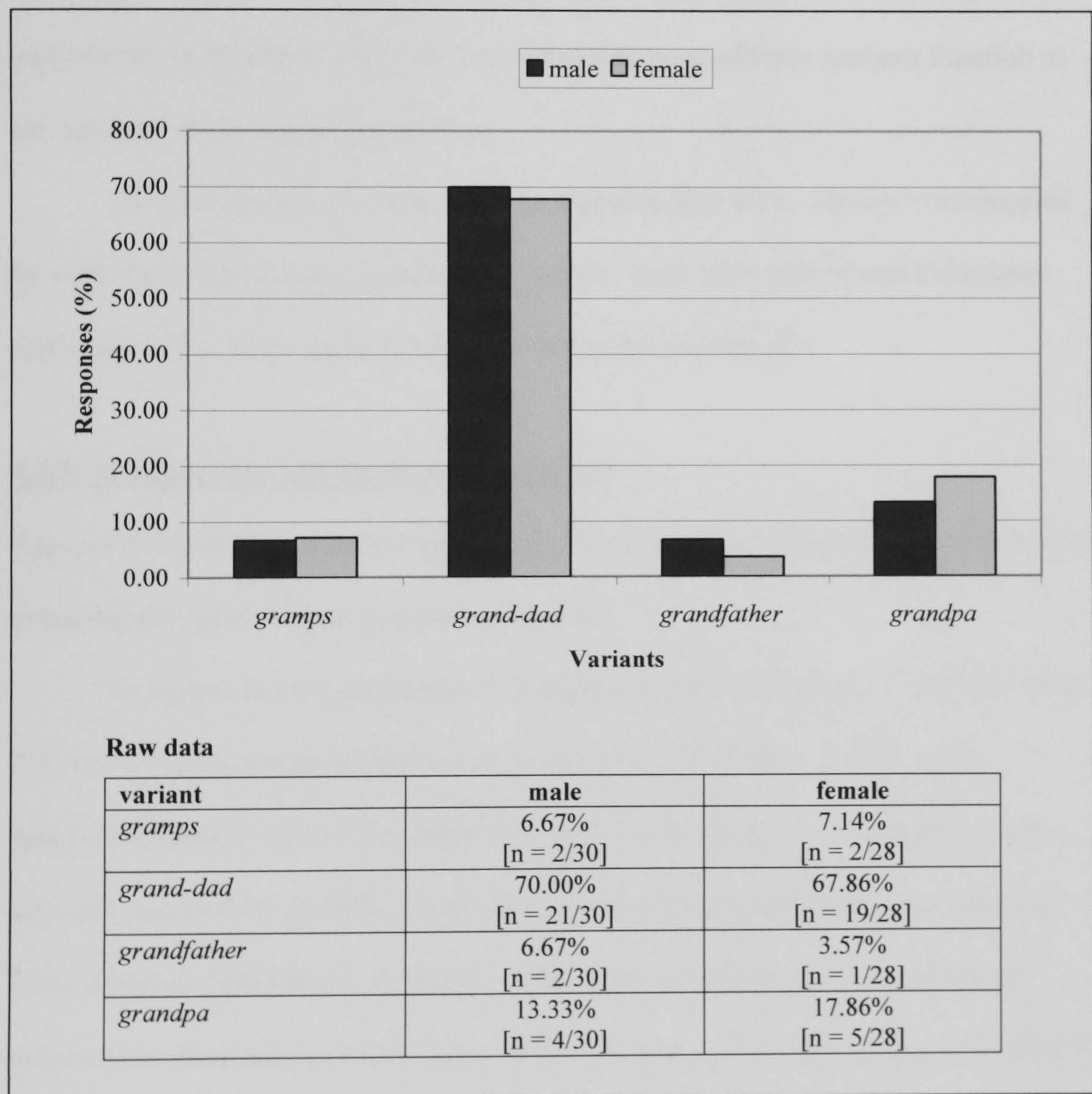
Grandmum is the only variant for *grandmother* reported only once, and is used by a male informant. Male and female informants both reported seven different variants in total for the notion word *grandmother*. Regardless of speaker sex, the

terms for *grandmother* beginning with <g> are more frequently reported overall than those beginning with <n>.

4.3.2 *Grandfather and Speaker Sex*

Four of the six variants for *grandfather* were reported more than once, and these are presented in Figure 4.7, according to the speaker variable of sex.

Figure 4.7 Most frequently reported variants for the notion word *grandfather* according to the speaker variable of sex



In contrast to the notion word *grandmother*, reported variants for the notion word *grandfather* exhibit no sex-based variation. Both male and female respondents

most frequently reported the variant *grand-dad*. The small difference in reported use of this term between males and females is not statistically significant at the 95% level, since the chi-squared value of 0.012 does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). Similarly, the small differences in reported use between males and females of *gramps* (chi-squared value of 0.200), *grandfather* (chi-squared value of 0.004), and *grandpa* (chi-squared value of 0.013) are not statistically significant either. Given that all of the most frequently occurring variants for *grandfather* are reported in roughly equal measures by male and female informants, it could be concluded that none of these variants function as sex markers in the Southampton area.

Two variants for *grandfather* were reported only once. *Gramp* was reported by a female informant and *grandpop* by a male. Both male and female informants reported a total of five variants for the notion word *grandfather*.

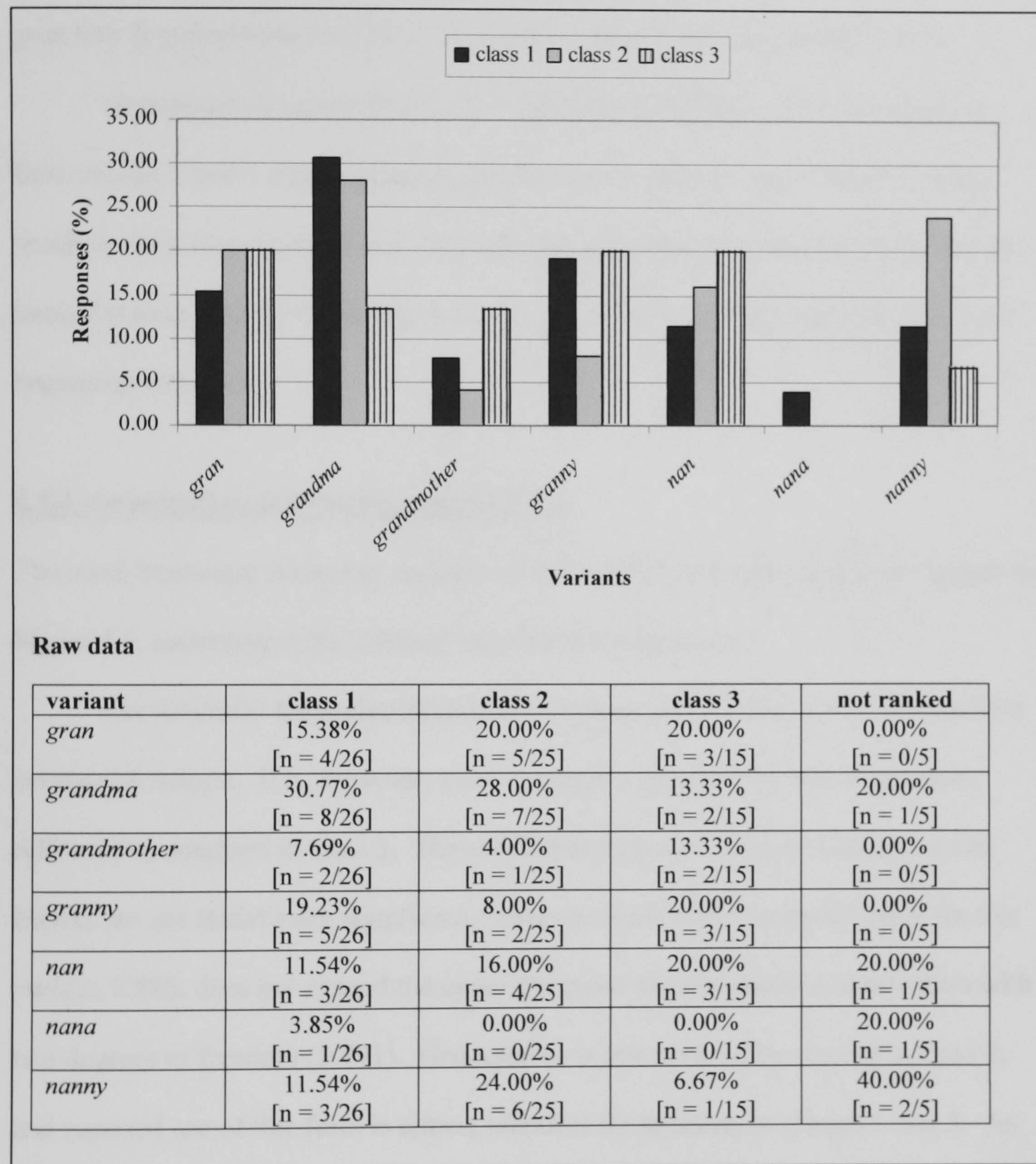
4.3.3 Grandmother and Speaker Social Class

Figure 4.8 shows the variants reported more than once in response to the notion word *grandmother*, according to speaker social class.

Grandma is the most frequently reported variant for speakers in classes 1 and 2, but by a larger margin in the former group than in the latter. In the case of speakers in class 3, there is no single term which is favoured, as *gran*, *granny* and *nan* each account for 20.00% of responses. Reported use of *grandma* decreases as level of social class lowers, though the differences in reported use of this variant between the three classes is not statistically significant, since the chi-squared value of 0.779 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). *Nan* displays the opposite tendency to *grandma*, with reported use of this variant increasing as level of social class falls. Again, however,

differences in reported use between the three classes are not statistically significant, the chi-squared value for *nan* being 0.110.

Figure 4.8 Most frequently reported variants for the notion word *grandmother* according to the speaker variable of social class



Reported use of *gran* is lowest among class 1 speakers, and is equal among those in classes 2 and 3. Reported use of *grandmother* is highest amongst those in class 3, and lowest amongst those in class 2. Reported use of *granny* is lower for class 2 informants than it is for their counterparts in classes 1 and 3. The opposite is

true of *nanny*, where reported use is higher in class 2 than it is in classes 1 and 3. *Nana* is only reported by a class 1 speaker. Differences in reported use of these variants between social classes are small, however, and are not statistically significant at the 95% level. The chi-squared values for these variants are as follows: *gran* 0.037; *grandmother* 0.336; *granny* 0.637; *nana* 0.401; and *nanny* 1.279.

Grandmum is reported once, by a speaker from class 3. In total, speakers from classes 1 and 3 report using seven different variants for *grandmother*, while speakers from class 2 report six. Overall, and regardless of speaker social class, the terms for *grandmother* beginning with <g> are more frequently reported than those beginning with <n>.

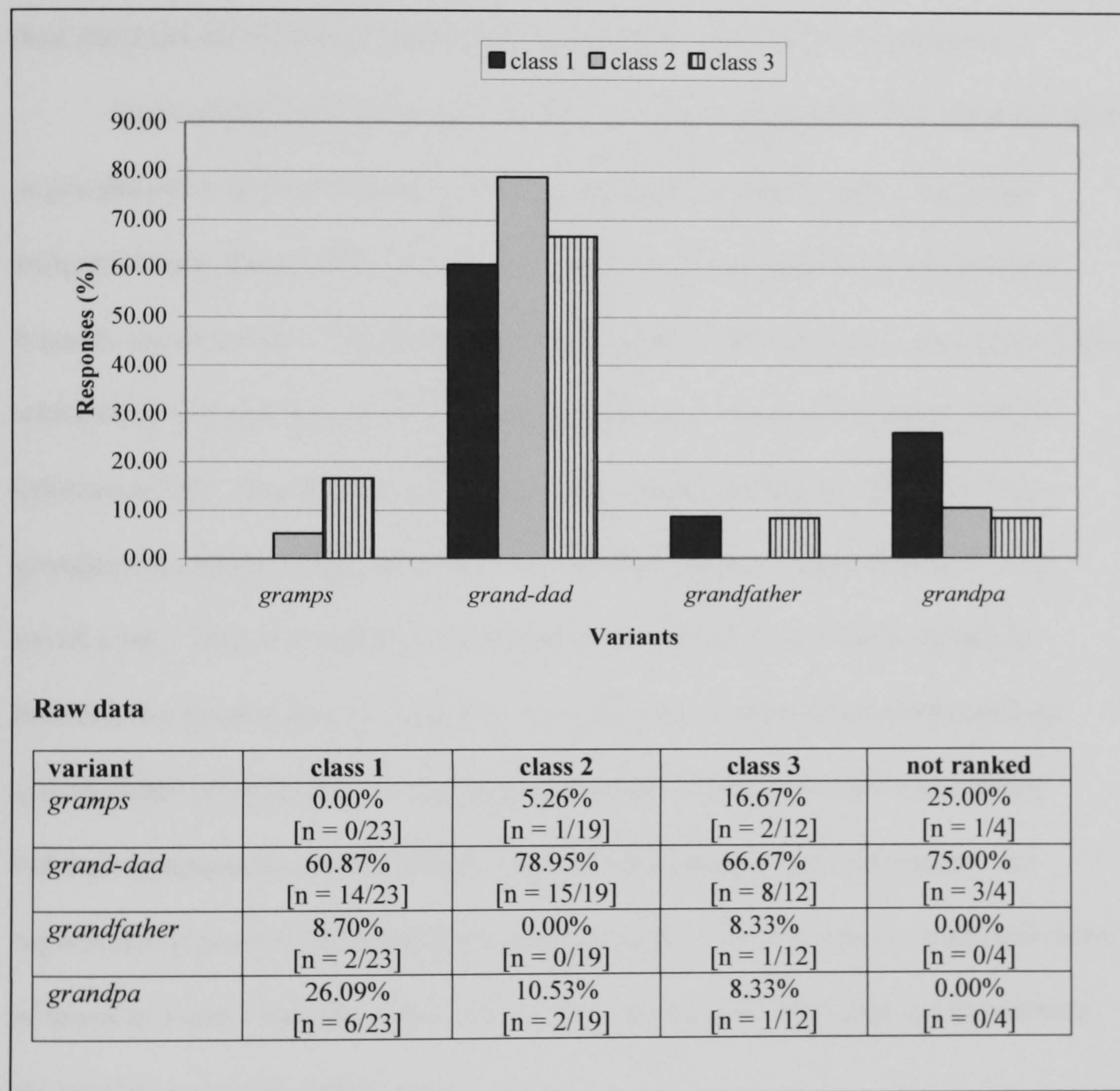
4.3.4 Grandfather and Speaker Social Class

The most frequently occurring variants for the notion word *grandfather* are shown in Figure 4.9, according to the speaker variable of social class.

Grand-dad is the preferred term for all three social classes, in each case by a substantial margin. It is, however, most frequently reported by class 2 speakers, followed by speakers in class 3. These differences in reported use between social classes are not statistically significant, however, since the chi-squared value for this variant, 0.885, does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). *Grandfather* is not reported by anyone in class 2, and reported use of this form is almost identical for speakers in classes 1 and 3. As in the case of *grand-dad*, however, differences in reported use of *grandfather* between social classes are not statistically significant (the chi-squared value is 0.395). Use of *gramps* increases as level of social class lowers. Overall, however, differences in reported use of this variant between the three social classes is small, and is not statistically significant (the observed chi-squared value is 1.802).

Reported use of *grandpa* decreases as level of social class falls, but, like all the most frequently occurring variants for the notion word *grandfather*, differences in reported use between social classes are not statistically significant (the chi-squared value is 1.188).

Figure 4.9 Most frequently reported variants for the notion word *grandfather* according to the speaker variable of social class



Two variants for *grandfather* are reported only once. *Gramp* is reported by a speaker from class 2, while *grandpop* is reported by a class 1 speaker. Each of the three social classes reports using a total of four variants.

4.3.5 Grandmother and Grandfather – A Summary

A comparison of the notion words *grandmother* and *grandfather* reveals that sex-based variation in reported lexical use is more evident in the case of *grandmother* than in that of *grandfather*. None of the variants elicited for *grandfather* appear to exhibit any significant signs of sex-based variation, while two of the most frequently reported variants for *grandmother*, *gran* and *nanny*, do exhibit sex-based variation, their reported use differing statistically significantly according to speaker sex.

None of the differences in reported use of the variants for either *grandmother* or *grandmother* between social classes are statistically significant. That these variants do not appear to be socially marked in the Southampton area is in some respects unsurprising. People generally refer to their grandparents using those terms which their own parents used for their grandparents. In cases, therefore, where informants have moved between social classes since childhood, their terms for grandparents might reflect the social class of their parents rather than their own social class. Terms of address might also reflect the nature of the relationship between the speaker and the referent. As such, use of very formal terms such as *grandmother* and *grandfather* might be the result of a more formal relationship between grandparent and grandchild than the relationship of an informant who reports using a more colloquial form, such as *nana*. They are not so much indicative of speaker social class, therefore, than of the closeness of the relationship between the speaker and their grandparent.

Though not directly linked to the possible influence of the speaker variables of sex or social class on lexical choice but to lexical variation in general, it is of interest to note that fewer variants were elicited for the notion word *grandfather* than were elicited for *grandmother*, the former producing six different terms and the latter

eight. Some speakers comment that they were unable to offer a response for one or the other notion word, their grandmother or grandfather not having been alive during the respondent's lifetime. The greater number of different variants for *grandmother* than for *grandfather* and the higher total number of tokens for the former than the latter, seventy-one overall for *grandmother* compared to fifty-eight for *grandfather*, can perhaps be explained by the longer life expectancy enjoyed by women than by men. It is possible that more variation exists in the terms for *grandmother* than for *grandfather* because, women living longer than men on average, more informants knew or know their grandmothers than their grandfathers, thus increasing the potential for lexical variation in names for them.

4.4 Long seat in main room of house

Four variants were elicited for the notion phrase *long seat in main room of the house*, of which three occurred more than once. In the following sections, *long seat in main room of house* is analysed according to the speaker variables of age and social class.

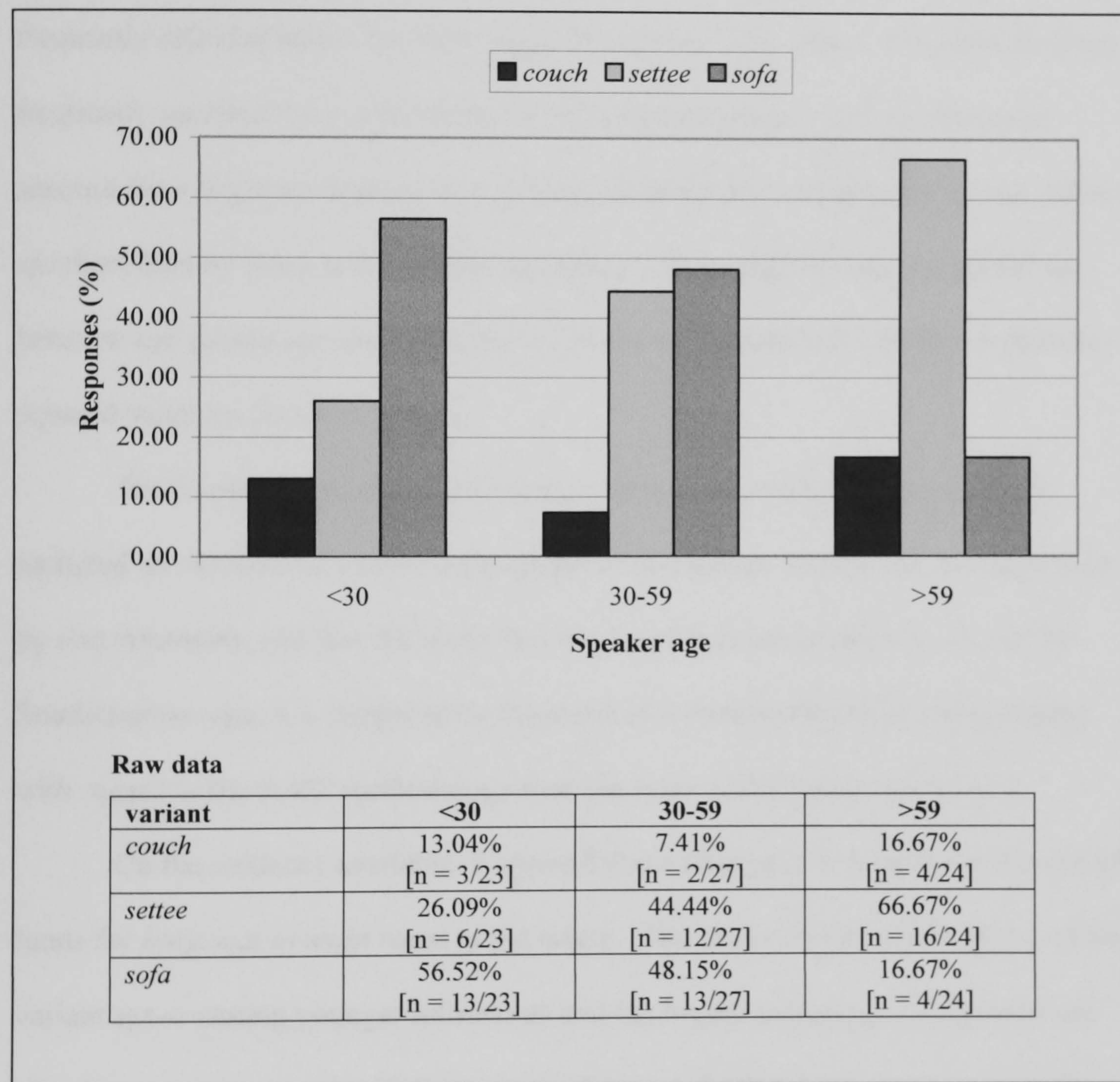
4.4.1 Long seat in main room of house and Speaker Age

In Figure 4.10, the most frequently occurring variants for the notion word *long seat in main room of house* are shown, according to the speaker variable of age.

It appears that the use of particular terms for *long seat in main room of house* correlates strongly with speaker age. Reported use of the word *settee* increases with increasing speaker age, and differences in use between the three age groups are significant at the 95% level, since the chi-squared value for *settee*, 7.826, exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). In fact this difference in use is also significant at the 97.5% level as the chi-squared

value also exceeds the upper 2.5% of a chi-squared distribution with two degrees of freedom (7.378). There is clearly a strong relationship between use of the variant *settee* and the age of the informant, with older informants more likely to report using this variant than their younger counterparts.

Figure 4.10 Most frequently reported variants for the notion phrase *long seat in main room of house* according to the speaker variable of age



Reported use of the word *sofa* also varies with speaker age, but in this case, reported use of the term decreases with increasing age. As in the case of *settee*, differences in use of *sofa* between the three age groups are significant at the 95% level, since the chi-squared value for *sofa*, 8.760, exceeds the upper 5% point of a

chi-squared distribution with two degrees of freedom (5.991). This difference in use is also significant at the 97.5% level, as the chi-squared value also exceeds the upper 2.5% of a chi-squared distribution with two degrees of freedom (7.378). There is marked age differentiation in the use of *sofa*, with younger speakers being more likely to report using this form than their older counterparts.

Among speakers under 30 and those aged 30 to 59, *sofa* is the most frequently reported term. For those aged 60 and over, it is *settee*. *Couch* is the least frequently reported term of the three for all three age groups. It does, however, account for a larger percentage of responses given by the youngest and by the oldest speakers than by those in the middle age group. These differences in reported use between age groups are small, however, and are not statistically significant (the chi-squared value for *couch* is 0.381).

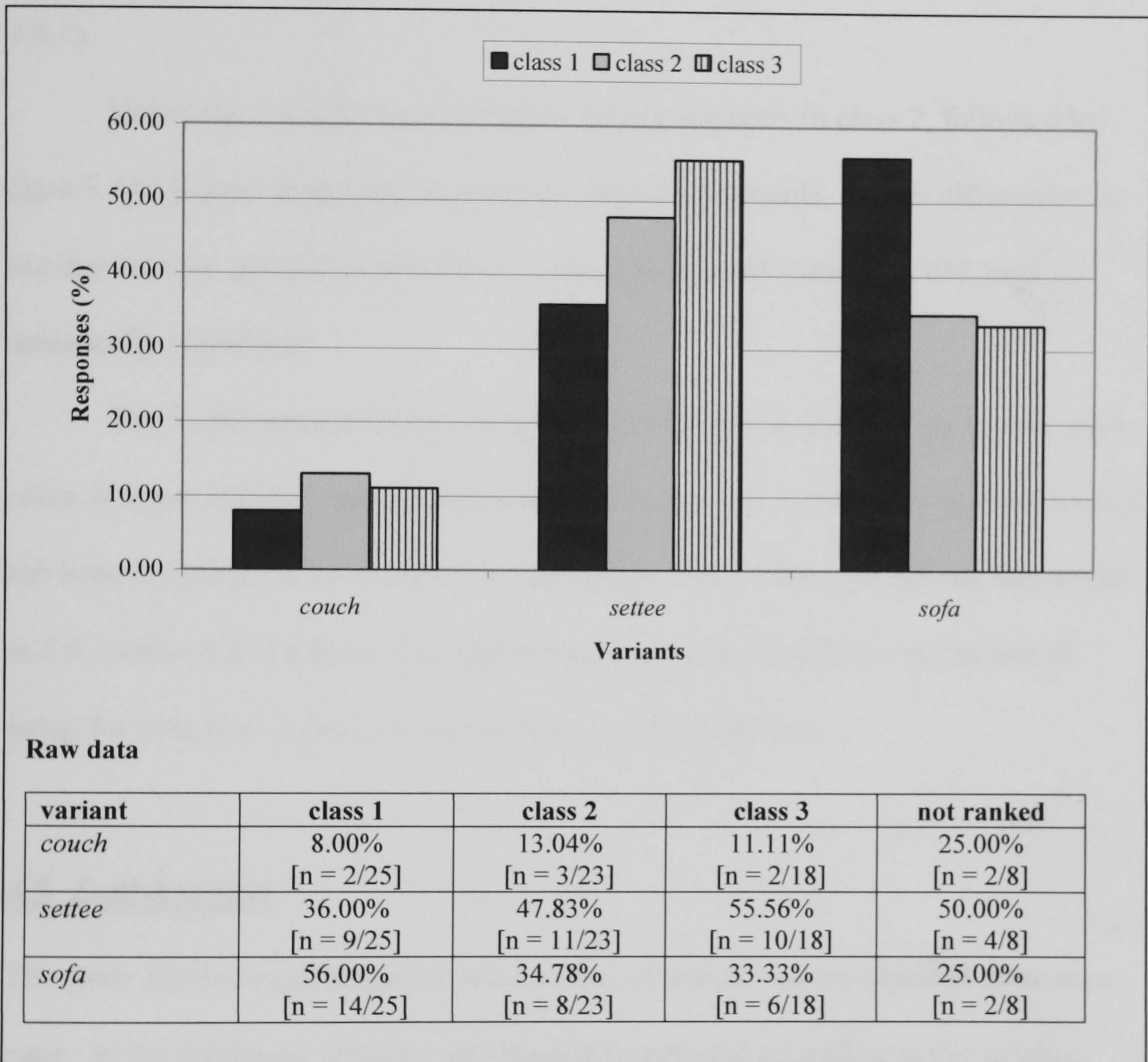
The fourth variant elicited for *long seat in main room of house* is *bench*, included on the SRN of a male in the under 30 age group. Given that it is only used by one informant, and that the writer has not heard it in use by anyone else in the Southampton area, it is judged to be the result of a misunderstanding, more telling with regard to the SuRE methodology than the lexis of the Southampton area.

On the evidence available, it appears that a change is in progress in the use of terms for *long seat in main room of the house*. The relatively low reported use of the variant *settee* among younger informants and the higher percentage of reported use of *sofa* among those under 60 than among those aged 60 and over suggests that the use of *settee* is on the decline and might ultimately be replaced by *sofa*. Though it is impossible to say this for certain given that the present study takes place in apparent time, the data suggest that *sofa* is supplanting *settee*.

4.4.2 Long seat in main room of house and Speaker Social Class

Figure 4.11 shows the three variants reported more than once in response to the notion phrase *long seat in main room of house*, according to speaker social class.

Figure 4.11 Most frequently reported variants for the notion phrase *long seat in main room of house* according to the speaker variable of social class



It is apparent from Figure 4.11 that reported use of the variants *settee* and *sofa* differs according to speaker social class. While *settee* is the most frequently reported variant for classes 2 and 3, *sofa* is the term preferred by informants in class 1. Reported use of *settee* increases the lower the speaker social class is, while the opposite is true of *sofa*. Differences in use between social classes of these two variants are not statistically significant, however, since the chi-squared values for

settee (1.694) and *sofa* (3.045) do not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). It might be the case that it is the age of the informants in class 1 which is responsible for the high reported use of *sofa* in the group, since 11 out of the 14 tokens of reported use of this variant in class 1 are reported by speakers under 60, for whom this is the preferred term (see section 4.4.1).

Use of the variant *couch* is highest among speakers in class 2, followed by class 3, and is least frequently reported by class 1 informants, though differences in use between the groups are small and, with a chi-squared value of 0.109, not statistically significant.

The fourth variant elicited in response to the notion phrase *long seat in main room of house* is *bench*, reported by a speaker in class 2. As stated in section 4.4.1, it has been judged to be the result of a misunderstanding. When comparing the results in 4.4.1 and 4.4.2, it appears that speaker age is a greater influence on the use of terms for *long seat in main room of house* than speaker class.

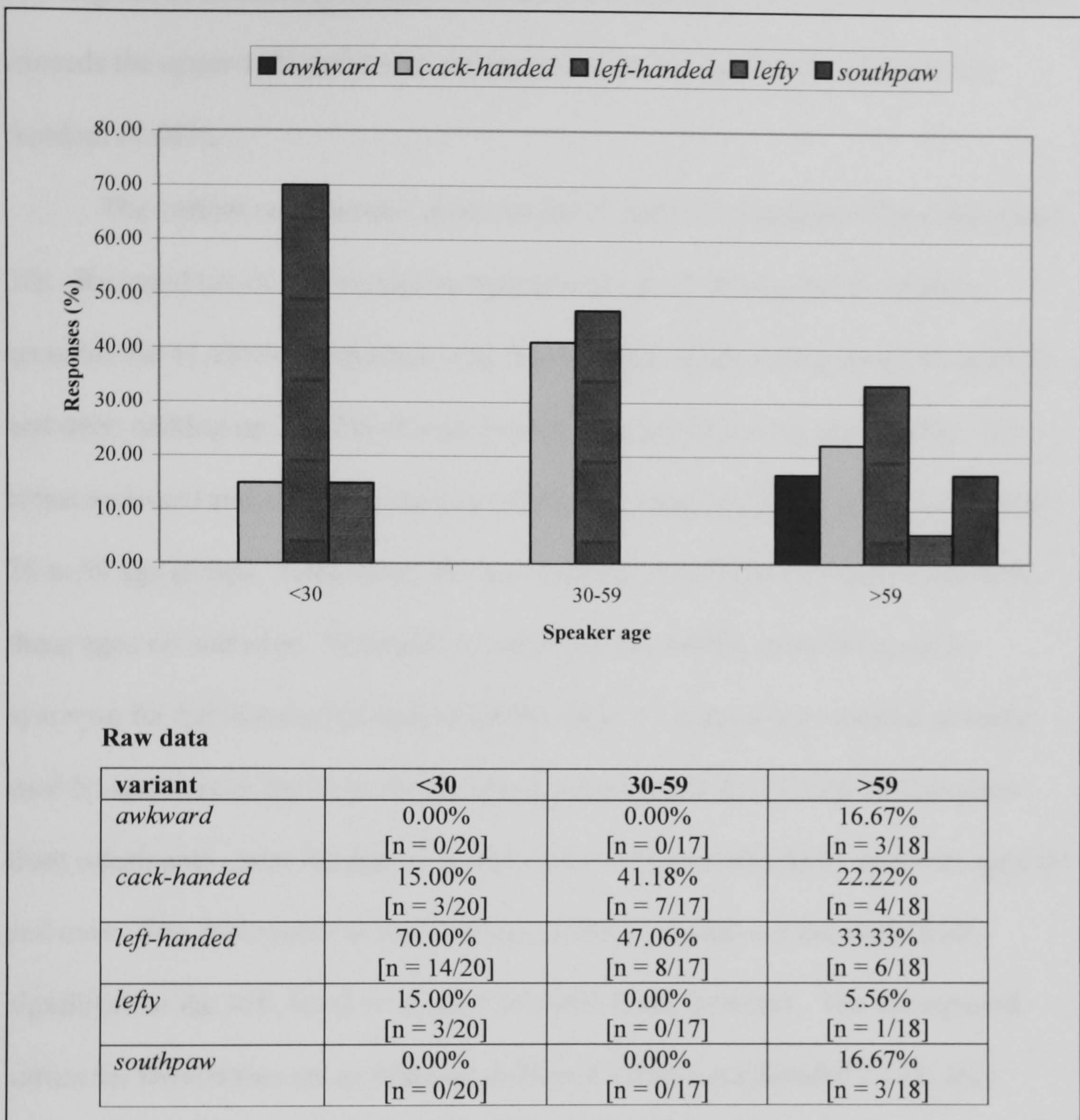
4.5 Left-handed

The study elicited eight terms for *left-handed*, of which five are reported more than once. In the following sections, *left-handed* is analysed according to the speaker variables of age and social class.

4.5.1 Left-handed and Speaker Age

Figure 4.12 shows the most frequently occurring variants for the notion word *left-handed* according to the speaker variable of age.

Figure 4.12 Most frequently reported variants for the notion word *left-handed* according to the speaker variable of age



Most frequently reported of the five variants for each of the three age groups is *left-handed*. Use of this term, however, decreases with increasing age. The greater reported use of this variant among younger informants than among their older counterparts perhaps indicates a change in progress. It is apparent that the wider range of variants reported by older speakers is perhaps being replaced by the standard term *left-handed* among their younger counterparts. Though difference in reported use of *left-handed* is not significant at the 95% level, since the chi-squared

value of 5.242 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991), it is nevertheless significant at the 90% level, as it exceeds the upper 10% point of a chi-squared distribution with two degrees of freedom (4.605).

The variant *cack-handed* accounts for 15.00% of responses offered by under 30s. Reported use of this variant increases in the 30 to 59 age group, where it accounts for 41.18% of responses, and then declines again among speakers aged 60 and over, making up 22.22% of responses from speakers in this age bracket. The terms *awkward* and *southpaw* are not reported by speakers in either the under 30 or 30 to 59 age groups. Each, however, accounts for 16.67% of responses offered by those aged 60 and over. It should be noted that the variant *awkward* is not a synonym for *left-handed* but instead means clumsy. *Lefty* is not reported as being used by speakers in the 30 to 59 age group but accounts for 15.00% of responses from informants under the age of 30 and 5.56% of those offered by speakers aged 60 and over. The differences in reported use of these variants are not statistically significant at the 95% level or, indeed, the 90% level, however. The chi-squared values for these terms are as follows: *awkward* 3.030; *cack-handed* 2.132; *lefty* 1.313; and *southpaw* 3.030.

The claim for possible lexical erosion here is further supported when the tail-end of the data are examined. Three other variants were elicited for *left-handed*. Of these, two (*cack* and *left hooker*) are reported by speakers in the 30 to 59 age group and one (*back-handed*) by a speaker in the 60 and over group. When analysed alongside the more frequently occurring variants, it is apparent that it is the oldest age group which employs the largest number of variants for *left-handed*. This group uses six different terms in total. Speakers aged between 30 and 59 use four variants

and the youngest respondents only three. On the evidence available, it would appear that lexical erosion is taking place in the Southampton area, particularly given that Upton et al. (1994: 241-42) record that the Survey of English Dialects collected eight-four different terms for *left-handed*. In light of the fact that the kind of stigma which was formerly attached to left-handedness is now lessening in the United Kingdom, it appears likely that terms which imply a difference or deviance are in decline. This theory is supported by the puzzlement expressed by many younger informants when asked for synonyms for *left-handed*. Many seemed unaware that there might be other terms in existence for this notion word. Being left-handed is no longer perceived as unusual, and the lexis used by younger people reflects this.

4.5.2 Left-handed and Speaker Social Class

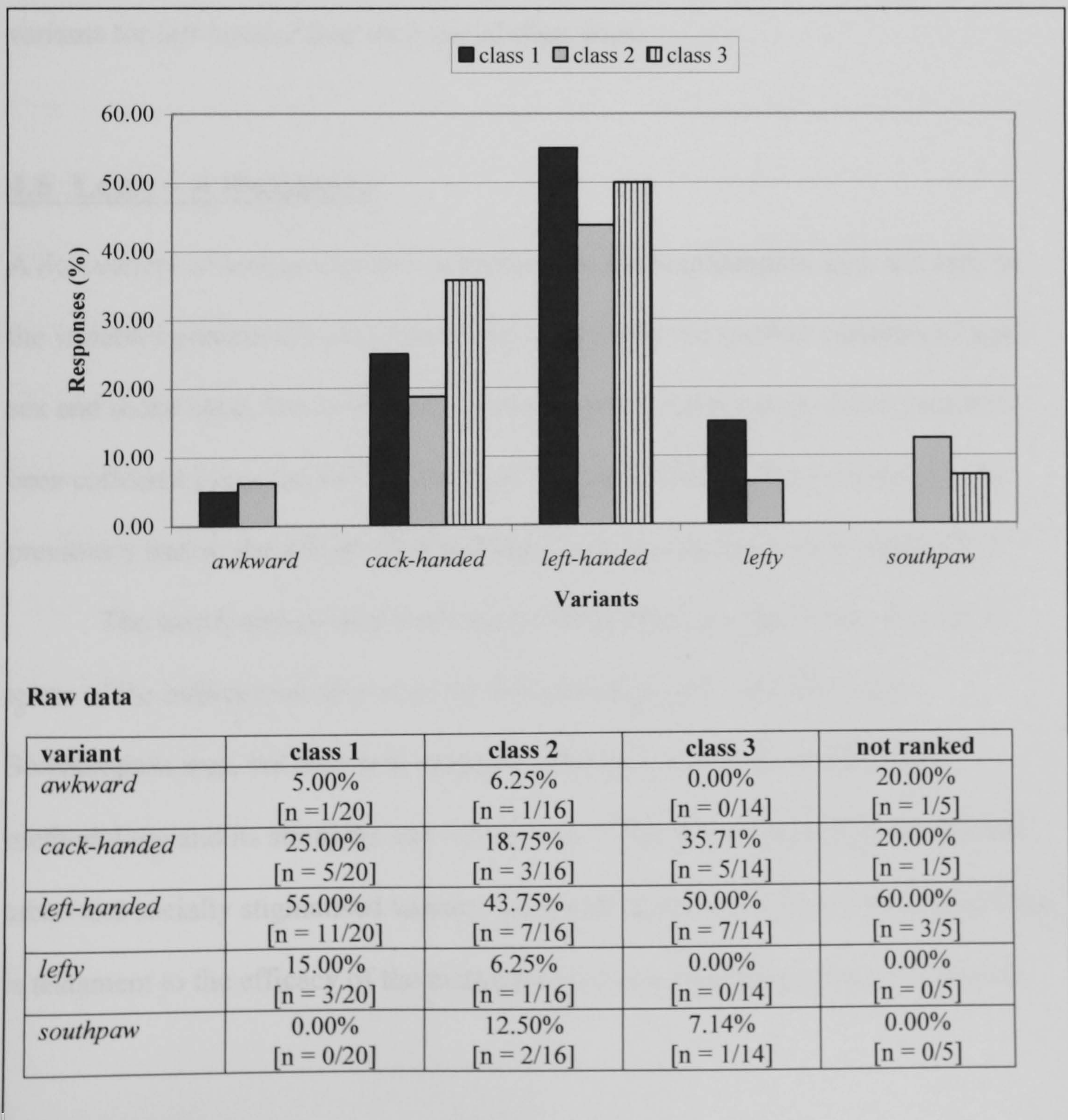
Figure 4.13 shows the variants reported more than once in response to the notion word *left-handed*, according to the extralinguistic variable of social class.

The variant *left-handed* is the preferred term for all three social classes. It is most frequently reported by speakers in class 1, followed by those in class 3, and least frequently reported by informants in class 2. These differences in reported use between classes small, however, and are not statistically significant at the 95% level since the chi-squared value for *left-handed*, 0.450, does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). *Cack-handed* is also reported by each of the three groups, but in the case of this variant it is those in class 3 who report it most frequently, followed by speakers in class 1 and, lastly, those in class 2. Again, these differences in use between classes are not statistically significant, as the chi-squared value for *cack-handed* is 0.439.

Awkward, *lefty* and *southpaw* are reported by only two out of the three classes, *awkward* and *lefty* by classes 1 and 2, and *southpaw* by classes 2 and 3.

Reported use of *awkward* is higher among class 2 speakers than among class 1 informants. This variant is not reported by anyone in class 3. With a chi-squared value of 0.156, differences in reported use of *awkward* between the three groups is not statistically significant. Reported use of *lefty* declines the lower the social class. Again, however, differences in reported use between the classes are not statistically significant, the chi-squared value for *lefty* being 0.964. *Southpaw* is not reported by any speakers in class 1, and is reported more frequently by those in class 2 than in class 3. It has a chi-squared value of 0.904, meaning that differences in reported use between classes are not statistically significant.

Figure 4.13 Most frequently reported variants for the notion word *left-handed* according to the speaker variable of social class



The three other variants elicited for the notion word *left-handed* are *back-handed* and *left-hooker*, both reported by speakers from class 2, and *cack*, presumably from *cack-handed*, reported by a speaker from class 3. Overall, it is speakers from class 2 who report using the widest range of variants, 7 in total, while speakers from classes 1 and 3 report 4 each.

Despite a range of variants for *left-handed* being elicited, all social class groups most frequently report using *left-handed* itself, by a large margin. This, coupled with the lack of statistically significant social class differentiation in reported use of any of the variants, suggests that terms for *left-handed* do not function as class markers. A comparison of section 4.5.1 with the present section suggests that the age of the speaker has a greater influence on reported use of variants for *left-handed* than their social class does.

4.6 Lexis – A Summary

A rich variety of lexical variation is apparent in the Southampton area, not only in the variables presented in this chapter for analysis by the speaker variables of age, sex and social class, but in all of the notion words and phrases on which data have been collected using the SuRE methodology Sense Relation Network sheets. As previously stated, the full set of lexical data collected can be seen in Appendix 2.

The words and phrases analysed in this chapter are significant not only in terms of the information they provide about the linguistic situation in the Southampton area, but also with regard to what they tell us about the SuRE methodology and its strengths and weaknesses. That the methodology has elicited taboo and socially stigmatised variants from informants, as well as very formal lexis is testament to the efficacy of the methodology in encouraging speakers to record

lexis of different styles, suitable for use with a range of audiences. Variants such as *bench* for *long seat in main room of house* and *bundle of joy* for *pregnant*, do indicate, however, that there is potential for misunderstanding in the methodology, as would be the case with any questionnaire which informants are asked to complete without a fieldworker present.

As a young, middle class female fieldworker, it might have been the case that I influenced some of the responses given. However, as both highly formal and vernacular, taboo variants have been collected (see section 4.2.1. for example), this has not proved problematic, and is an issue that would be faced any fieldworker.

In some instances, lexical variants function as markers of particular speaker groups. Sex-based variation is apparent in terms used by males and females for *toilet* and *grandmother*. Age-based variation can also be seen in the notion words/phrases *long seat in main room of house* and *left-handed*. *Sofa* and *left-handed* are the terms preferred by speakers under 30, and it seems that these variants are supplanting *settee* and the wide range of variants for *left-handed* used by older informants. It would appear that changes are in progress in the most frequently reported variants for these notion words/phrases. The present study is an apparent time one, however, and a real-time study would be needed to confirm that these changes are taking place. Reported use of lexis in the current study appears to be less influenced by the social class of the speakers than it is by their age and sex. Johnson (1996: 81) states that 'words rarely are subject to stigmatization based upon a stereotypical association with the speech of lower status groups'. This might explain the relative lack of social class differentiation in the reported use of lexical data in the present study.

Chapter 5

5.0 Grammar

In this chapter, a selection of grammatical features is analysed according to the speaker variables of sex, age and social class. In sections 5.2 to 5.5, reported speaker use of four nonstandard grammatical features is examined, and in section 5.6, reported use of a further two grammatical features is compared with actual speaker use of these features. Reported use refers to speaker claims to use a variable when talking to a friend (see section 3.10.2).

As discussed in section 3.8, the Southampton area study has 30 male and 30 female informants, and 20 speakers in each of the three age groups (under 30, 30 to 59, and 60 and over). The speaker sample is comprised of 3 social class groups, class 1 being the highest, and class 3 the lowest. Class 1 has 20 informants, class 2 has 18 informants, and class 3 has 16 informants (see section 3.9 for a discussion of the social stratification of Southampton area speakers).

Chi-squared testing (see section 3.12) has been undertaken when differences in reported use of a variant between speaker groups is small.

5.1 Grammar Scores

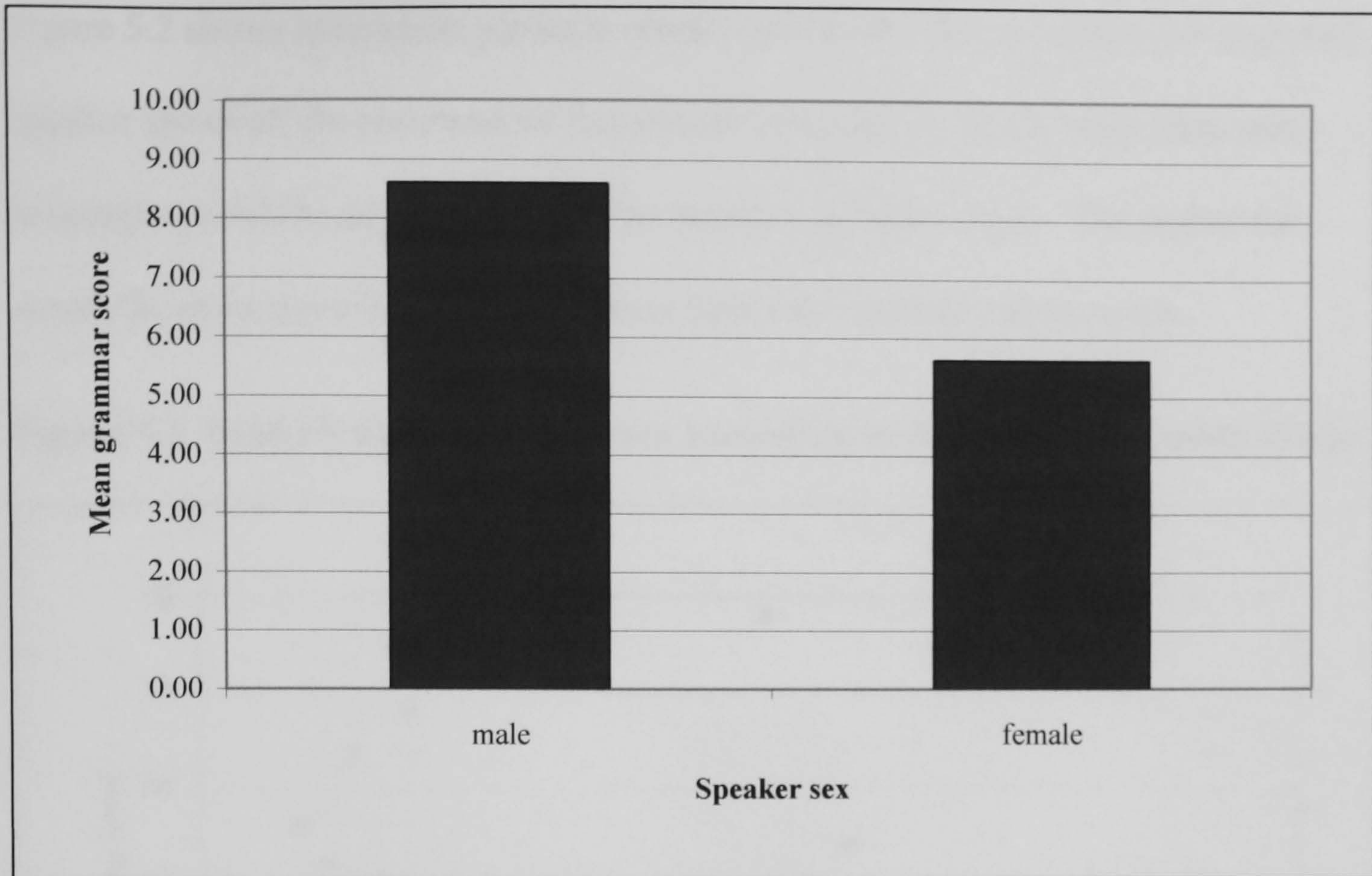
It is impracticable to analyse individually by the speaker variables of sex, age and social class each of the grammatical variables investigated in the Southampton area language questionnaire. An examination of the grammatical data as a whole, however, provides a useful preliminary insight into the possible relationship between these extralinguistic variables and the reported use of nonstandard grammar prior to the analysis of individual grammatical variables. A grammar score, which takes into consideration reported speaker use of all the nonstandard features investigated by the

Southampton area language questionnaire. has been calculated for each informant, as discussed in section 3.10.2. Informants are given a score of 1 if they claim to employ a particular nonstandard feature when talking to a friend and a score of 0 if they do not. They are then given an overall score out of 45 (the total number of grammatical variables under examination, excluding control sentences). The higher the score, the more nonstandard grammatical forms they claim to use.

5.1.1 Grammar Scores and Speaker Sex

For male informants, individual grammar scores range between 0 and 29. For female informants, the range is smaller, between 0 and 18. The ranges are, however, somewhat misleading, since they are affected by outlying scores. This is evident particularly in the case of the male informants, for whom the range is greatly increased by an outlying grammar score of 29 when the majority of scores in this speaker group are under 14.

A mean grammar score is less affected than a range by such outliers, and has been calculated, as discussed in section 3.10.2, for both male and female informants. The grammar scores of all informants who fall into a particular speaker group, in this case male or female, are added together and the resultant number is then divided by the total number of informants in that category, to produce a mean grammar score. Figure 5.1 shows the results. A high mean grammar score corresponds to high reported use of nonstandard grammatical variables, while a low mean grammar score corresponds to low reported use of nonstandard grammatical forms which, by extension, equates to a claim to use standard grammatical forms.

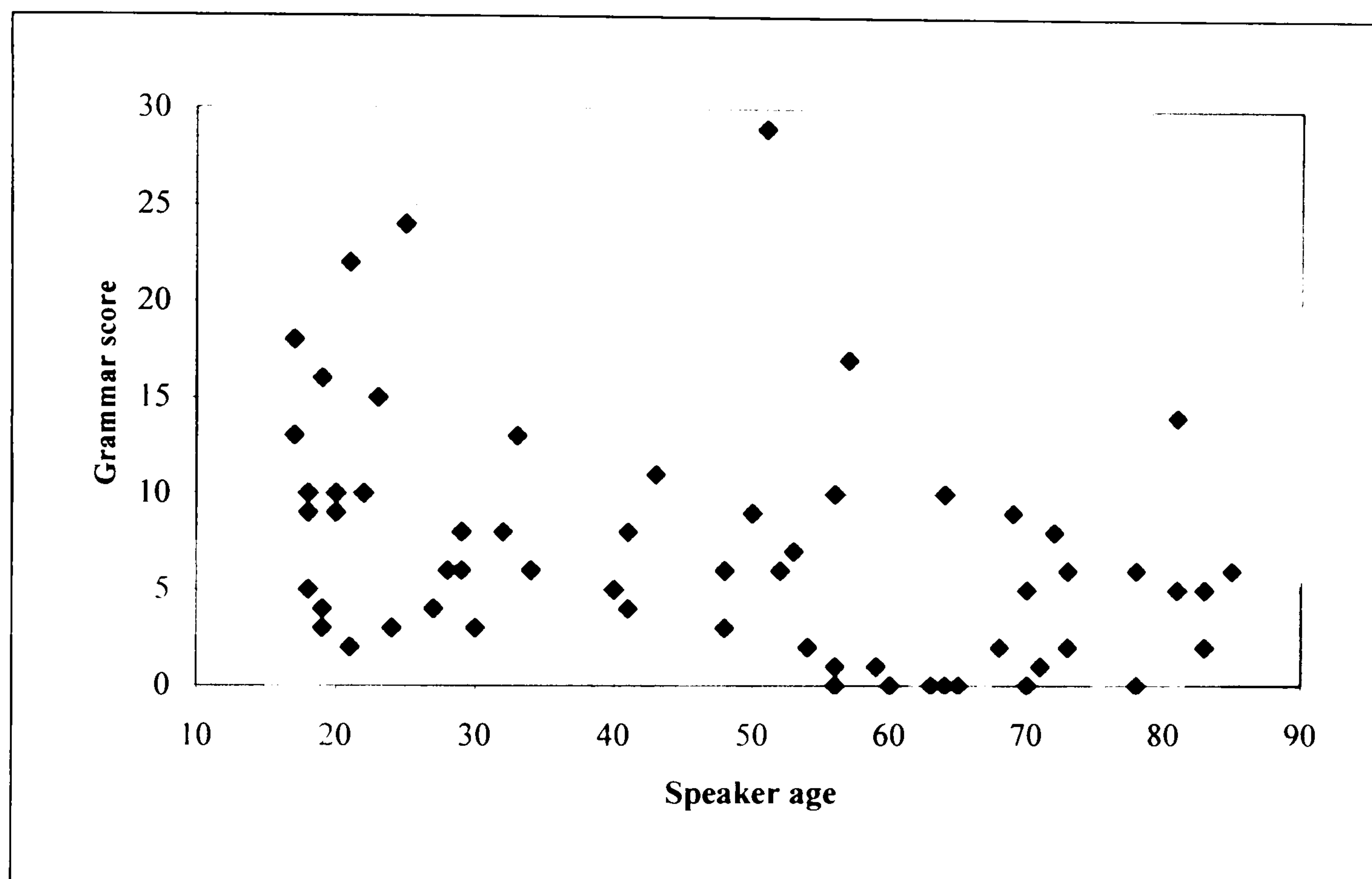
Figure 5.1 Mean grammar scores according to the speaker variable of sex

Given that previous research has shown that males tend to use more nonstandard or stigmatised variants than their female counterparts (Labov 1966; Trudgill 1972), the mean grammar scores are unsurprising. The mean score for male informants is 8.60, while the mean score for females is only 5.63. On average, men in the current study reported using a greater number of nonstandard grammatical features than did their female counterparts. Though the grammatical data analysed here are concerned with reported rather than actual usage, the findings support those of earlier studies. An examination of selected grammatical features according to speaker sex will show, however, whether this pattern is consistent for individual variants.

5.1.2 Grammar Scores and Speaker Age

Figure 5.2 shows individual grammar scores, which take into consideration reported speaker use of all the nonstandard features investigated by the Southampton area language questionnaire, according to the speaker variable of age. The higher the score, the more nonstandard grammatical forms the speaker claims to use.

Figure 5.2 Individual grammar scores according to the speaker variable of age



Individual grammar scores for informants under the age of 30 range from 2 to 24. For informants between the ages of 30 and 59, the range is between 0 and 29. For those aged 60 and over, the range is between 0 and 14. The ranges are, however, somewhat misleading, since they are affected by outlying scores. This can be seen particularly in the 30 to 59 age group, where the range is greatly increased by an outlying grammar score of 29 when all other scores for this group are under 18.

Variation in reported use of nonstandard grammatical variables according to speaker age is most clearly shown using a mean average for each age group. In order

to calculate the mean grammar score for each age group, the individual grammar scores for every informant belonging to a particular age group are added together and then divided by twenty, the number of informants in each age range. The results are shown in Figure 5.3.

Figure 5.3 Mean grammar scores according to the speaker variable of age

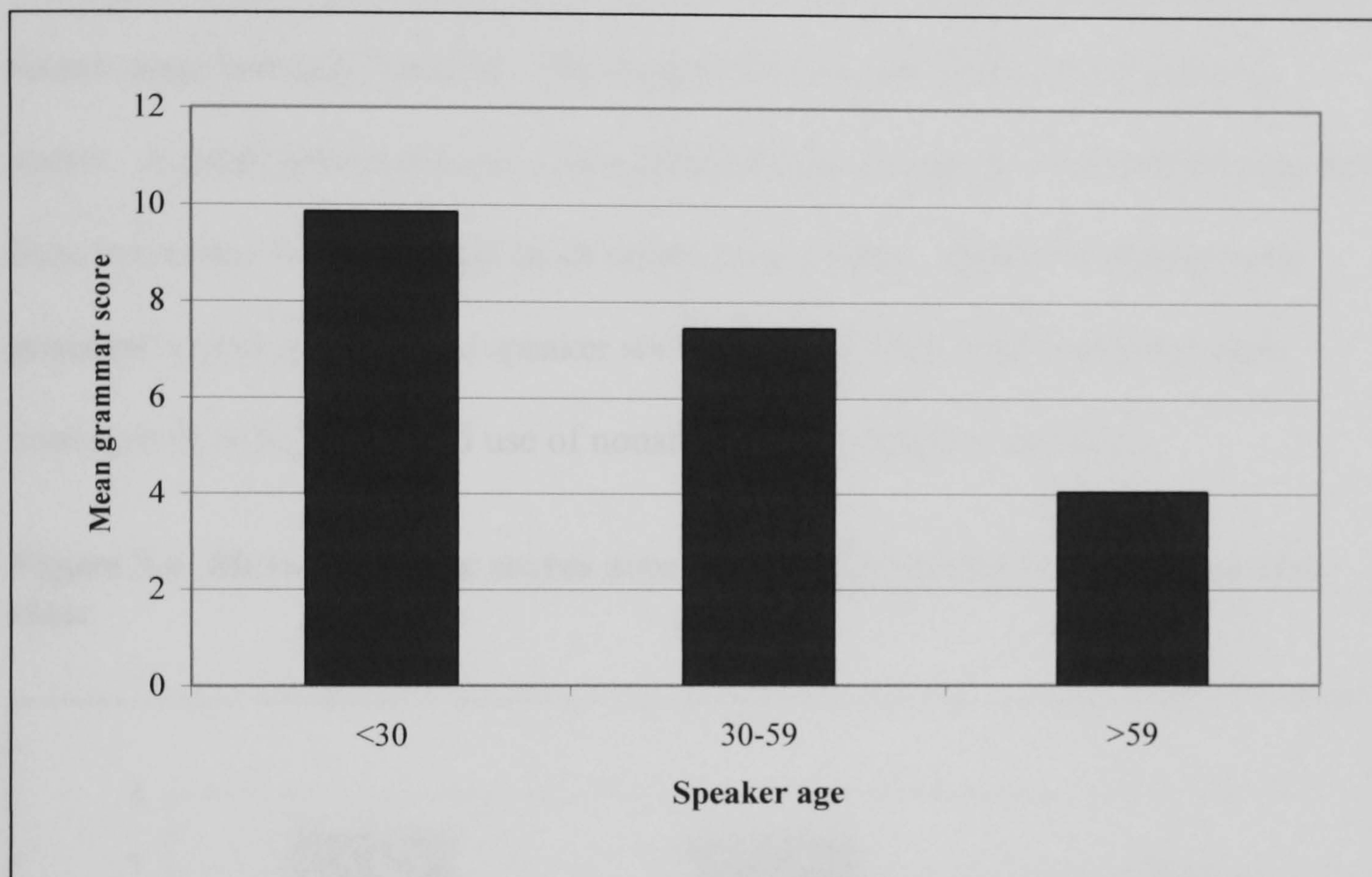


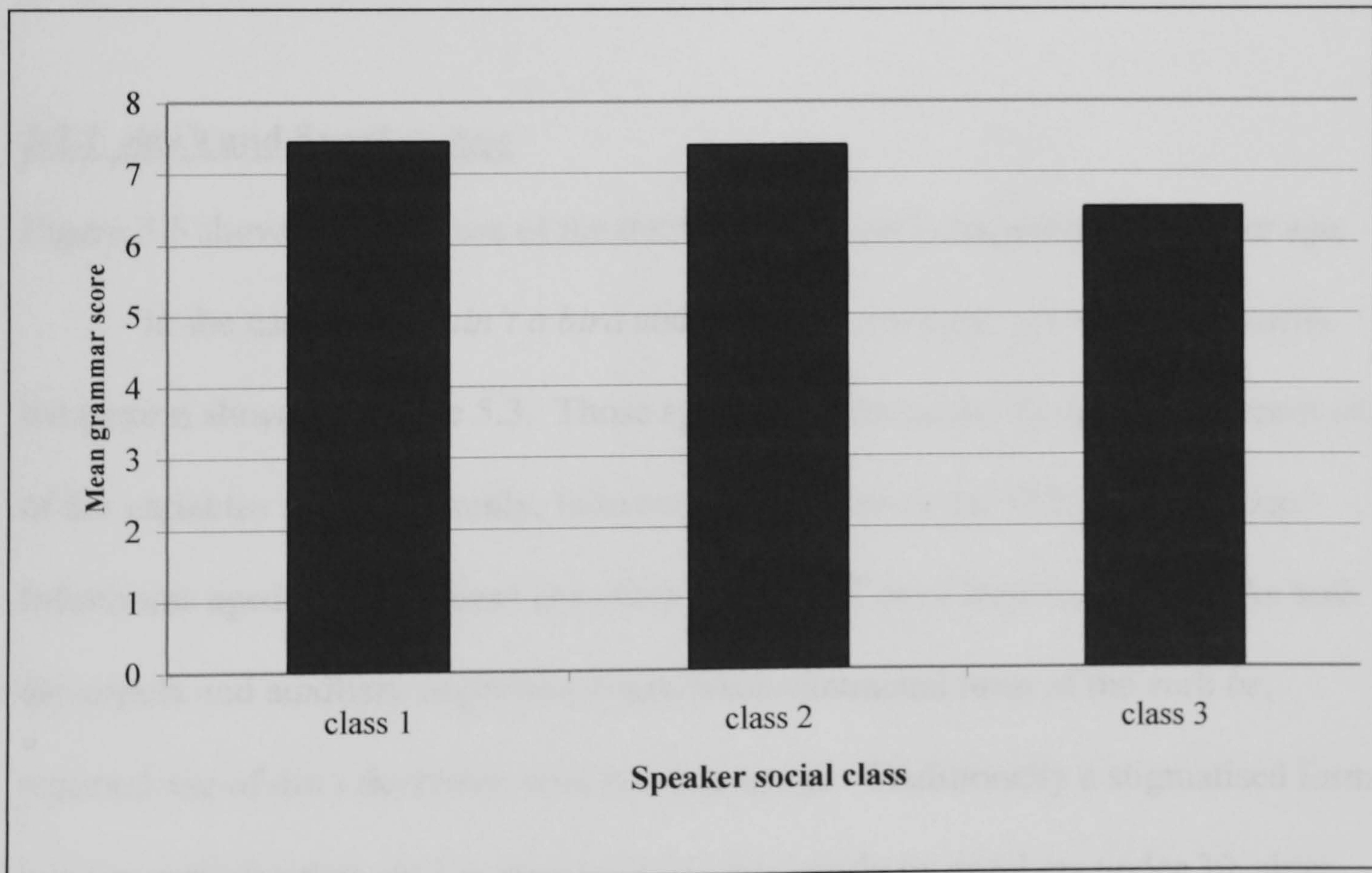
Figure 5.3 shows very clearly that the mean grammar score decreases steadily with increasing age. Informants under 30 report using on average 9.85 of the 45 grammatical features under examination. For speakers aged between 30 and 59, the average is 7.45. For those aged 60 and over, it is 4.05. It was perhaps to be expected that the age group with the highest average reported use of nonstandard grammatical features would be the under 30s, since use of nonstandard or stigmatised features is often associated with the young (Labov 1972a: 257; Chambers and Trudgill 1998: 79). An examination of a selection of nonstandard variables according to the speaker

variable of age will show, however, whether this pattern is consistent for individual variants.

5.1.3 Grammar Scores and Speaker Social Class

For class 1 informants, individual grammar scores range between 1 and 22. For class 2 informants, the range is larger, between 0 and 29. For class 3 informants, grammar scores range between 0 and 24. The ranges are, however, affected by outlying scores. A mean grammar score is less affected than a range by such outliers, and has been calculated for informants in all three social classes. Figure 5.4 shows mean grammar scores according to speaker social class. A high mean grammar score corresponds to high reported use of nonstandard grammatical variables.

Figure 5.4 Mean grammar scores according to the speaker variable of social class



Class 1 has a mean score of 7.45, class 2 a mean score of 7.39, and class 3 a mean grammar score of 6.5. Though nonstandard or low prestige variants are usually

more frequently found in the speech of those from lower socioeconomic backgrounds (Trudgill 1974: 61-63, Chambers 2003: 57-58), it is class 3 informants who, on average, report using the fewest nonstandard grammatical variables. Differences in mean grammar scores between the three social classes are small, however. Analysis of selected grammatical variables according to the speaker variable of social class will show whether the pattern shown in Figure 5.4 is repeated for individual variants.

5.2 *Ain't*

Ain't is investigated in the Southampton area language questionnaire as a negative present tense contracted form of the verbs *be*, both as the copula (*that ain't a bird*) and the auxiliary (*that ain't working*), and *have*, as the auxiliary (*I ain't got a clue*). In the following sections, *ain't* is analysed according to the extralinguistic variables of age and social class.

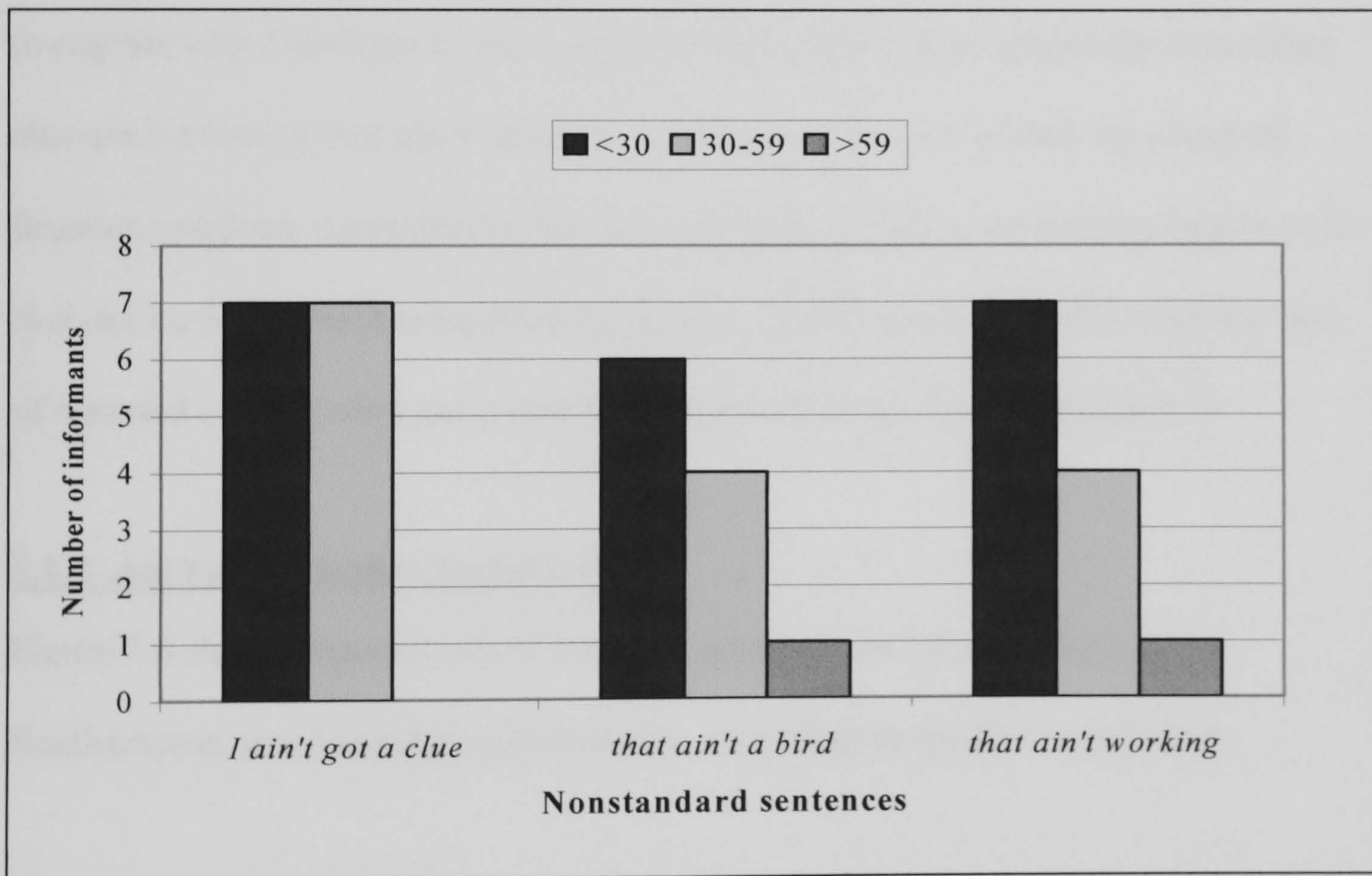
5.2.1 *Ain't* and Speaker Age

Figure 5.5 shows reported use of the three forms of *ain't*, according to speaker age.

In the case of *that ain't a bird* and *that ain't working*, *ain't* follows exactly the pattern shown in Figure 5.3. Those speakers in the under 30 age group report use of the variables most frequently, followed by speakers in the 30 to 59 age group. Informants aged over 59 report use of these forms of *ain't* least frequently. As both the copula and auxiliary negative present tense contracted form of the verb *be*, reported use of *ain't* decreases with increasing age. Traditionally a stigmatised form, it is unsurprising that *ain't* is reported most frequently by speakers under 30, since use of nonstandard features is often associated with the young (Labov 1972a: 257; Chambers and Trudgill 1998: 79). That *ain't* is reported more by younger

informants than by their older counterparts could mean that a change is in progress. It might be the case that use of the form is increasing, with young speakers leading this increase. Alternatively, the pattern shown might be an example of age-grading. Given that the present study occurs in apparent time, and that the data are concerned with reported rather than actual use, it is impossible to give a definitive explanation. It should be noted, however, that differences in reported use of *that ain't a bird* and *that ain't working* between age groups are not statistically significant at the 95% level. The chi-squared values of 2.699 and 3.984 respectively, do not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Figure 5.5 Reported use of *ain't* when talking to a friend according to the speaker variable of age



The findings for *ain't* as the negative present tense contracted auxiliary form of *have*, however, are slightly different. Though *I ain't got a clue* follows the same overall pattern as *ain't* as a form of the verb *be*, in that there is a sharp decline in reported use of the form between the youngest and oldest age groups, reported use of *ain't* as a form of *have* is the same across the youngest and the middle age groups.

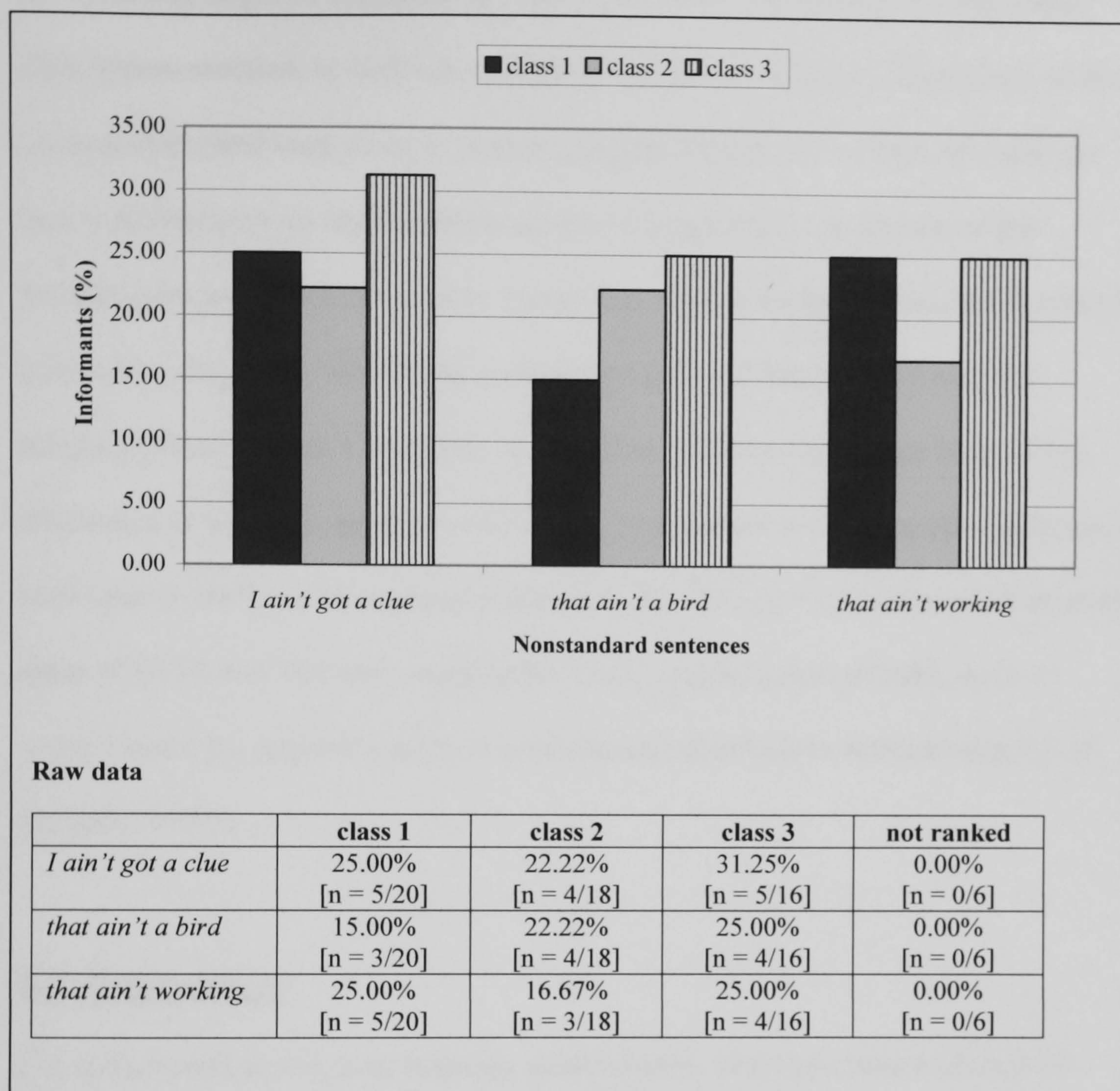
This is compared to a decline in reported use among 30 to 59-year-olds in the case of *that ain't a bird* and *that ain't working*. The second difference between the *be* and *have* forms of *ain't* is that, unlike *that ain't a bird* and *that ain't working*, *I ain't got a clue* is not reported by any informants in the 60 and over age group. The differences between age groups in reported use of *I ain't got a clue* are statistically significant at the 95% level, as the chi-squared value of 6.731 exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

One possible explanation for the difference in reported use between the *be* forms and the *have* form of *ain't* is that *I ain't got a clue*, unlike *that ain't a bird* and *that ain't working*, could perhaps be described as a stock phrase which might be employed by people who would not usually employ *ain't*. This would certainly help to explain why *I ain't got a clue* is reported marginally more frequently than either *that ain't a bird* or *that ain't working*: *I ain't got a clue* is reported by a total of fourteen speakers, compared to *that ain't working*, which is reported by twelve, and *that ain't a bird*, which is reported by eleven. It does not explain the complete lack of reported use of *I ain't got a clue* by speakers over the age of 59, however.

5.2.2 Ain't and Speaker Social Class

Figure 5.6 shows reported use of the three forms of *ain't* investigated by the Southampton area language questionnaire, according to speaker social class.

Figure 5.6 Reported use of *ain't* when talking to a friend according to the speaker variable of social class



As one might expect, given that nonstandard grammatical forms tend to be found more frequently in working class speech than in middle class speech (Trudgill 1974: 61-63; Chambers 2003: 57-58), it is those speakers in class 3 who, overall, report using *ain't* most frequently. However, it is only *that ain't a bird* which follows exactly the pattern of use that might be expected, with reported use increasing as speaker social class lowers. In the case of *I ain't got a clue*, though reported use of the form is lower among those in class 1 than it is in class 3, it is lowest for those in class 2. Similarly, reported use of *that ain't working* is lowest for class 2 speakers, though reported use in this instance is identical for class 1 and class

3 speakers. The lower reported use by class 2 speakers of *I ain't got a clue* and *that ain't working* might be explained by Labov's (1972b: 138) theory of lower middle class hypercorrection, in which he suggests that the less secure social position of the lower middle class leads them to employ prestige features more frequently than do their counterparts from higher socioeconomic backgrounds. In the case of the Southampton area study, it could be argued that this insecurity has resulted in class 2 informants claiming to use certain nonstandard grammatical constructions less frequently than do class 1 speakers. It should be noted, however, that none of the differences in reported use of *ain't* between social classes are statistically significant. *I ain't got a clue* has a chi-squared value of 0.074, *that ain't a bird* has a chi-squared value of 0.133, and *that ain't working* has a chi-squared value of 0.082, none of which exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

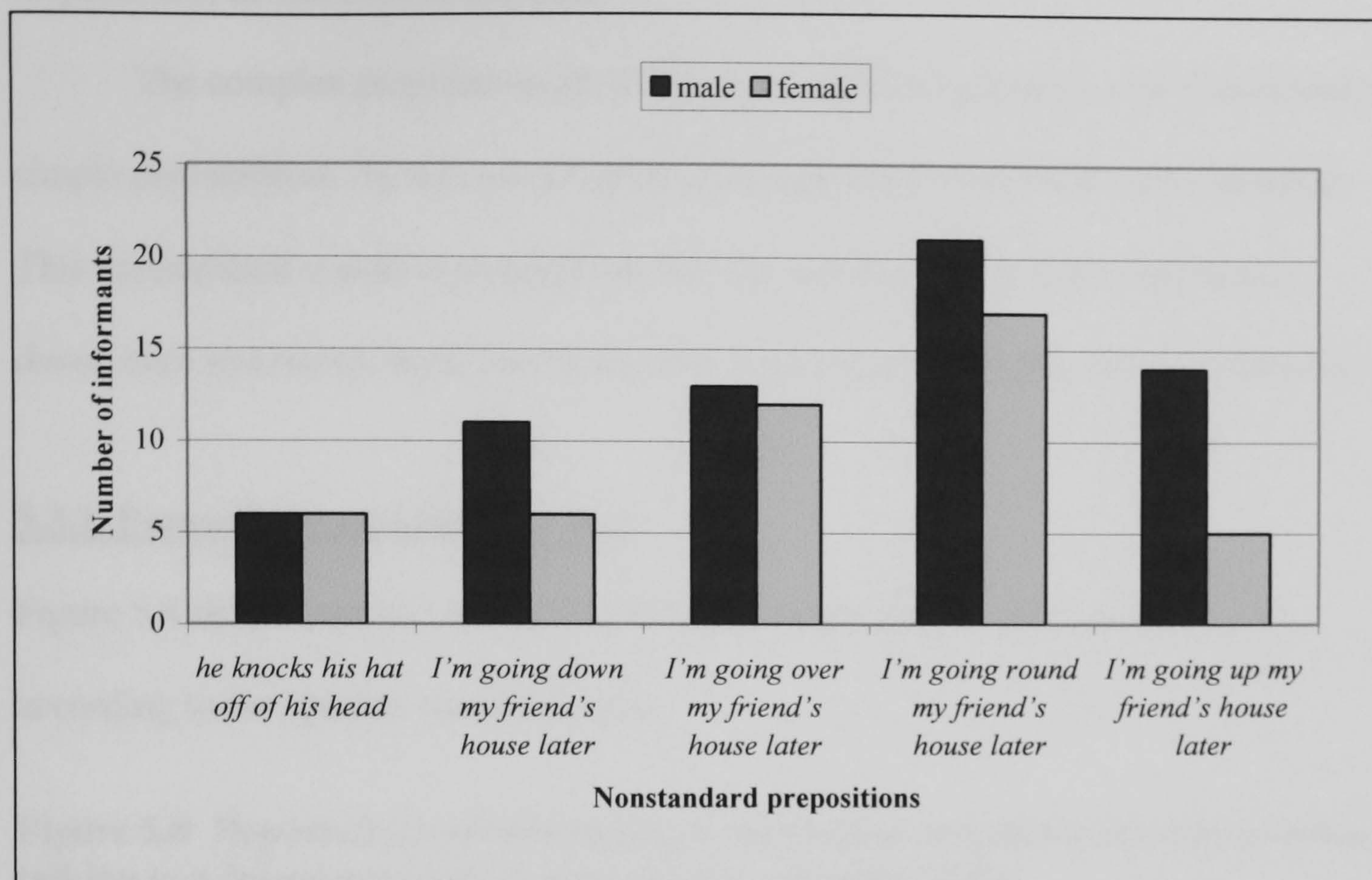
5.3 Prepositions

Using the Southampton area language questionnaire, data have been collected on both the use of simple prepositions where Standard English has complex prepositions, *I'm going up/down/over/round my friend's house later*, and the use of a complex preposition where Standard English has a simple preposition, *he knocks his hat off of his head*. In the following sections, reported use of these prepositions is analysed according to the speaker variables of sex, age and social class.

5.3.1 Prepositions and Speaker Sex

Figure 5.7 shows reported use of nonstandard simple and complex prepositions according to the speaker variable of sex.

Figure 5.7 Reported use of nonstandard simple and complex prepositions when talking to a friend according to the speaker variable of sex



In the case of each of the four nonstandard simple prepositions examined by the Southampton area language questionnaire, *I'm going up/down/over/round my friend's house later*, more male informants reported the use of these forms than did their female counterparts, supporting the overall trend shown in Figure 5.1. The greatest variation in male and female usage is found in *I'm going up my friend's house later*, reported by fourteen males and only five females, followed by *down*, reported by eleven males and six females, then *round*, reported by twenty-one males and seventeen females, and finally *over*, reported by thirteen males and twelve females. It is only *up*, however, where the difference in reported use between the sexes is statistically significant. The chi-squared value for *up* is 4.929, which exceeds the upper 5% point of a chi-squared distribution with one degree of freedom (3.841), making the difference in reported use between men and women significant at the 95% level. Men are more likely to use *up* than their female counterparts. In

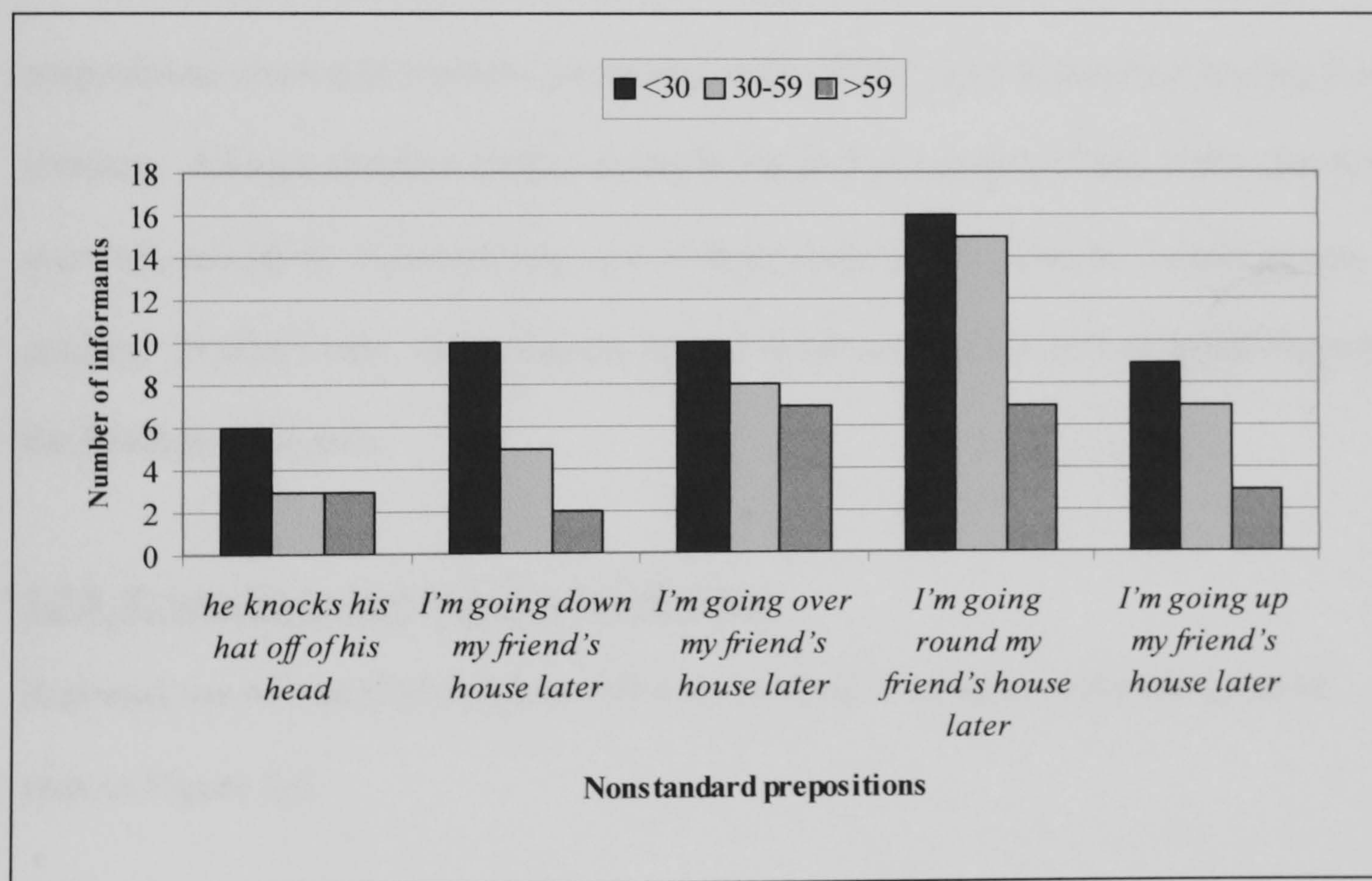
the case of *down*, *over* and *round*, the chi-squared values of 1.313, 0, and 0.646 respectively, do not exceed this point.

The complex preposition *off of* displays a different pattern to the nonstandard simple prepositions. In the case of *off of*, male and female reported use is identical. This nonstandard complex preposition, like the nonstandard simple prepositions *down*, *over* and *round*, does not function as a marker of sex in the Southampton area.

5.3.2 Prepositions and Speaker Age

Figure 5.8 shows reported use of nonstandard simple and complex prepositions according to the speaker variable of age.

Figure 5.8 Reported use of nonstandard simple and complex prepositions when talking to a friend according to the speaker variable of age



It is apparent from Figure 5.8 that reported use of nonstandard simple and complex prepositions declines with increasing speaker age. In the case of each of the five prepositions under examination, it is speakers in the under 30 age group who report the use of these nonstandard forms most frequently. The simple prepositions,

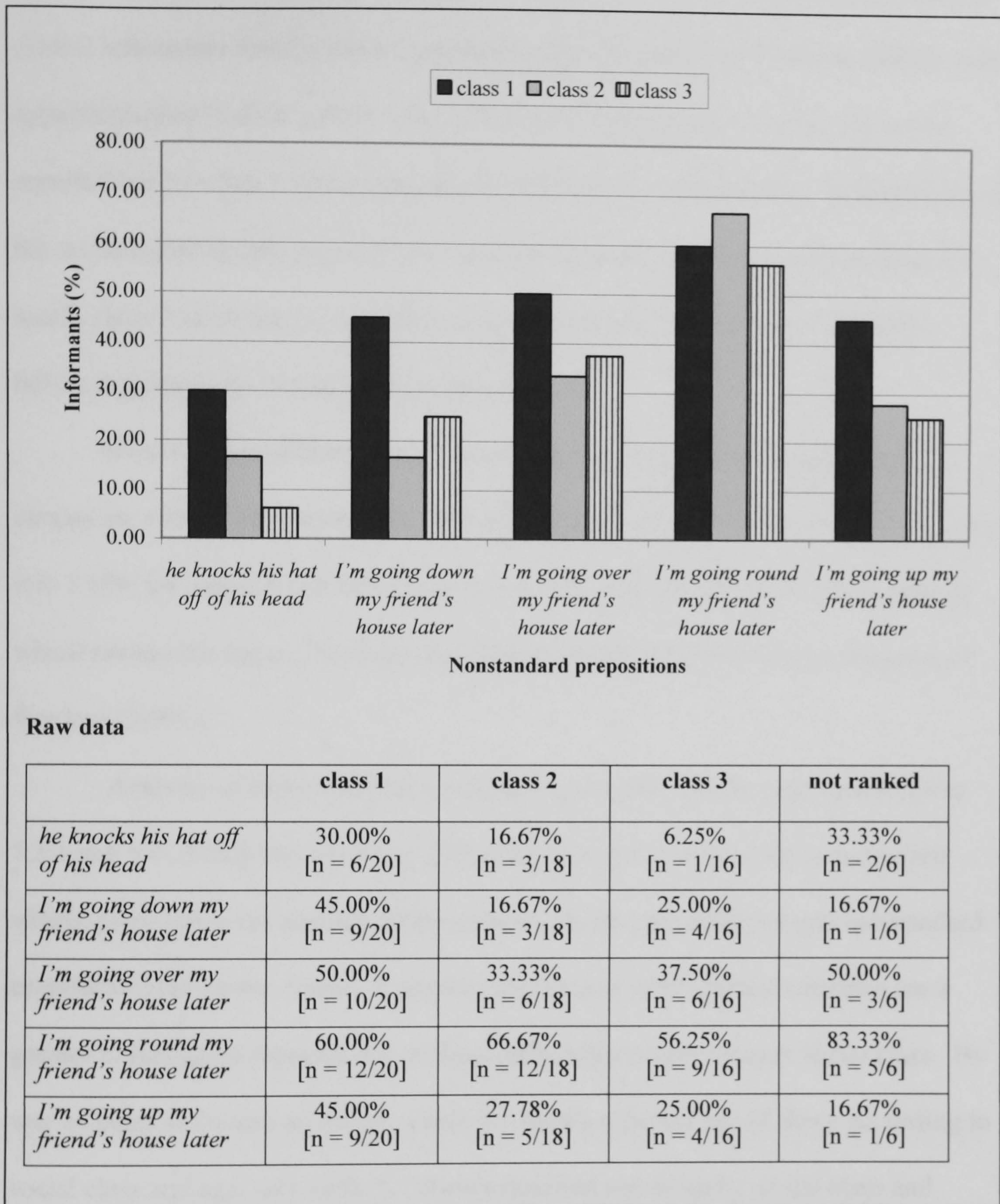
up, *down*, *over* and *round*, follow the same pattern as the mean grammar scores of Figure 5.3. The youngest speakers report the highest use of nonstandard simple prepositions, followed by speakers in the 30 to 59 age group. Speakers aged 60 and over have the lowest reported use. In the case of the complex preposition *off of*, while reported use does decrease with increasing speaker age, in that there is a decline in the level of reported use between the youngest and oldest age groups, reported use is the same across the 30-59 and 60 and over age groups. It is only in the case of *down* and *round*, however, that differences in reported use between age groups are statistically significant. *Down* has a chi-squared value of 8.044, while *round* has a chi-squared value of 10.478, both of which far exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

One possible explanation for these findings is that use of the nonstandard prepositions *down* and *round* is increasing, with the youngest informants leading the increase. A larger speaker sample would be needed to ascertain if this is the case for *over*, *up*, and *off of*. Alternatively, use of these prepositions could be subject to age-grading. In either case, these features appear to function as markers of speaker age in the Southampton area.

5.3.3 Prepositions and Speaker Social Class

Reported use of nonstandard prepositions according to speaker social class can be seen in Figure 5.9.

Figure 5.9 Reported use of nonstandard simple and complex prepositions when talking to a friend according to the speaker variable of social class



In the case of three of the four nonstandard simple prepositions, *down*, *over* and *up*, and the nonstandard complex preposition *off of*, it is speakers in class 1 who most frequently report using these variants. Exactly the opposite pattern to the one that might be expected for nonstandard variants is seen in the reported use of *off of* and *up*, with reported use decreasing, as opposed to increasing, as speaker social

class lowers. In the case of *down* and *over*, while claimed use decreases between class 1 and class 2, it rises again in class 3. The lower reported use of the variant by class 2 informants than by class 1 speakers might be explained by lower middle class hypercorrection (Labov 1972b: 138). This does not explain, however, the lower reported use by class 3 informants than by their class 1 counterparts. Reported use of the nonstandard simple preposition *round*, by contrast, increases between classes 1 and 2, class 2 being the group which reports using this variant most frequently, before dipping to its lowest level in class 3.

None of these differences in reported use of prepositions between social classes are statistically significant. The chi-squared value for *down* is 2.502, for *over* it is 1.189, for *round* it is 0.403, for *up* it is 1.975, and for *off of* it is 1.966, none of which exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Analysis of these variants by speaker sex and by speaker age (see sections 5.3.1 and 5.3.2) indicates clear sex differentiation, in the case of the nonstandard simple preposition *up*, and age differentiation, in the case of the simple nonstandard prepositions *down* and *round*. It appears that these extralinguistic variables are a greater influence on reported use of these prepositions than speaker social class. By way of support for this assertion, Table 5.1 shows reported use of *down* according to social class and age, and Table 5.2 shows reported use of *up* by social class and speaker sex. Both Tables show reported use by age or sex as a proportion of the total number of tokens of reported use by a particular social class.

Table 5.1 Reported use of *I'm going down my friend's house later* when talking to a friend according to the speaker variables of social class and age

	class 1	class 2	class 3
<30	88.89% [n = 8/9]	33.33% [n = 1/3]	25.00% [n = 1/4]
30-59	0.00% [n = 0/9]	33.33% [n = 1/3]	75.00% [n = 3/4]
>59	11.11% [n = 1/9]	33.33% [n = 1/3]	0.00% [n = 0/4]

Table 5.2 Reported use of *I'm going up my friend's house later* when talking to a friend according to the speaker variables of social class and sex

	class 1	class 2	class 3
male	66.67% [n = 6/9]	60.00% [n = 3/5]	100.00% [n = 4/4]
female	33.33% [n = 3/9]	40.00% [n = 2/5]	0.00% [n = 0/4]

Table 5.1 shows that the majority of class 1 speakers who report using *down* are under 30, while Table 5.2 shows that men, regardless of class, report the nonstandard simple preposition *up* more frequently than their female counterparts. These Tables support the assertion that age and sex appear to be a greater influence on reported use of certain nonstandard prepositions than speaker social class.

5.4 Nouns of Measurement with Zero Plurals

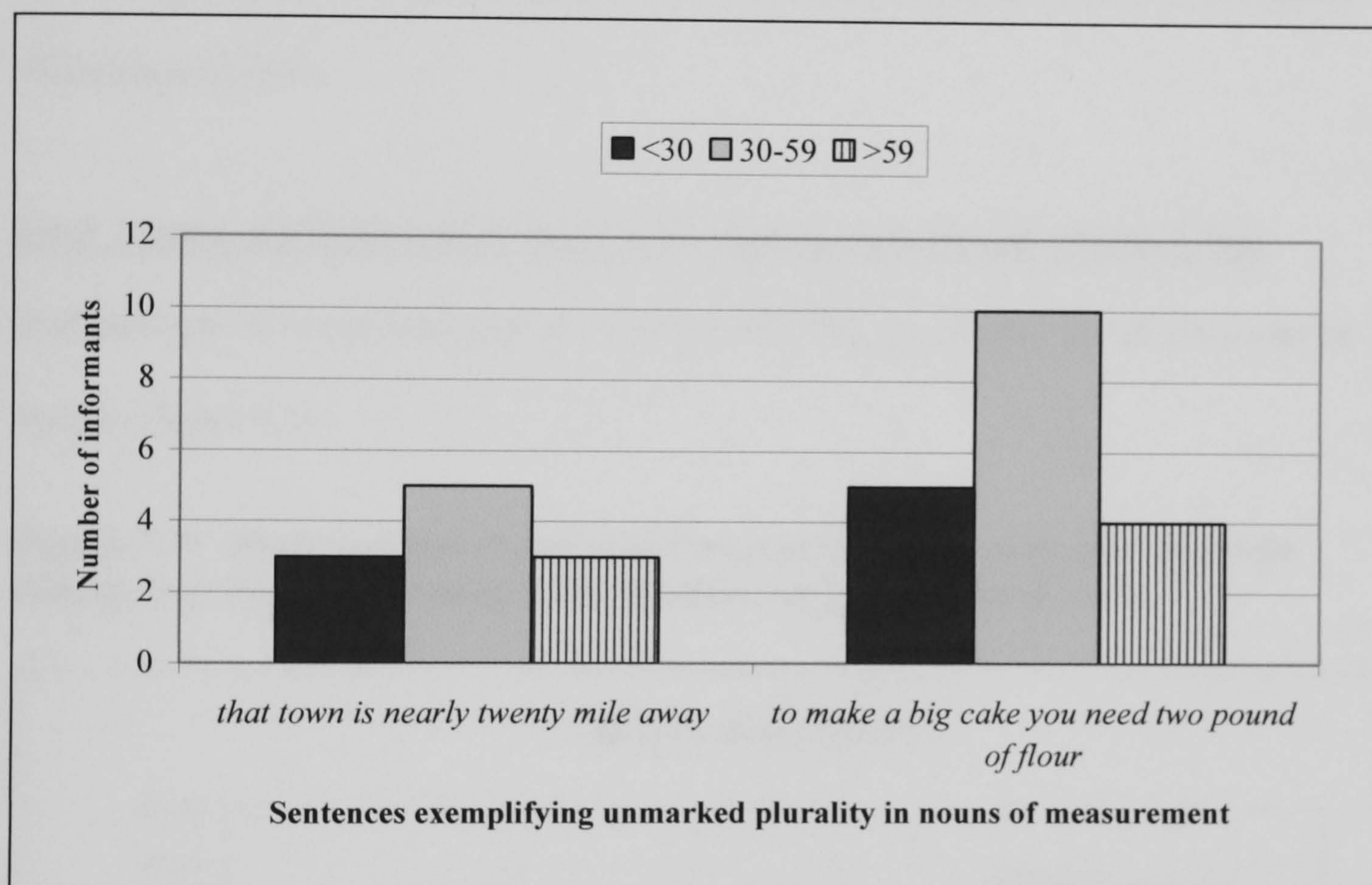
Three sentences exemplifying nouns of measurement with zero plurals are featured in the Southampton area language questionnaire: *that town is nearly twenty mile away*; *to make a big cake you need two pound of flour*; and *this string is three inch long*. *Two pound* is the most frequently reported of the three, followed by *twenty mile*. *Three inch* is reported by only four informants and so is excluded from analysis here on the grounds that to subdivide such a small sample would be

uninstructive. In the following sections, *twenty mile* and *two pound* are analysed according to the extralinguistic variables of speaker age and social class.

5.4.1 Nouns of Measurement with Zero Plurals and Speaker Age

Reported use of *twenty mile* and *two pound* according to speaker age can be seen in Figure 5.10.

Figure 5.10 Reported use of nouns of measurement with zero plurals when talking to a friend according to the speaker variable of age



The pattern which emerges when reported use of unmarked plurality in nouns of measurement is analysed by speaker age is very different from that found in the mean grammar scores of Figure 5.3. Unlike the mean grammar scores, which show reported use of nonstandard grammatical forms to be highest among the youngest speakers, and lowest among the oldest speakers, it is the speakers in the 30 to 59 age group who have the highest reported use of this particular feature, with use across the

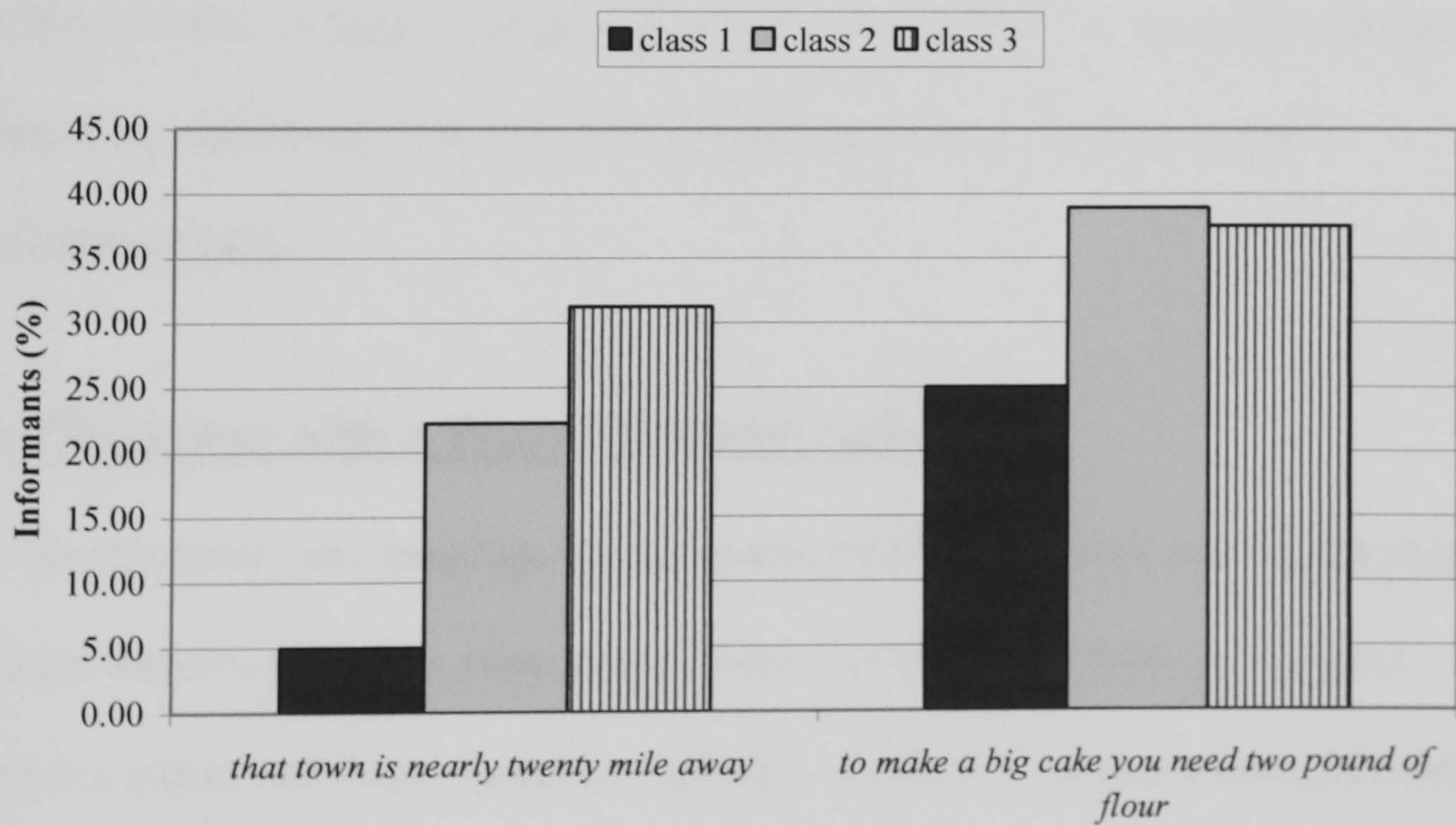
older and younger age groups being equal, in the case of *twenty mile*, or nearly equal, in the case of *two pound*.

High reported use of *twenty mile* and *two pound* among the middle age group is surprising, given that this is the age group typically most influenced linguistically by the standard (Chambers and Trudgill 1998: 79). The differences in reported use between age groups of these two forms are not, however, statistically significant. The chi-squared value for *twenty mile* is 0.250, and for *two pound* it is 4.775, neither of which exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

5.4.2 Nouns of Measurement with Zero Plurals and Speaker Social Class

Reported use of *twenty mile* and *two pound* according to speaker social class can be seen in Figure 5.11.

Figure 5.11 Reported use of nouns of measurement with zero plurals when talking to a friend according to the speaker variable of social class



Sentences exemplifying unmarked plurality in nouns of measurement

Raw data

	class 1	class 2	class 3	not ranked
<i>that town is nearly twenty mile away</i>	5.00% [n = 1/20]	22.22% [n = 4/18]	31.25% [n = 5/16]	16.67% [n = 1/6]
<i>to make a big cake you need two pound of flour</i>	25.00% [n = 5/20]	38.89% [n = 7/18]	37.50% [n = 6/16]	16.67% [n = 1/6]

As one might expect, given that nonstandard grammatical forms are most commonly associated with the speech of those from lower socioeconomic classes (Trudgill 1974: 61-63; Chambers 2003: 57-58), it is class 1 which has the lowest percentage of informants claiming to use the invariant plural noun forms. Reported use of *that town is nearly twenty mile away* increases as the class of the speakers lowers. *To make a big cake you need two pound of flour* patterns slightly differently. Though, as in the case of *twenty mile*, reported use of this form is higher for classes 2 and 3 than it is for class 1, there is no increase in reported use between the middle and lowest classes. Rather, reported use of *two pound* by the middle group is marginally higher, accounting as it does for 38.89% of class 2 responses and 37.50% of class 3 responses. It is still speakers from the highest social class, however, who report *two pound* least frequently. Again, however, differences in reported use between social classes of these two forms are not statistically significant. The chi-squared value for *twenty mile* is 2.598, and for *two pound* it is 1.000, neither of which exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). A larger speaker sample would be needed to determine whether nouns of measurement with zero plurals function as class markers in the Southampton area.

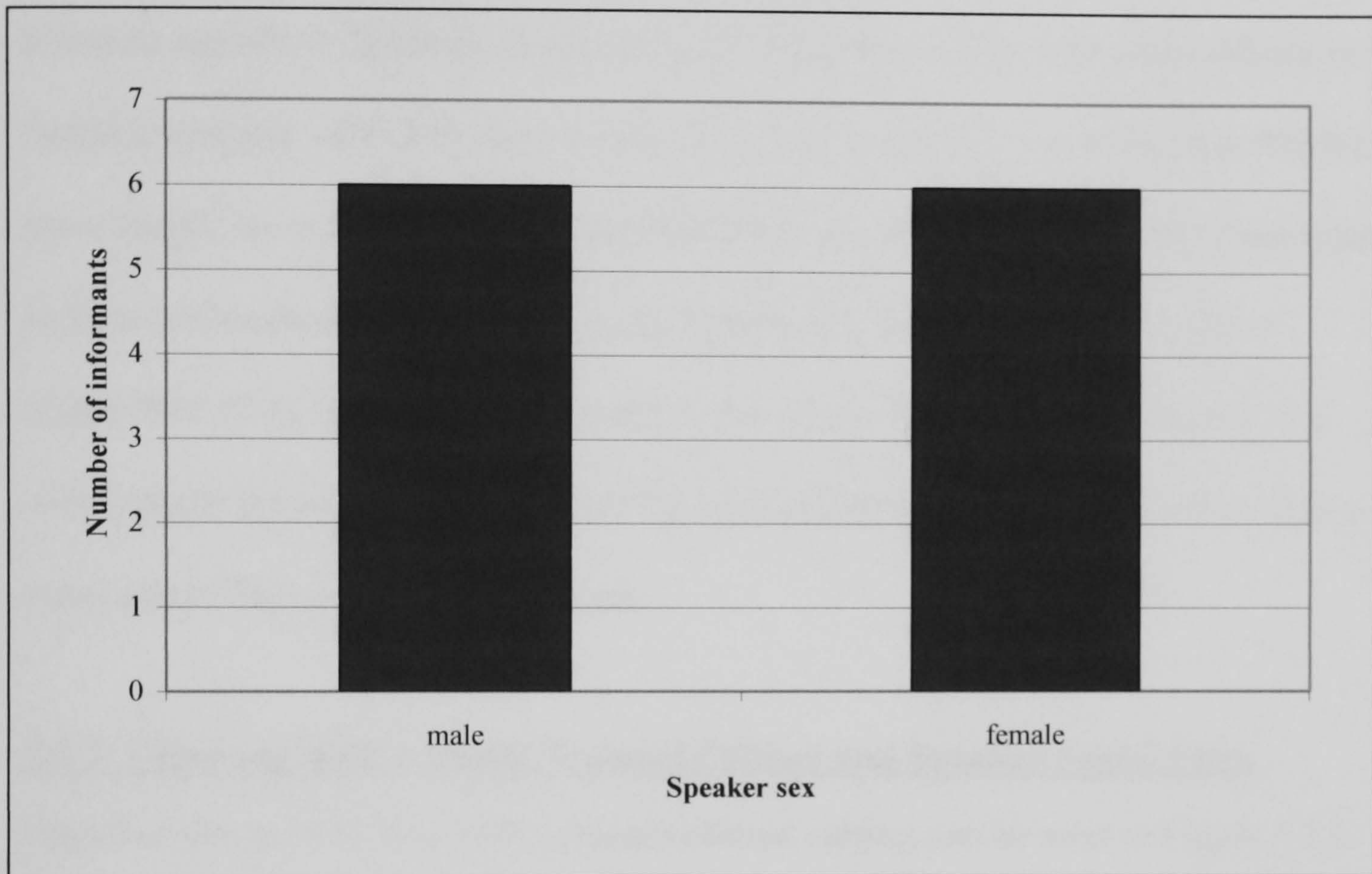
5.5 There was with a Plural Notional Subject

The Southampton area language questionnaire investigated *there was* with a plural notional subject, *there was some singers here a minute ago*, where in Standard English a plural verb form would be expected. In the following sections, this feature is analysed according to the extralinguistic variables of sex and social class.

5.5.1 *There was* with a Plural Notional Subject and Speaker Sex

Figure 5.12 shows reported use of *there was* with a plural notional subject, according to the speaker variable of sex.

Figure 5.12 Reported use of *there was* with a plural notional subject when talking to a friend according to the speaker variable of sex



As in the case of *he knocks his hat off of his head, there was some singers here a minute ago* shows no signs of sex-based variation, as the construction was reported by exactly the same number of male and female informants. Six male and six female respondents reported using *there was* with a plural notional subject when talking to a friend.

As discussed in section 1.7.4.1, research conducted by Cheshire (1999: 70) in Milton Keynes shows that *there was* does not show the kind of 'sharp pattern of variation' exhibited by features such as multiple negation. Instead, invariant *there* constructions are used 'by both female and male speakers in both social class groups [working-class and middle-class] almost one hundred per cent of the time' (Cheshire

1999: 70). Relatively small though the sample size is, the Southampton area data nevertheless broadly support Cheshire's findings from Milton Keynes.

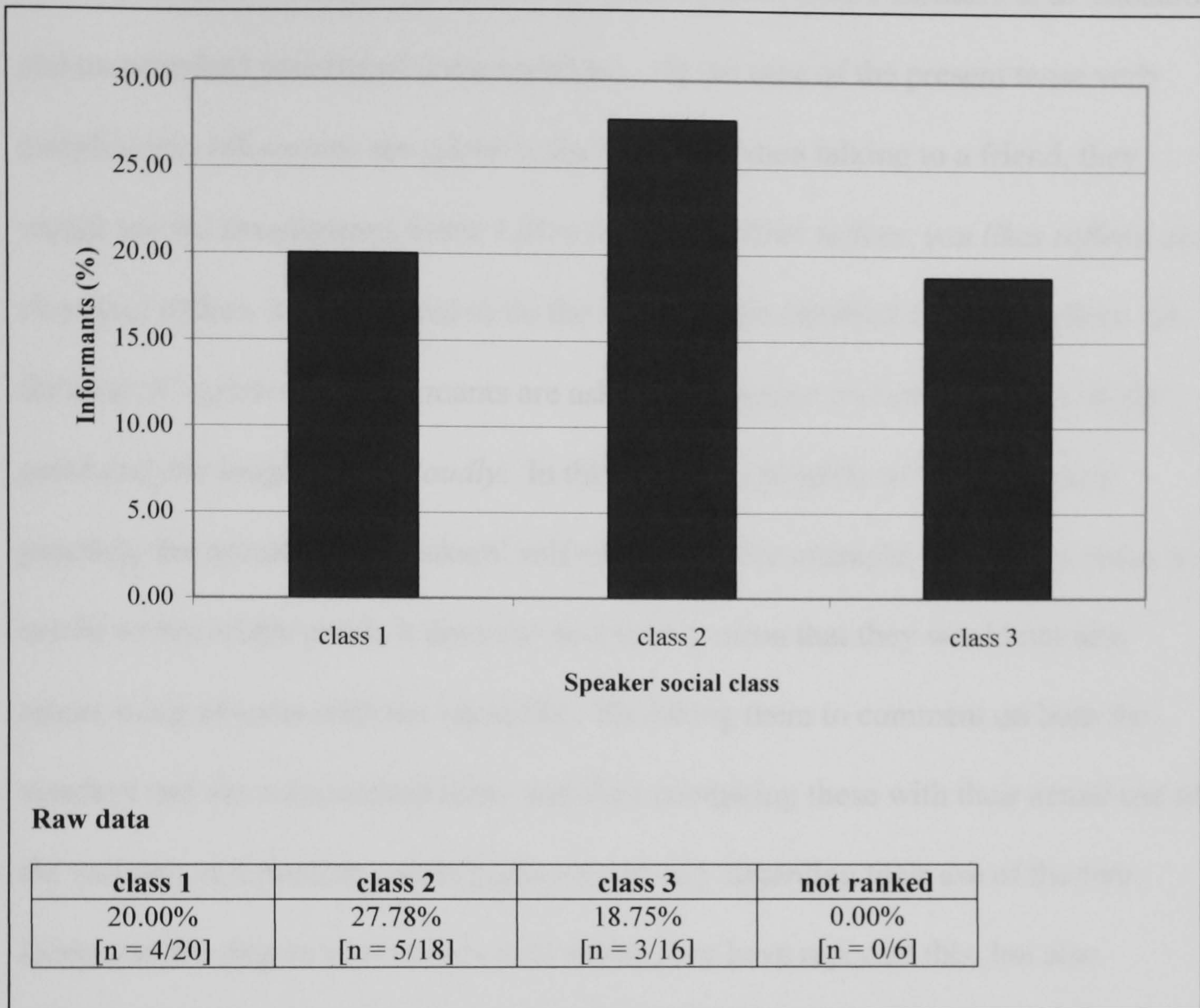
Given the widespread nature of its use in the Survey of British Dialect Grammar, Cheshire et al. (1989: 200) argue that *there was* with a plural subject is not nonstandard but rather a stylistic feature of English characteristic of informal speech. It can be seen from Appendix 4 that when reported use of this form when talking to a friend is coupled with affirmative responses to the statement 'I wouldn't use this but some people do use it in Southampton/Eastleigh Borough', the feature is widespread in the Southampton area, being reported as being used or heard used locally by nearly 70% of all informants. Lack of sex-based variation in the use of *there was* with a plural subject provides support for its classification as a stylistic rather than a nonstandard feature, in speech at least.

5.5.2 *There was* with a Plural Notional Subject and Speaker Social Class

Reported use of *there was* with a plural notional subject can be seen in Figure 5.13, according to speaker social class.

There was with a plural notional subject is least frequently reported by the lowest social class. Reported use is marginally higher in class 1, and is highest in class 2. These differences in reported use between social classes are not, however, statistically significant. The chi-squared value of 0.082 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). The findings of the Southampton area study appear to support Cheshire's (1999: 70) assertion that invariant *there was* does not exhibit sharp social class variation, and is to be found in the speech of both working- and middle-class people.

Figure 5.13 Reported use of *there was* with a plural notional subject when talking to a friend according to the speaker variable of social class



5.6 Speaker Self-reporting Versus Actual Use

Recordings of each of the 60 Southampton area informants have been listened to in full for tokens of present tense verb morphology and *-ly* adverbs, in order to compare informants' actual use of these forms with their claimed use of them in the Southampton area language questionnaire. These two features have been chosen for further analysis as the SuRE interviews have yielded sufficient tokens of these variables for meaningful conclusions to be drawn regarding both the use of these forms in the Southampton area and the matter of speaker self-reporting versus actual use. This might not have been the case had other less frequently-occurring variables included in the Southampton area language questionnaire been selected.

An additional advantage to selecting present tense verb morphology and *-ly* adverbs for further analysis is that the grammar questionnaire includes both standard and nonstandard variants of these variables. In the case of the present tense verb morphology, informants are asked to say whether, when talking to a friend, they would use the nonstandard forms *I likes toffees*, *we likes toffees*, *you likes toffees*, and *they likes toffees*, and are asked to do the same for the standard *she likes toffees*. In the case of *-ly* adverbs, informants are asked to comment on both *he writes really quick* and *she laughs really loudly*. In this way, it is possible to ascertain more precisely the accuracy of speakers' self-reporting. For example, if speakers claim to use *he writes really quick*, it does not necessarily mean that they would not also report using adverbs with the *-ly* suffix. By asking them to comment on both the standard and the nonstandard form, and then comparing these with their actual use of the variants, information can be gathered not only regarding their use of the two forms and the degree of accuracy with which they have reported this, but also regarding their attitude to the variables under examination.

5.6.1 Present Tense Verb Morphology

Table 5.3 shows reported use of standard and nonstandard present tense verb morphology by the 60 Southampton area informants.

Table 5.3 Reported use of standard and nonstandard present tense verb morphology

Variable	Number of informants claiming to use this form when talking to a friend
<i>I likes toffees</i>	3/60
<i>You likes toffees</i>	0/60
<i>She likes toffees</i>	35/60
<i>We likes toffees</i>	0/60
<i>They likes toffees</i>	1/60

Reported use of nonstandard verb forms, *I likes toffees*, *you likes toffees*, *we likes toffees* and *they likes toffees*, is very low. Conceivable explanations for this low level of reported use include the possibility that nonstandard *-s* is simply not very common in the Southampton area or, alternatively, that these are forms that speakers did not wish to disclose using.

That the standard form *she likes toffees* was reported by only thirty-five out of a possible sixty informants is notable, particularly as the other two control sentences (see section 3.5.2), *she laughs really loudly* and *I've got four dozen eggs*, were reported by 46 and 45 informants respectively. This unexpectedly low level of reported use is either the result of some informants misreading the language questionnaire, or it suggests that a nonstandard alternative is used by some of those informants who did not report using *she likes toffees*.

Table 5.4 shows reported use of both the standard third person singular form and nonstandard *I likes* and *they likes* by speaker age and sex.

Table 5.4 Reported use of standard and nonstandard present tense verb morphology according to the speaker variables of age and sex

	<i>I likes toffees</i>		<i>They likes toffees</i>		<i>She likes toffees</i>	
	Male	Female	Male	Female	Male	Female
<30	1/10	0/10	0/10	0/10	5/10	6/10
30-59	0/10	0/10	0/10	0/10	7/10	9/10
>59	2/10	0/10	0/10	1/10	3/10	5/10

Overall, a higher number of male speakers than female reported using nonstandard present tense morphology. However, the total number of reports is low, and the difference in reported use between sexes is not statistically significant, as the chi-squared value of 0.268 does not exceed the upper 5% of a chi-squared distribution with one degree of freedom (3.841). Reported use of *she likes toffees* is much higher than reported use of the nonstandard forms. As might be expected

given that women tend to report employing more standard or prestige forms than their male counterparts (Trudgill 1972: 187), more female informants than male informants reported using the standard. However, this difference in reported use of *she likes toffees* between the sexes is again not statistically significant, as the chi-squared value of 1.097 does not exceed the upper 5% of a chi-squared distribution with one degree of freedom (3.841). In the case of both male and female informants, it is those in the 30 to 59 age group who reported using this form most frequently. The differences in reported use of *she likes toffees* between the age groups are statistically significant, as the chi-squared value of 6.720 exceeds the upper 5% of a chi-squared distribution with two degrees of freedom (5.991). This supports previous findings which state that the influence of the standard is greatest on those in the middle age group (Chambers and Trudgill 1998: 79).

Table 5.5 shows reported use of both the standard third person singular form and *I likes* and *they likes* by speaker social class. As the numbers of informants are not the same in the three analytic classes, results are given as a percentage of the total number of informants in the class, as well as in tokens.

Table 5.5 Reported use of standard and nonstandard present tense verb morphology according to the speaker variable of social class

	<i>I likes toffees</i>	<i>They likes toffees</i>	<i>She likes toffees</i>
Class 1 – Managerial and professional occupations	0% [n = 0/20]	0% [n = 0/20]	50% [n = 10/20]
Class 2 – Intermediate occupations	5.56% [n = 1/18]	0% [n = 0/18]	61.11% [n = 11/18]
Class 3 – Routine and manual occupations	12.50% [n = 2/16]	6.25% [n = 1/16]	50% [n = 8/16]
Not ranked	0% [n = 0/6]	0% [n = 0/6]	100% [n = 6/6]

In the case of both of the nonstandard forms, reported use is highest in class 3. No-one in class 1 reports either of these forms. In class 2, one informant reports

using *I likes toffees* but none of the informants in this group report using *they likes toffees*. These differences in reported use are not statistically significant, however, as the chi-squared value for reported use of nonstandard present tense verb morphology according to social class is 2.300, which does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

In the case of the standard form *she likes toffees*, reported use by classes 1 and 3 is identical, while for class 2 it is slightly higher, a finding that might be explained by Labov's (1972b: 124-138) theory of lower middle class hypercorrection. This theory states that the greater sensitivity of speakers from the second highest social class to the social significance of linguistic variables sometimes results in greater use or reported use of prestigious or standard forms by this group than by the highest socioeconomic group. The differences in reported use between social classes are not statistically significant, however, as the chi-squared value of 0.596 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Table 5.6 shows actual use of standard and nonstandard present tense verb morphology during the SuRE interviews. Results are expressed as percentages as well as tokens in order to account for the varying number of tokens of each variable collected.

It is only in the case of the first person singular present tense that any tokens of nonstandard present tense verb use are recorded. Overall, the low level of actual use of nonstandard *-s* found in the interviews supports the low level of claimed use of nonstandard present tense verbs, shown in Table 5.3. The uniform use of a standard third person singular verb form by all informants is, however, very different to the reported use of this form.

Table 5.6 Actual use of standard and nonstandard present tense verb morphology

Variable	Standard tokens	Nonstandard tokens
First person singular present tense	99.85% [n = 1959/1962]	0.15% [n = 3/1962]
Second person present tense	100% [n = 829/829]	0% [n = 0/829]
Third person singular present tense	100% [n = 379/379]	0% [n = 0/379]
First person plural present tense	100% [n = 218/218]	0% [n = 0/218]
Third person present tense plural	100% [n = 352/352]	0% [n = 0/352]
Total number of tokens	3737	3

Table 5.7 shows speaker self-reporting of the present tense verb forms investigated by the Southampton area questionnaire. The term ‘over-report’ refers to when a speaker claims to use a particular form but no tokens of the said form are found in the subsequent interview, while the term ‘under-report’ refers to when a speaker claims not to use a form but produces it in the interview (Trudgill 1972: 184-185; Newbrook 1999: 100). ‘Accurate claim’ refers to when a speaker’s claim to use, or not to use, a particular form is borne out by their use or non-use respectively of that form in the interview.

Table 5.7 Speaker self-reporting of present tense verb morphology

Variable	Accurate claim	Over-report	Under-report	Discounted
<i>I likes toffees</i>	54/60	3/60	3/60	N/A
<i>You likes toffees</i>	60/60	0/60	0/60	N/A
<i>She likes toffees</i>	31/60	0/60	24/60	5/60
<i>We likes toffees</i>	60/60	0/60	0/60	N/A
<i>They likes toffees</i>	59/60	1/60	0/60	N/A

It can be seen from Table 5.7 that the number of speakers who accurately reported their use of the four nonstandard sentences under investigation is very high.

There is 100%, or near 100%, accuracy of reporting for all of these nonstandard forms.

A total of four informants over-reported their use of the nonstandard forms, claiming to use nonstandard *-s* when no tokens of this were found in the subsequent interview, while three informants under-reported their use, claiming not to use nonstandard *-s* but producing it in the interview. The use of nonstandard *-s* by the three informants who under-reported was, however, variable: the vast majority of their utterances were standard.

The number of accurate claims of reported use of the standard third person singular present tense is low when compared to the number of accurate claims of reported use of nonstandard *-s*. Thirty-one out of sixty informants accurately reported their own use of standard *-s*. Of the remaining twenty-nine informants, twenty-four under-reported their use of the form, claiming not to use standard *-s* but employing it in their interviews. The other five responses have been discounted, four of these because the informants did not employ the construction, standard or otherwise, in the course of the SuRE interview, so rendering an assessment of the accuracy of their self-reporting impossible, and the fifth because it was a double response, the informant having both claimed to use and claimed not to use *she likes toffees*. This speaker did not produce any tokens of *she likes toffees* in the SuRE interview and so is not included in the count of actual use shown in Table 5.6 either.

Table 5.8 provides a breakdown of under- and over-reporting of present tense verb morphology according to speaker age and sex.

Table 5.8 Under- and over-reporting of present tense verb morphology use according to the speaker variables of age and sex

	<i>I likes toffees</i>				<i>They likes toffees</i>				<i>She likes toffees</i>			
	Male		Female		Male		Female		Male		Female	
	Under	Over	Under	Over	Under	Over	Under	Over	Under	Over	Under	Over
<30	1/10	1/10	0/10	0/10	0/10	0/10	0/10	0/10	5/10	0/10	4/10	0/10
30-59	1/10	0/10	0/10	0/10	0/10	0/10	0/10	0/10	3/10	0/10	1/10	0/10
>59	0/10	2/10	1/10	0/10	0/10	0/10	0/10	1/10	7/10	0/10	4/10	0/10

Given that males tend to attach covert prestige to nonstandard or stigmatised forms (Trudgill 1972: 187), to find more males than females over-reporting their use of the nonstandard form *I likes toffees* is perhaps unsurprising. Overall, however, the number of under- and over-reports of *I likes toffees* is small, and the difference in over-reporting between the sexes is not statistically significant, as the chi-squared value of 1.404 does not exceed the upper 5% point of chi-squared distribution with one degree of freedom (3.841).

There are no over-reports of *she likes toffees*. Fifteen of the twenty-four under-reports of the third person singular present tense form were made by male informants and nine by female informants. As *she likes toffees* is a standard form, this is perhaps to be expected in the case of the male informants, given that they have been shown to claim to use fewer prestigious forms than they actually do (Trudgill 1972: 187). In Trudgill's Norwich phonological self-evaluation test, male informants were 'much more likely to *under-report* [and] female informants to *over-report* [original italics]' (Trudgill 1972: 187) their use of prestige forms. Of this finding, Trudgill states:

This, then, is the objective evidence which demonstrates that male speakers, at least in Norwich, are at a subconscious or perhaps simply private level very favourably disposed towards non-standard speech forms.

(Trudgill 1972: 187)

This tendency on the part of men to opt for nonstandard forms provides a plausible explanation for the Southampton area findings. were it not for the fact that male informants in the present study do not have high levels of over-reporting of the four sentences illustrating nonstandard present tense morphology. It might be the case that nonstandard *-s* is not a sufficiently salient variant in the Southampton area for male informants to select it as a covertly prestigious form.

That nearly a third of all female informants would under-report their use of *she likes toffees* is notable. Women tend to favour standard or prestige forms and have been shown to select these in self-reporting exercises, often over-reporting their use of such variables (Trudgill 1972: 187-188). One explanation for the unexpectedly high number of under-reports of *she likes toffees*, particularly among female informants, is a methodological one. It might be that, seeing *she likes toffees* grouped with four sentences exemplifying nonstandard *-s* in the present tense verb morphology, informants simply presumed that this sentence was also nonstandard and so did not claim to use it.

The highest number of under-reports, eleven in total, is found in the 60 and over age group, followed by nine in the under 30 age group, and four in the 30 to 59 age group. That the lowest number of under-reports is found in the middle age group is to be expected, given that the standard has a greater influence on this group than on any other age group (Chambers and Trudgill 1998: 79). The differences in under-reporting of *she likes toffees* between age groups are significant at the 90% level, as the chi-squared value of 5.417 exceeds the upper 10% point of a chi-squared distribution with two degrees of freedom (4.605). Speakers aged 30 to 59 are less likely than those under 30 or over 59 to under-report their use of *she likes toffees*.

In the case of *they likes toffees*, there are no under-reports, and the sole over-report is made by a woman in the over 59 age group. However, as there is just one report, no meaningful conclusions can be drawn.

Table 5.9 provides a breakdown of under- and over-reporting of present tense verb morphology according to speaker social class.

Table 5.9 Under- and over-reporting of present tense verb morphology use according to the speaker variable of social class

	<i>I likes toffees</i>		<i>They likes toffees</i>		<i>She likes toffees</i>	
	Under	Over	Under	Over	Under	Over
Class 1 – Managerial and professional occupations	5.00% [n = 1/20]	0.00% [n = 0/20]	0.00% [n = 0/20]	0.00% [n = 0/20]	50.00% [n = 10/20]	0.00% [n = 0/20]
Class 2 – Intermediate occupations	0.00% [n = 0/18]	5.56% [n = 1/18]	0.00% [n = 0/18]	0.00% [n = 0/18]	33.33% [n = 6/18]	0.00% [n = 0/18]
Class 3 – Routine and manual occupations	12.50% [n = 2/16]	12.50% [n = 2/16]	0.00% [n = 0/16]	6.25% [n = 1/16]	50.00% [n = 8/16]	0.00% [n = 0/16]
Not ranked	0.00% [n = 0/6]	0.00% [n = 0/6]	0.00% [n = 0/6]	0.00% [n = 0/6]	0.00% [n = 0/6]	0.00% [n = 0/6]

There is both under- and over-reporting of *I likes toffees*. In class 1, 5.00% of informants under-report the form, but none over-report it. In class 2, however, the opposite is true; 5.56% of informants over-report the form, but none under-report it. In class 3, equal proportions of informants under- and over-report the form. Overall, however, the number of under- and over-reports of *I likes toffees* is small, and differences in under- and over-reporting of *I likes toffees* is not statistically significant. The chi-squared value for under-reporting of this form according to social class is 0.854, and for over-reporting it is 1.065, neither of which exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

There are no under-reports of *they likes toffees*. The sole over-report is found in class 3, but as there is only one report, no conclusions can be drawn from this.

There is no over-reporting of *she likes toffees*. Half of all informants in classes 1 and 3 under-report their use of the standard form, but this drops to a third for the informants in class 2. The fact that class 2 informants under-report their use of *she likes toffees* less than their class 1 counterparts might be explained by Labov's (1972b: 124-138) theory of lower middle class hypercorrection. Again, however, these differences in under-reporting of *she likes toffees* between social classes are not statistically significant, as the chi-squared value of 1.350 does not exceed the upper 5% point of chi-squared distribution with 2 degrees of freedom (5.991).

Overall, a comparison of speakers' reported and actual use of standard and nonstandard present tense verb morphology is informative on both a methodological and a linguistic level. From a methodological perspective, the difficulties of trying to collect grammatical data in free conversation in an interview conducted in a relatively limited time-span, discussed in section 3.5.2, are highlighted. Even in the case of a high frequency variable like present tense verb morphology, one cannot guarantee that speakers will produce the construction under investigation. This is apparent in the case of the third person singular present tense form where, as discussed, the claims of four speakers have had to be discounted as the speakers did not produce the form, standard or otherwise, in the interview, rendering an assessment of the accuracy of their claimed use impossible. This, coupled with the largely accurate speaker self-reporting of the four nonstandard present tense verb forms investigated, supports the use of a grammar questionnaire which can be used to collect data on a large number of variables in a short period of time.

However, the shortcomings of such questionnaires must also be acknowledged. As previously discussed, it might well be the case that some of those informants who under-reported their use of the standard form *she likes toffees* did so because they failed to read the questionnaire sufficiently thoroughly. In light of this potential for methodological flaws, and given previous findings which show that speakers do not necessarily accurately report their own language use (Labov 1966: 455; Trudgill 1972), it is wise, though not always practicable for reasons discussed in 3.5.2, to undertake a comparison of actual and reported use. That said, even inaccurate speaker self-reporting is of great interest for what it reveals to us about how different informants consider a variable.

Nonstandard *-s* in the present tense verb form is found in the Southampton area, albeit only in the speech of three out of the sixty informants. It was in the first person singular and the third person plural forms that nonstandard *-s* in present tense verb forms was recorded by the SED in Hampshire (Orton and Wakelin 1967-68: 445, 679), and this study shows that the nonstandard first person singular form is still present. Of course, it is impossible to say categorically whether or not use of nonstandard *-s* is as limited as these findings suggest. Had the SuRE interviews been longer and more tokens been elicited, perhaps there would have been more instances of nonstandard *-s*, both in the first person singular and, perhaps, in the other forms investigated. It might not be a coincidence that the tokens of nonstandard *-s* elicited are found in the first person singular, as it is this form for which, overall, most tokens were collected. Similarly, despite the intention that the interview should be as informal as possible, and the questioning method was designed to achieve this (see section 3.7), the very fact that it was an interview situation may have influenced speakers' use of the form (Labov 1972b: 209; Milroy 1987b: 41).

5.6.2 Adverbs

Table 5.10 shows speakers' reported use of sentences in the Southampton area language questionnaire, illustrating both standard and nonstandard *-ly* adverbs.

Table 5.10 Reported use of *-ly* adverbs

Variable	Number of informants claiming to use this form when talking to a friend
<i>He writes really quick</i>	23/60
<i>She laughs really loudly</i>	46/60

Table 5.11 shows the reported use of both the standard and the nonstandard form by speaker age and sex.

Table 5.11 Reported use of standard and nonstandard adverbs according to the speaker variables of age and sex

	<i>He writes really quick</i>		<i>she laughs really loudly</i>	
	Male	Female	Male	Female
<30	5/10	6/10	9/10	10/10
30-59	6/10	1/10	6/10	7/10
>59	4/10	1/10	8/10	6/10

Reported use of *he writes really quick* shows marked sex variation, with nearly twice as many male informants as female claiming to use this variable. It should be noted, however, that this difference in reported use between sexes is not statistically significant, as the chi-squared value of 2.538 does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). There is no such sex differentiation in the reported use of *she laughs really loudly*. It is those speakers in the under 30 age group who most frequently report both *he writes really quick* and *she laughs really loudly*. The chi-squared value for *he writes really quick* according to speaker age is 3.948, and for *she laughs really loudly* it is 3.936.

Neither value is statistically significant, as neither exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Table 5.12 shows the reported use of both the standard and the nonstandard form by speaker social class.

Table 5.12 Reported use of standard and nonstandard adverbs according to the speaker variable of social class

	<i>He writes really quick</i>	<i>She laughs really loudly</i>
Class 1 – Managerial and professional occupations	35.00% [n = 7/20]	90.00% [n = 18/20]
Class 2 – Intermediate occupations	38.89% [n = 7/18]	72.22% [n = 13/18]
Class 3 – Routine and manual occupations	43.75% [n = 7/16]	56.25% [n = 9/16]
Not ranked	33.33% [n = 2/6]	100.00% [n = 6/6]

In the case of *he writes really quick*, reported use of this nonstandard form increases as the social class of speakers becomes lower. The differences in reported use according to social class are not statistically significant, however, as the chi-squared value of 0.286 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). The reverse is true of *she laughs really loudly*. The higher the social class of the informants, the higher the level is of reported use of the standard form. As in the case of *he writes really quick*, however, differences in reported use of *she laughs really loudly* according to social class are not statistically significant (the chi-squared value is 3.685).

Table 5.13 shows actual use of standard and nonstandard adverbial forms during the SuRE interviews.

Table 5.13 Actual use of standard and nonstandard *-ly* adverbs

Tokens of standard adverbs	Tokens of nonstandard adverbs
1951	8

Table 5.13 shows that the standard *-ly* form is overwhelmingly the preferred construction for the Southampton area informants. While this is surprising in some respects, given that Edwards et al. (1984: 24) claim that use of adverbs without the *-ly* suffix are more common in most dialects than is the standard form, it mirrors Tagliamonte and Ito's (2002: 248) findings in York which showed the standard form to be the dominant one. It should be remembered that the tokens in the present survey have been collected in an interview situation, albeit one that aimed to be as informal as possible. Speakers are likely to have been monitoring their speech more than they would in everyday conversation with friends, for example. There is no way of saying definitively that those speakers who did not employ the nonstandard form in the interview do not do so outside the interview context. However, one can conclude that most speakers favoured the standard *-ly* form in an interview. Even the six informants who did employ nonstandard adverbs did so variably, with the majority of their tokens being standard. Table 5.14 shows the accuracy of speaker self-reporting of adverbial construction.

Table 5.14 Speaker self-reporting of adverbial constructions

Variable	Accurate claim	Over-report	Under-report	Discounted
<i>He writes really quick</i>	34/60	21/60	4/60	1/60
<i>She laughs really loudly</i>	46/60	0/60	12/60	2/60

The majority of claims to use or not to use *he writes really quick* and *she laughs really loudly* are accurate ones. The tendency amongst the remaining informants was

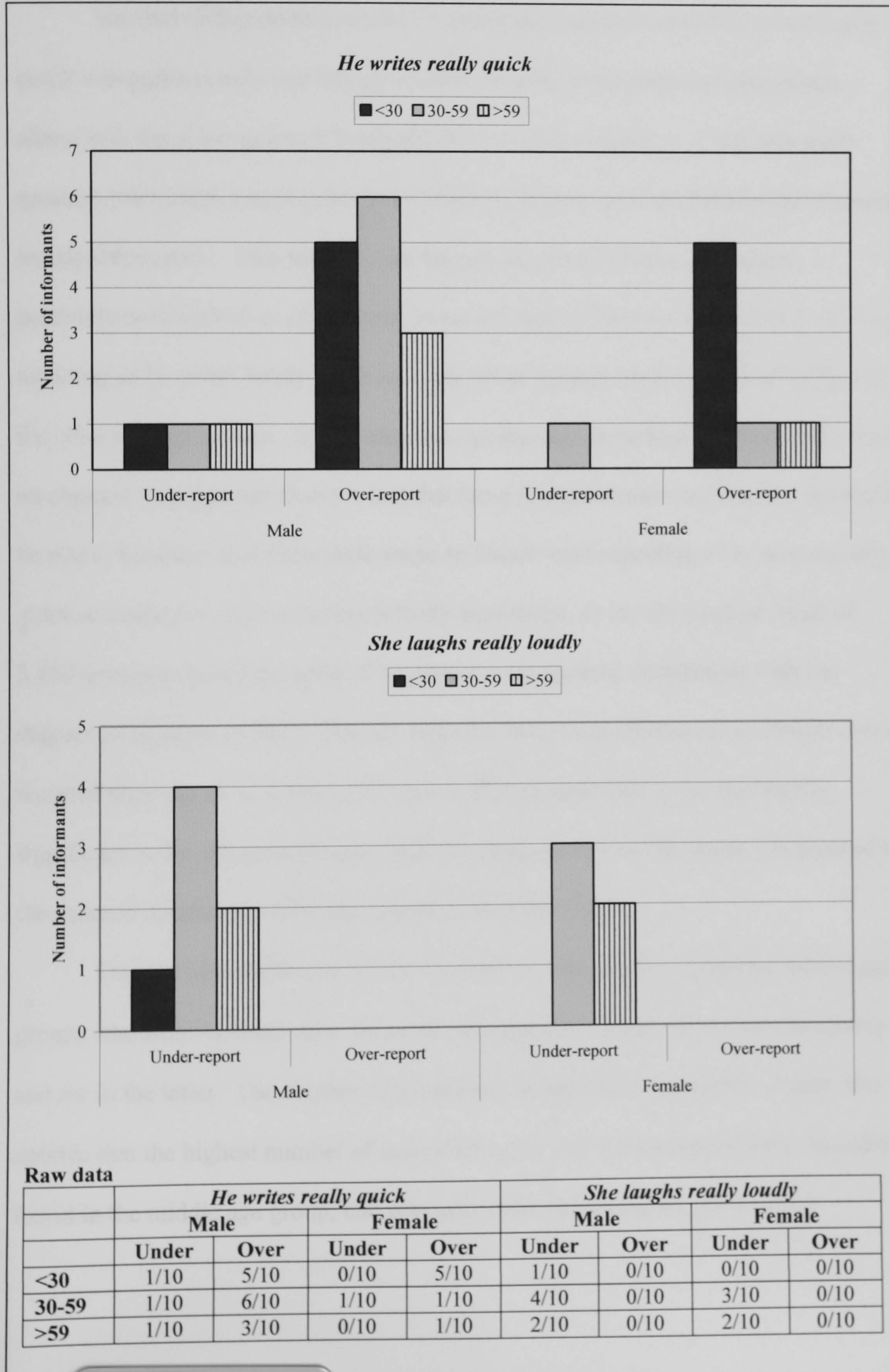
to claim to use nonstandard forms when, from the data available, they actually use the standard forms. In the case of *he writes really quick*, twenty-one speakers over-reported their use of the nonstandard form, compared to only four speakers who did not claim to use the nonstandard form but actually did. Of course it is not possible to say categorically that speakers do not use a particular construction; conclusions are drawn from the data available. In the case of the standard *-ly* adverb, twelve informants did not claim to use the standard form when in fact they do. The three discounted claims are double responses, which make it impossible to ascertain the accuracy of the speakers' reported use because they have both claimed to use and claimed not to use the forms under investigation.

Figure 5.14 shows the under- and over-reporting of *he writes really quick* and *she laughs really loudly* according to the speaker variables of age and sex.

There is no over-reporting of *she laughs really loudly*. Notably, given that the middle age group has been shown to be the one most influenced by the standard (Chambers and Trudgill 1998: 79), the highest number of under-reports is found in this group. Seven speakers aged between 30 and 59 under-reported their use of *she laughs really loudly*, compared with four speakers aged 60 and over, and one aged under 30. These differences in under-reporting between age groups are not statistically significant, however. The chi-squared value for under-reporting of *she laughs really loudly* according to speaker age is 3.984, which does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). There is little difference in under-reporting between the sexes, with seven men and five women claiming not to use *she laughs really loudly* but employing it in the interview. Again, this difference is not statistically significant, as the chi-squared

value of 0.104 does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841).

Figure 5.14 Under- and over-reporting of standard and nonstandard adverbial constructions according to the speaker variables of age and sex



In the case of *he writes really quick*, there are more over-reports, claims to use the nonstandard form which are not borne out by the interview data, than under-reports. Of course it is impossible to say categorically that those speakers who over-reported their use of this form never employ it.

The fact of five females under 30 over-reporting their use of *he writes really quick*, compared to only one female speaker in each of the other two age groups, aligns with the observation of Trudgill (1972: 192) that '[i]t is [...] not only male speakers who attach covert prestige to W[orking]C[lass] speech, but also the younger female informants'. This tendency on the part of young females to evaluate positively nonstandard or stigmatised features might explain the higher level of over-reporting of *he writes really quick* amongst those women under the age of 30 than in the other two age groups. In this instance, speaker age is perhaps a greater influence on claimed language use than the fact that these five informants are female. It should be noted, however, that these differences in female over-reporting of *he writes really quick* according to age are not statistically significant, as the chi-squared value of 3.400 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). Overall, twice as many male informants as female over-reported their use of *he writes really quick*, though again this is not statistically significant as the chi-squared value of 2.637 does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841).

There is little difference in the numbers of males in the young and middle age groups who over-reported their use of the nonstandard *-ly* adverb, five in the former and six in the latter. The number drops to three in the oldest age group. Again, it is notable that the highest number of male over-reports of a nonstandard form should be found in the middle age group, one normally most influenced by the standard

(Chambers and Trudgill 1998: 79). However, the differences in male over-reporting of *he writes really quick* according to age are not statistically significant. The chi-squared value of 0.837 does not exceed the upper 5% point of a chi-squared distribution with 2 degrees of freedom (5.991).

That more male informants than female under-reported their use of *he writes really quick* is perhaps surprising, but the overall number of under-reports, four in total, is too small to draw any meaningful conclusions. The difference in under-reporting of *he writes really quick* according to speaker sex is not statistically significant, as the chi-squared value of 0.268 does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841).

Table 5.15 shows the under- and over-reporting of *he writes really quick* and *she laughs really loudly* according to the speaker variable of social class.

Table 5.15 Under- and over-reporting of standard and nonstandard adverbial constructions according to the speaker variable of social class

	<i>He writes really quick</i>		<i>She laughs really loudly</i>	
	Under	Over	Under	Over
Class 1 – Managerial and professional occupations	5.00% [n = 1/20]	30.00% [n = 6/20]	10.00% [n = 2/20]	0.00% [n = 0/20]
Class 2 – Intermediate occupations	5.56% [n = 1/18]	33.33% [n = 6/18]	22.22% [n = 4/18]	0.00% [n = 0/18]
Class 3 – Routine and manual occupations	12.50% [n = 2/16]	43.75% [n = 7/16]	37.50% [n = 6/16]	0.00% [n = 0/16]
Not ranked	0.00% [n = 0/6]	33.33% [n = 2/6]	0.00% [n = 0/6]	0.00% [n = 0/6]

Under-reporting of the nonstandard *he writes really quick* is fairly consistent across classes 1 and 2, and increases in class 3. The small differences in under-reporting between classes are not statistically significant as the chi-squared value of 0.113 does not exceed the upper 5% point of a chi-squared distribution with two

degrees of freedom (5.991). Over-reporting of the zero adverb form also increases as social class becomes lower. Again, however, the small differences in over-reporting of *he writes really quick* between classes are not statistically significant, as the chi-squared value of 0.778 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). There is no over-reporting of *she laughs really loudly*. Under-reporting of this standard form, however, increases as the social class of the speaker falls. However, with a chi-squared value of 2.541, these differences in under-reporting of *she laughs really loudly* between social classes are not statistically significant.

Zero forms of *-ly* adverbs, as in the example *he writes really quick*, are present in the Southampton area, albeit only in the speech of six out of a possible sixty informants. As in the case of the present tense verb morphology, had the interviews been longer, more tokens of this feature might have been elicited.

Adverbs with the *-ly* suffix are, however, dominant.

5.7 Grammar – A Summary

Overall, reported use of nonstandard grammatical features in the Southampton area study is limited. Despite this, however, some of the features analysed in this chapter do correlate with the speaker variables of sex, age and social class. *I ain't got a clue* patterns significantly with speaker sex, being more likely to be reported by those under 60, than by those and 60 and over. The nonstandard simple prepositions *I'm going down my friend's house later* and *I'm going round my friend's house later* correlate with speaker age, being used most frequently by those under 30. Male informants are more likely to use *I'm going up my friend's house later* than are their female counterparts. It is, however, significant that *there was* with a plural notional

subject does not vary significantly according to sex or social class, as this supports Cheshire's (1999: 70) assertion that *there was* shows signs of being a lexical unit which is not subject to sharp social differentiation.

An examination of actual use shows that nonstandard *-s* in the present tense verb morphology is present in the speech of the Southampton area, as are adverbs with an adjectival form. Only very limited tokens of both, however, were found in the SuRE recordings, and the standard forms appear to be the dominant ones in the Southampton area.

Overall, a comparison of reported and actual use of present tense verb morphology and *-ly* adverbs has proved to be informative. In the case of some variables, namely *you likes toffees*, *we likes toffees*, *they likes toffees*, and even *I likes toffees*, accuracy of reported use is very high. However, in the case of other variables, namely *he writes really quick* and *she likes toffees*, only slightly over half of all informants accurately reported their own use. This variable accuracy raises important methodological issues, namely how best to collect grammatical data, and the level of formality induced by the SuRE technique. The high level of accuracy of reported use for many of the sentences illustrating present tense verb morphology supports the use of a grammar questionnaire as an effective means to collect large quantities of comparable data in a short space of time.

Though the relatively low number of accurate claims for some of the other variables investigated suggests that we might gain a false impression of language use in the Southampton area if we rely on this data alone, in other respects it is very telling. Disparity between reported and actual use can provide information regarding the way in which informants' perceive particular variables, and sometimes highlights variability in these attitudes between different speaker groups. Even when reported

and actual use are compared, it is impossible say definitively that speakers do not use a particular form which they might have reported but which is not found in data collected during the interview.

Aside from methodological issues, a comparison of reported and actual use of standard and nonstandard *-ly* adverbs suggests that there is a tendency among informants to over-report their use of nonstandard variables and, conversely, to under-report their use of standard variables. This tendency, however, is greatest amongst male informants, supporting previous findings which state that males tend to find covert prestige in nonstandard or low-prestige forms (Trudgill 1972: 188), and amongst those from the lowest social class.

Chapter 6

6.0 Phonology

In this chapter, L Vocalization, rhoticity, and the BATH and PRICE vowels are analysed according to the speaker variables of sex, age and social class. As discussed in section 3.8, the Southampton area study has 30 male and 30 female informants, and 20 speakers in each of the three age groups (under 30, 30 to 59, and 60 and over). The speaker sample is comprised of 3 social class groups, class 1 being the highest, and class 3 the lowest. Class 1 has 20 informants, class 2 has 18 informants, and class 3 has 16 informants (see section 3.9 for a discussion of the social stratification of Southampton area speakers).

As stated in 3.10.3, in all but the case of the BATH vowel, 15 tokens of the phonological features to be analysed by extralinguistic variables have been collected per speaker. For the BATH vowel, as close to 15 tokens as possible per speaker have been collected. In order to maximise the number of tokens of this variable collected, tokens of the START vowel (in non-rhotic speakers) and of the PALM vowel have been included with BATH, since in Received Pronunciation, these share the same stressed vowel, /ɑ:/ (Wells 1982: 143). Eighteen of the sixty informants for the Southampton area study produced fewer than 15 tokens, though of these eighteen, all but six produced 10 or more tokens, the number for which Milroy (1987b: 135) states that researchers should aim when collecting phonological data. In total, 806 tokens of BATH have been collected for analysis, whereas for the other phonological variables, 900 tokens have been collected. The BATH vowel is very worthy of analysis, being a particularly significant phonological variable in the Southampton area, and over 800 tokens of this vowel are more than sufficient to enable reliable conclusions to be drawn regarding its use by different speaker groups.

Chi-squared testing (see section 3.12) has been undertaken when differences in use of a variant between speaker groups are small.

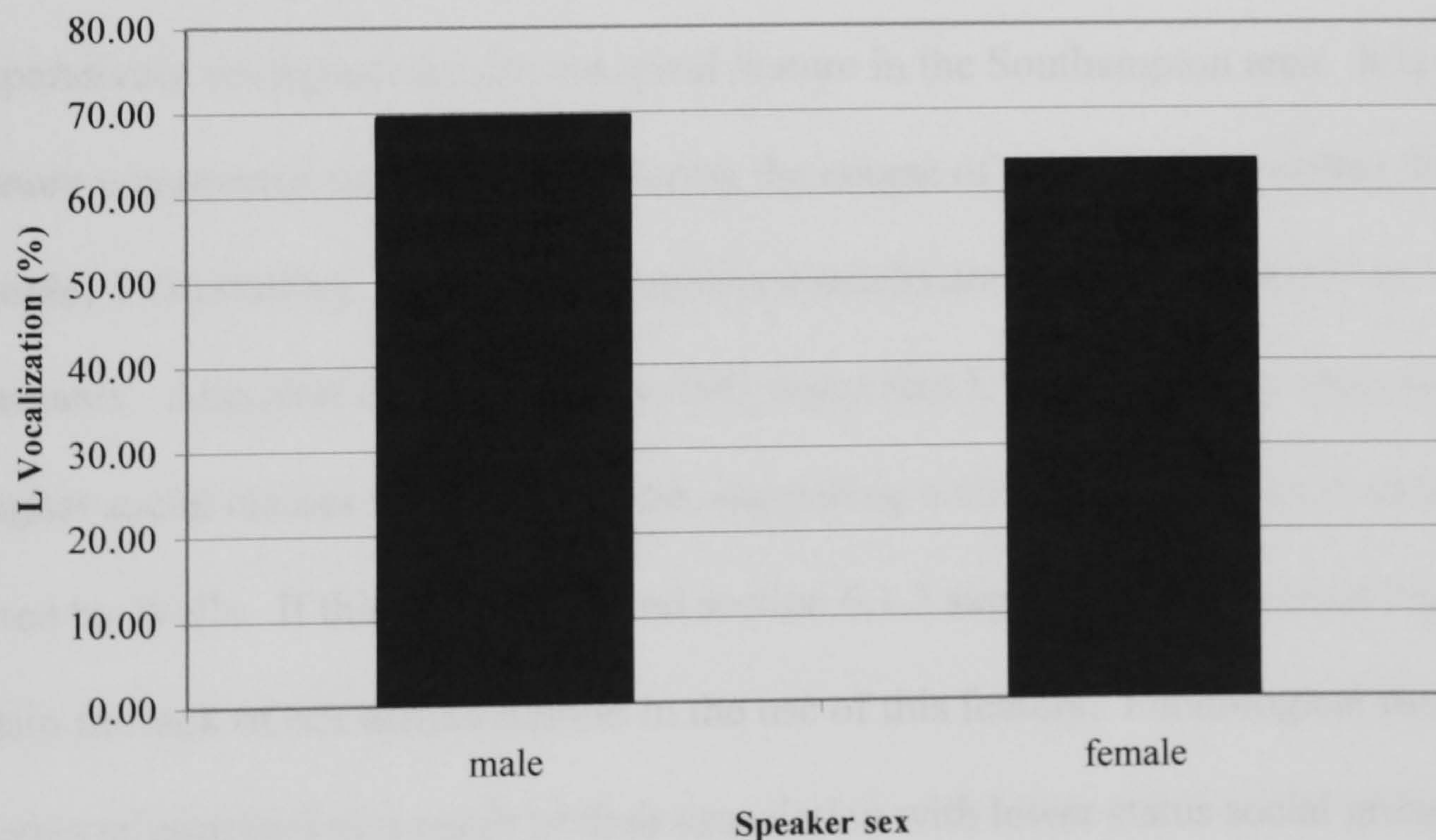
6.1 L Vocalization

Data have been collected on L Vocalization in the Southampton area in word-final, word-final intervocalic, word-medial, and preconsonantal positions. In the following sections, L Vocalization is analysed according to the extralinguistic variables of sex, age and social class.

6.1.1 L Vocalization and Speaker Sex

Figure 6.1 shows L Vocalization according to the extralinguistic variable of speaker sex.

Figure 6.1 L Vocalization according to the speaker variable of sex



Raw data

speaker sex	L Vocalization
male	69.56% [n = 313/450]
female	63.78% [n = 287/450]

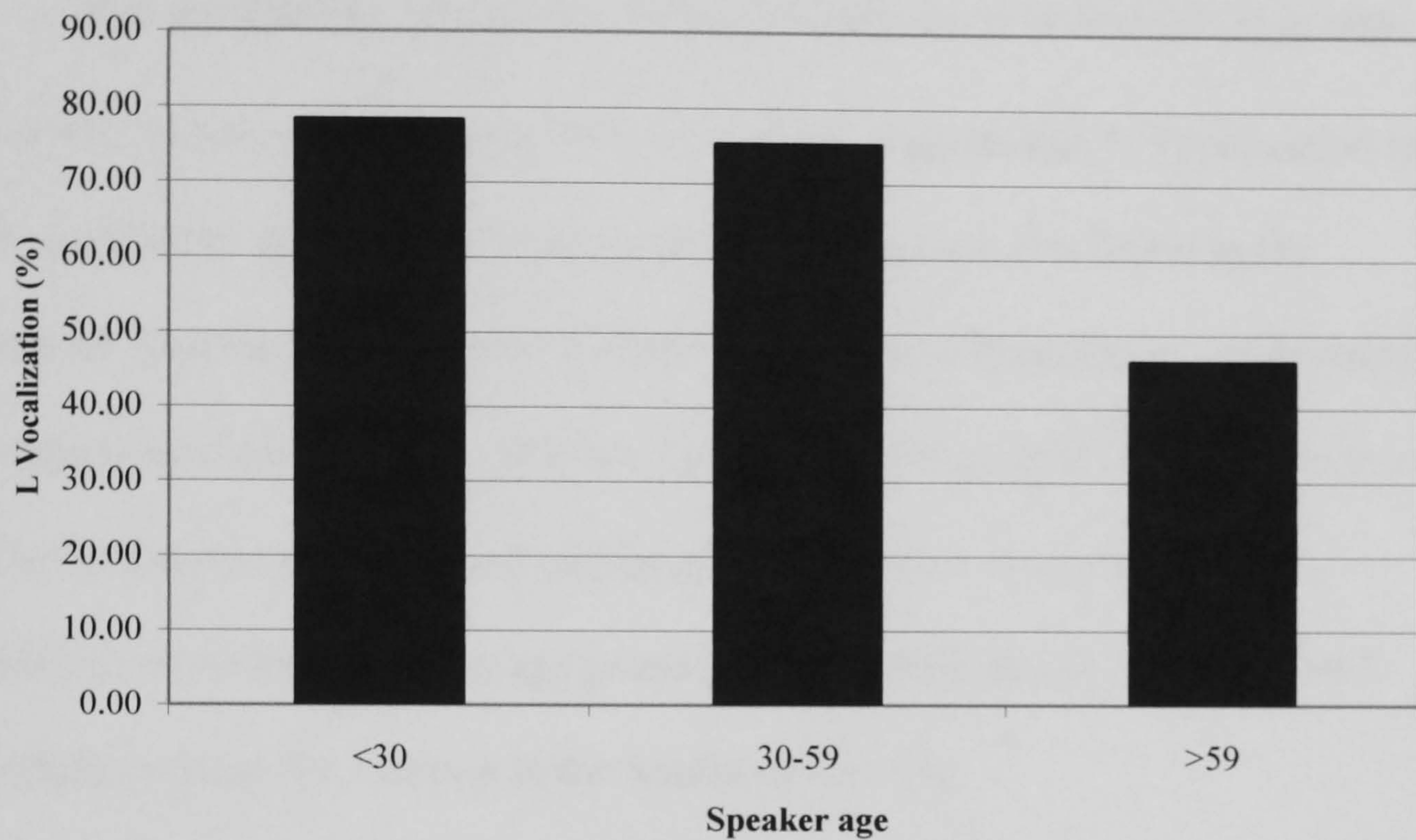
It is apparent from Figure 6.1 that the use of a vocoid for [ɫ] is slightly higher among male speakers than it is among their female counterparts. 69.56% of tokens of [ɫ] collected from the former speaker group are realised as a vocoid, compared to 63.78% of tokens from the latter group. This difference in L Vocalization between male and female informants is not, however, statistically significant, as the chi-squared value of 3.125 does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841).

Since Wells (1982: 20) argues that L Vocalization is a feature 'to which low prestige is attached', one might have expected use of this variant to have been significantly higher among male informants than among their female counterparts, as men have traditionally been found to employ more stigmatised or low prestige variants than women (Chambers 2003: 116). However, given the high levels of vocalization among both speaker groups (approximately two-thirds of all tokens of [ɫ] are realised as a vocoid by both groups) it would appear that L Vocalization is a comparatively unstigmatised phonological feature in the Southampton area. It is not a feature commented on by speakers during the course of the interview, unlike, for example, T Glottalling, about which negative remarks are made by a number of informants. Altendorf and Watt (2004: 196) argue that L Vocalization is spreading to higher social classes in the South East, suggesting a change in its status from that offered by Wells. If this is the case (and section 6.1.3 suggests it is), it would help to explain the lack of sex differentiation in the use of this feature. Phonological forms are often stigmatised as a result of their association with lower status social groups. If L Vocalization is used by people across the social spectrum, there would be no reason for this variant to be avoided by female speakers on the grounds that it lacks prestige.

6.1.2 L Vocalization and Speaker Age

Figure 6.2 shows L Vocalization according to speaker age.

Figure 6.2 L Vocalization according to the speaker variable of age



Raw data

speaker age	L Vocalization
<30	78.33% [n = 235/300]
30-59	75.33% [n = 226/300]
>59	46.33% [n = 139/300]

The use of a vocoid for [ɫ] in non-prevocalic positions is highest among those speakers under the age of 30. 78.33% of tokens of [ɫ] collected from this speaker group are realised as a vocoid, compared to 75.33% of tokens collected from those aged 30 to 59, and 46.33% of tokens from speakers aged 60 and over. This difference in use between age groups is highly statistically significant. The chi-squared value of 84.330 far exceeds both the upper 5% point of a chi-squared

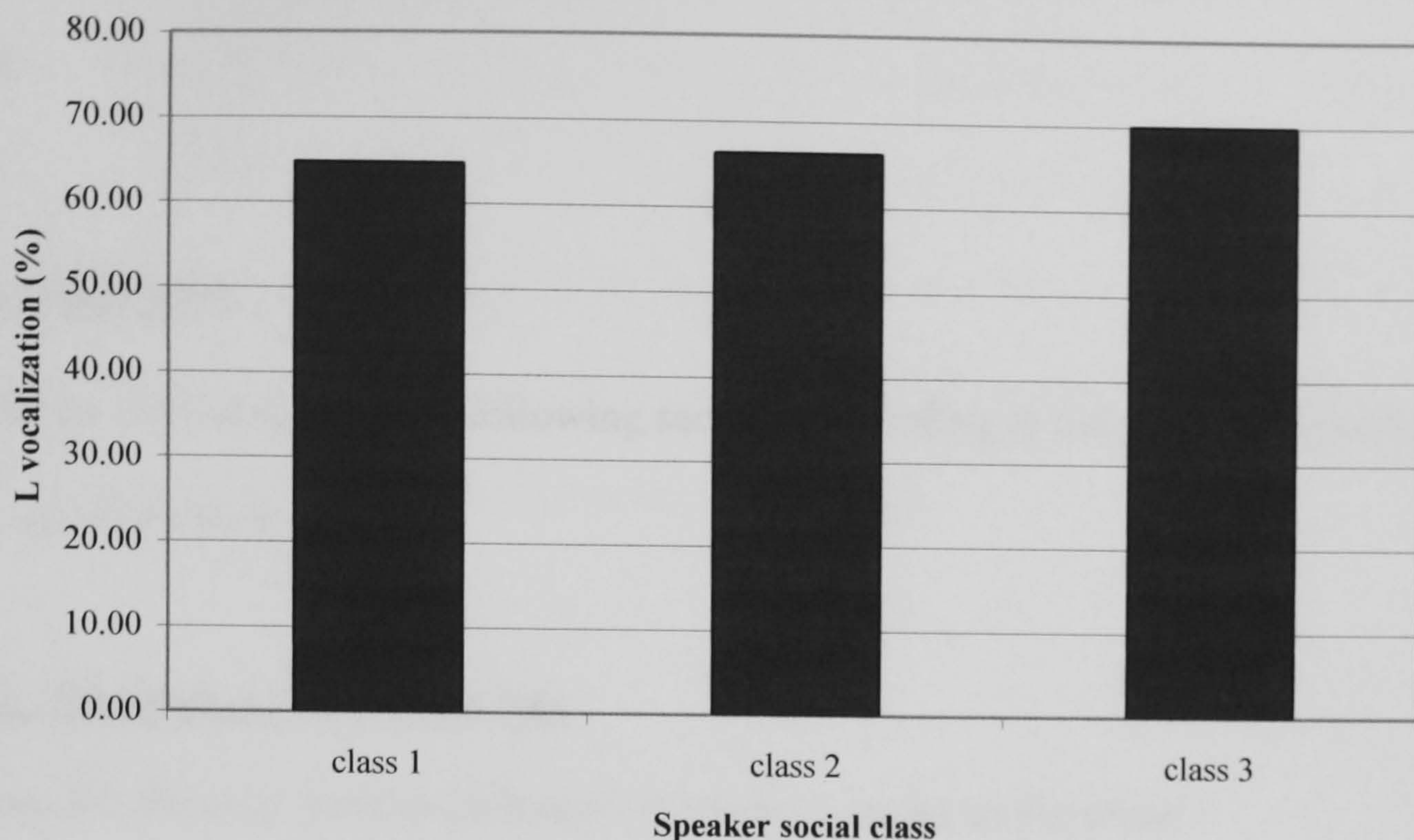
distribution with two degrees of freedom (5.991), and indeed the upper 0.1% point (13.82), making it significant at the 99.9% level. There is a very strong correlation between L Vocalization and speaker age, with use of a vocoid for [ɫ] being more common among those under 60 than among those aged 60 and over.

That even the age group who least frequently use a vocoid for [ɫ] in non-prevocalic positions do so nearly 50% of the time suggests that L Vocalization is a well-established variant in the Southampton area. Indeed, it is found in the Incidental Material of the Survey of English Dialects in Hampshire (Leeds Archive of Vernacular Culture 2005). The fact that the percentage of L Vocalization is so similar among the youngest and middle age groups, and the percentage of L Vocalization among the oldest age group is considerably lower, suggests that L Vocalization is on the increase in the Southampton area.

It is of interest to note that, while Tollfree (1999: 174), in her study of South East London English, found vocalization in word-final intervocalic position only in the speech of informants aged between 15 and 30, in the present study, vocalization was found in this phonetic environment in the speech of a 56-year-old male from Eastleigh Borough. Analysis of a greater number of tokens from the existing speaker sample, or a larger speaker sample, would be needed to ascertain whether this is an anomalous finding or one which shows that L Vocalization is possible in word-final intervocalic position in the Southampton area, and, if so, how the feature patterns with speaker age.

6.1.3 L Vocalization and Speaker Social Class

L Vocalization according to speaker social class is shown in Figure 6.3.

Figure 6.3 L Vocalization according to the speaker variable of social class**Raw data**

speaker social class	L Vocalization
class 1	64.67% [n = 194/300]
class 2	66.30% [n = 179/270]
class 3	70.00% [n = 168/240]
not ranked	65.56% [n = 59/90]

It is apparent from Figure 6.3 that L Vocalization, a feature traditionally associated with working class speech (Wells 1982: 314), is most frequently found in the speech of those in class 3, and that use decreases the higher the speakers' social class. However, the difference in use between the three groups is small, and overall use is high regardless of social class, with L Vocalization accounting for over 60% of all realisations of /t/. The differences in L Vocalization between social classes is not statistically significant, as the chi-squared value of 1.754 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). The

findings of the present study appear to support Altendorf and Watt's (2004:196) assertion that L Vocalization 'is spreading [...] to higher social classes' in the South East.

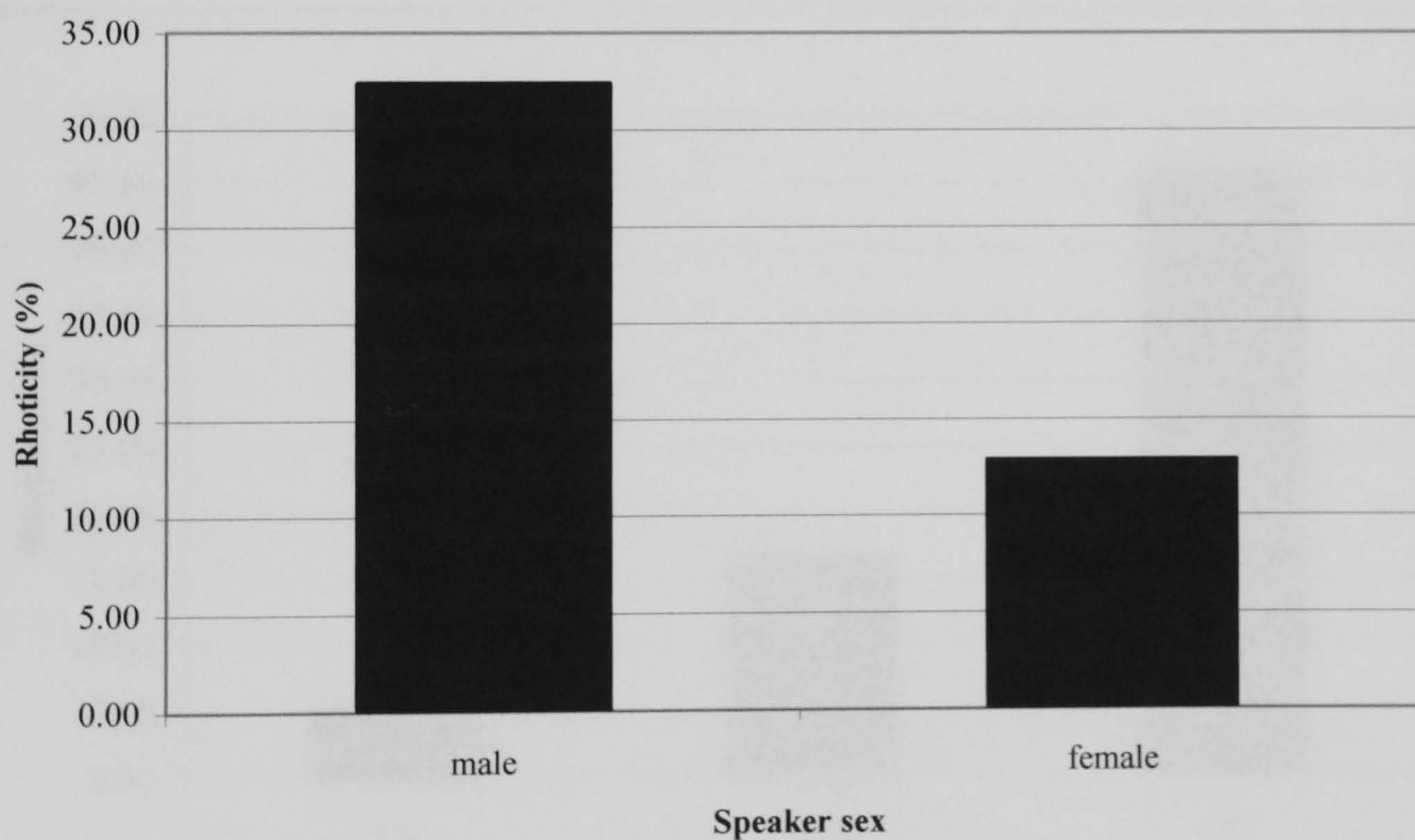
6.2 Rhoticity

Rhoticity is examined in the following sections according to the speaker variables of sex, age and social class.

6.2.1 Rhoticity and Speaker Sex

Figure 6.4 shows a marked difference in rhoticity between the sexes.

Figure 6.4 Rhoticity according to the speaker variable of sex



Raw data

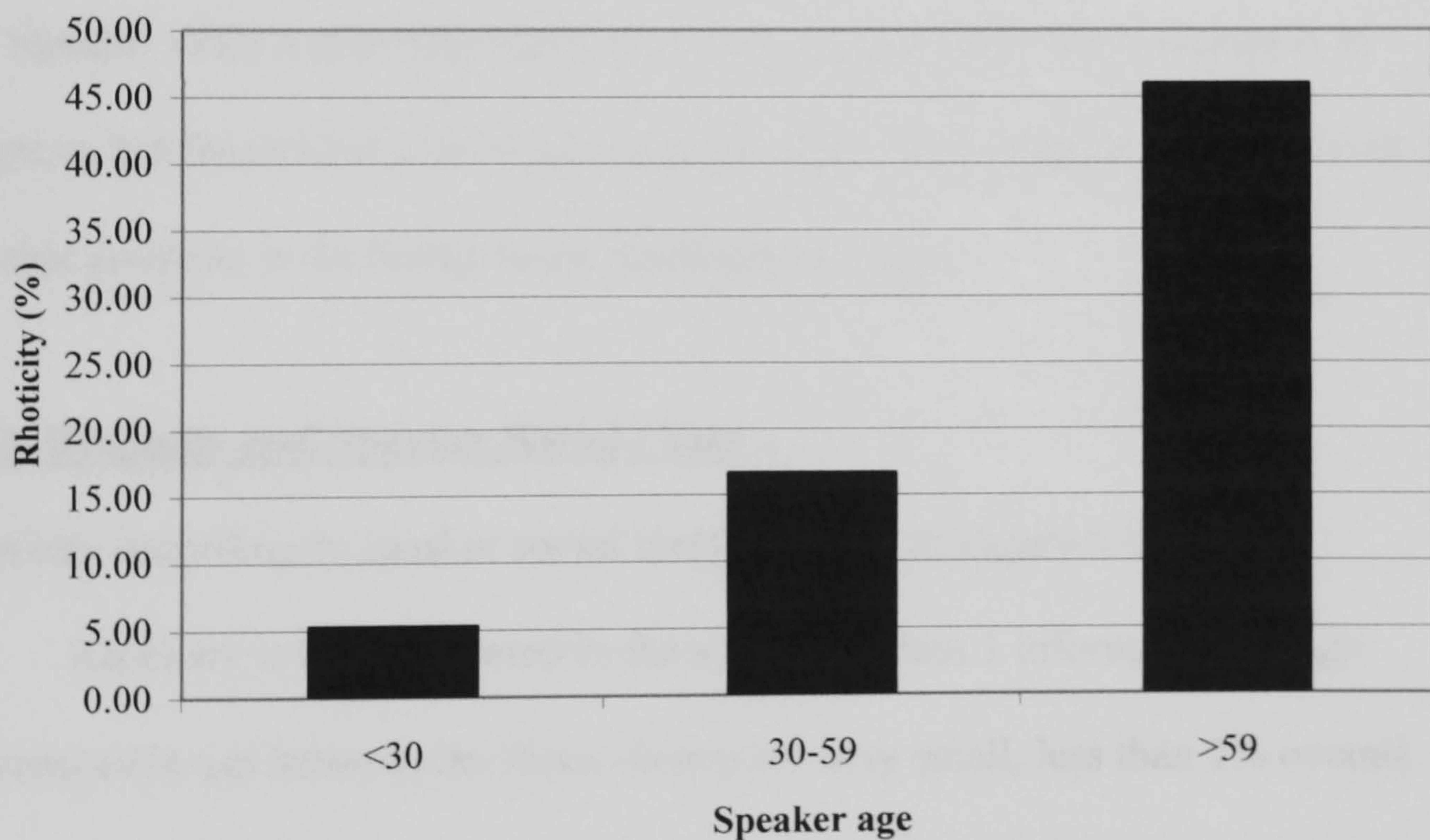
speaker sex	rhoticity
male	32.44% [n = 146/450]
female	12.89% [n = 58/450]

32.44% of tokens of (r) in preconsonantal and absolute-final environments collected from male speakers are realised as /r/, compared with only 12.89% of tokens offered by female informants. This difference in rhoticity between sexes is highly statistically significant. The chi-squared value of 47.978 far exceeds the upper 5% point of a chi-squared distribution with one degree of freedom (3.841), and indeed the upper 0.1% point (10.83), making it significant at the 99.9% level. Men in the Southampton area are far more likely to be rhotic than their female counterparts.

6.2.2 Rhoticity and Speaker Age

Figure 6.5 shows rhoticity according to speaker age.

Figure 6.5 Rhoticity according to the speaker variable of age



Raw data

speaker age	rhoticity
<30	5.33% [n = 16/300]
30-59	16.67% [n = 50/300]
>59	46.00% [n = 138/300]

It is apparent from Figure 6.5 that, like L Vocalization, rhoticity in the Southampton area patterns correlates with speaker age. Unlike L Vocalization, however, it is the oldest group who most frequently employ this feature, with 46% of tokens of (r) in preconsonantal and absolute-final environments collected from speakers aged over 59 being realised as /r/. This figure decreases to 16.67% of tokens offered by speakers aged between 30 and 59, dropping again to 5.33% of tokens from those aged under 30. These differences in use between age groups are highly statistically significant. The chi-squared value of 150.761 far exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991), and indeed the upper 0.1% point (13.82), making it significant at the 99.9% level. In the Southampton area, as speaker age increases, so too does rhoticity.

It would appear likely that rhoticity is on the decline in the Southampton area. No informants are wholly rhotic, and very few speakers under the age of 30 employ this feature. Only a real-time study could say categorically that a change is in progress, but the evidence from this apparent time study suggests that this is the case and that rhoticity is declining in the Southampton area.

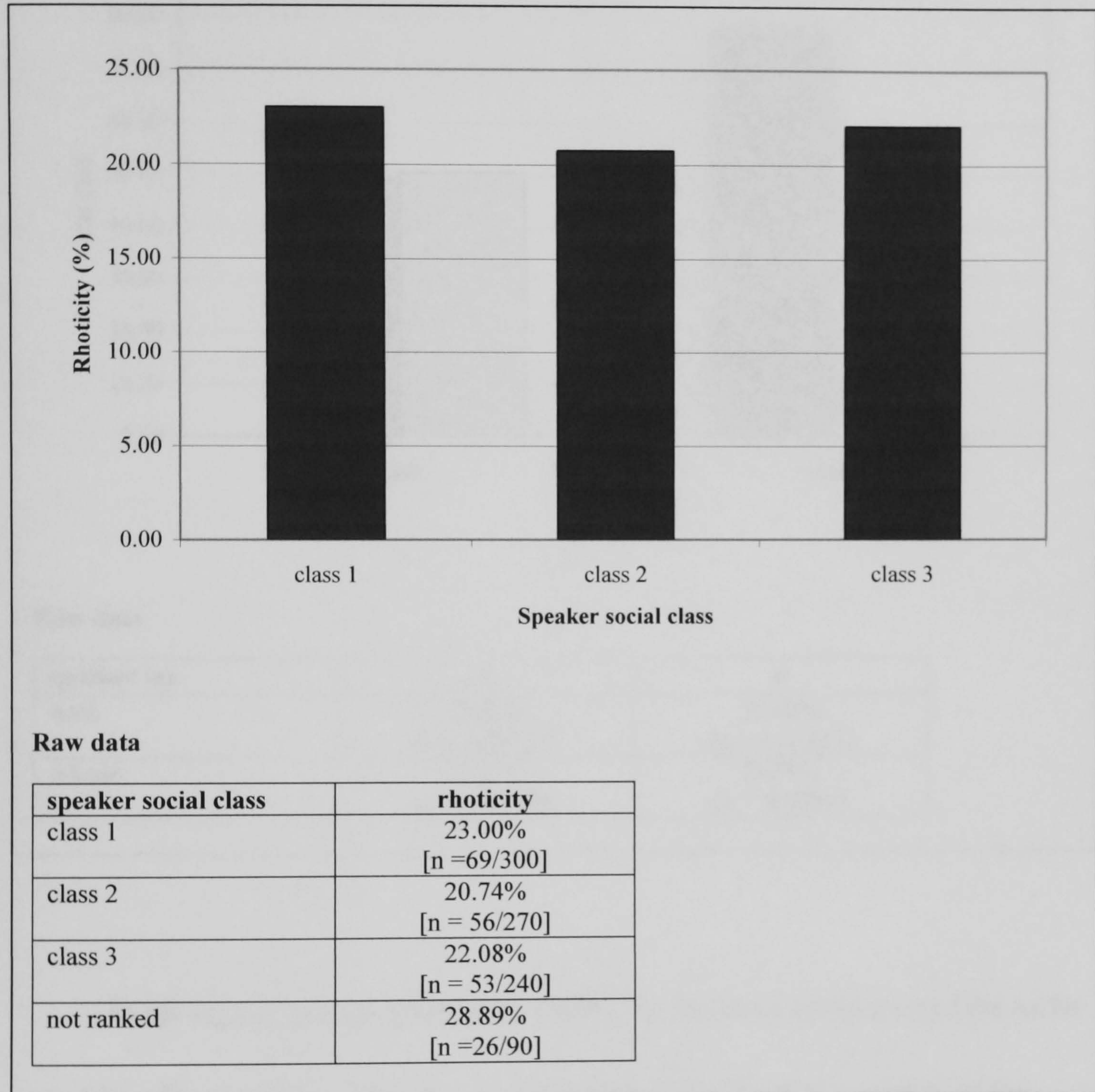
6.2.3 Rhoticity and Speaker Social Class

Rhoticity according to speaker social class is shown in Figure 6.6.

Rhoticity is most common in the speech of class 1 informants, though differences in use between the three classes are very small, less than 2% overall. These differences between social classes are not statistically significant, as the chi-squared value of 0.425 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). It does not appear that social class affects rhoticity in the Southampton area. Analysis in sections 6.2.1 and 6.2.2

suggests that the sex and age of the speaker are a greater influence on rhoticity than is their social class.

Figure 6.6 Rhoticity according to the speaker variable of social class



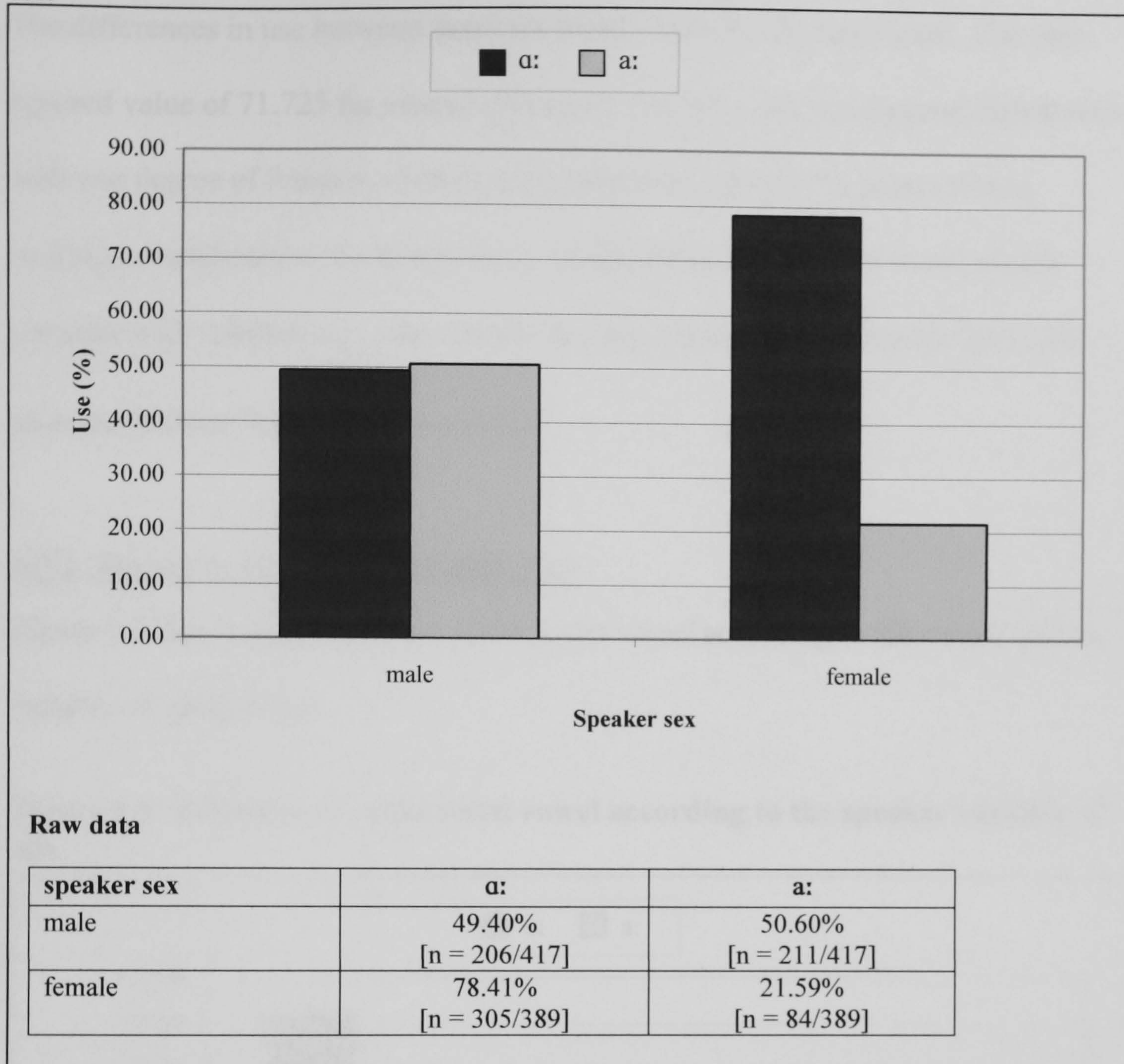
6.3 The BATH Vowel

In the following sections, the BATH vowel is analysed according to the speaker variables of sex, age and social class.

6.3.1 The BATH Vowel and Speaker Sex

Figure 6.7 shows realisations of the BATH vowel according to the sex of the speaker.

Figure 6.7 Realisations of the BATH vowel according to the speaker variable of sex



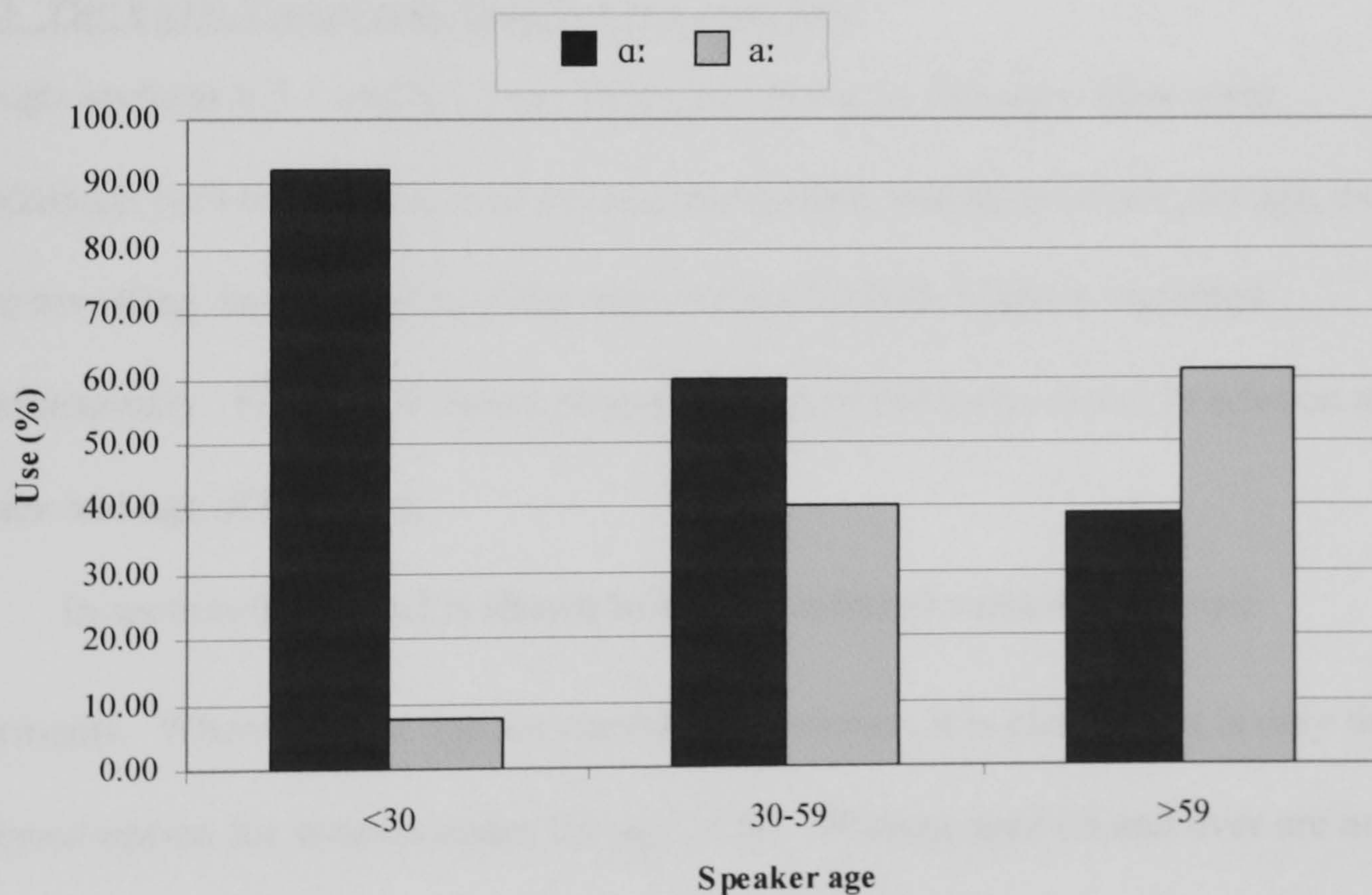
In the case of female informants, [ɑ:] is the preferred realisation of the BATH vowel by a large margin. The [a:] variant accounts for less than a quarter of the tokens collected from female speakers. By contrast, use of [ɑ:] and [a:] by male speakers is almost identical, with the latter form being the preferred option by little over 1%. Men are more than twice as likely as their female counterparts to employ the fronted variant. Given that females tend to employ more prestige forms than males (Chambers 2003: 116), and that [ɑ:] is the Received Pronunciation realisation of the BATH vowel (*OED*), it is unsurprising that this form is preferred by female

informants in the Southampton area, and that it is used by more females than males. The differences in use between sexes are highly statistically significant. The chi-squared value of 71.725 far exceeds the upper 5% point of a chi-squared distribution with one degree of freedom (3.841), and indeed the upper 0.1% point (10.83), making it significant at the 99.9% level. Realisations of the BATH vowel clearly correlate with speaker sex, with women far more likely than men to use [ɑ:], and men more likely than women to use [a:].

6.3.2 The BATH Vowel and Speaker Age

Figure 6.8 shows pronunciations of the BATH vowel according to the extralinguistic variable of speaker age.

Figure 6.8 Realisations of the BATH vowel according to the speaker variable of age



Raw data

speaker age	ɑ:	a:
<30	92.34% [n = 241/261]	7.66% [n = 20/261]
30-59	59.78% [n = 165/276]	40.22% [n = 111/276]
>59	39.03% [n = 105/269]	60.97% [n = 164/269]

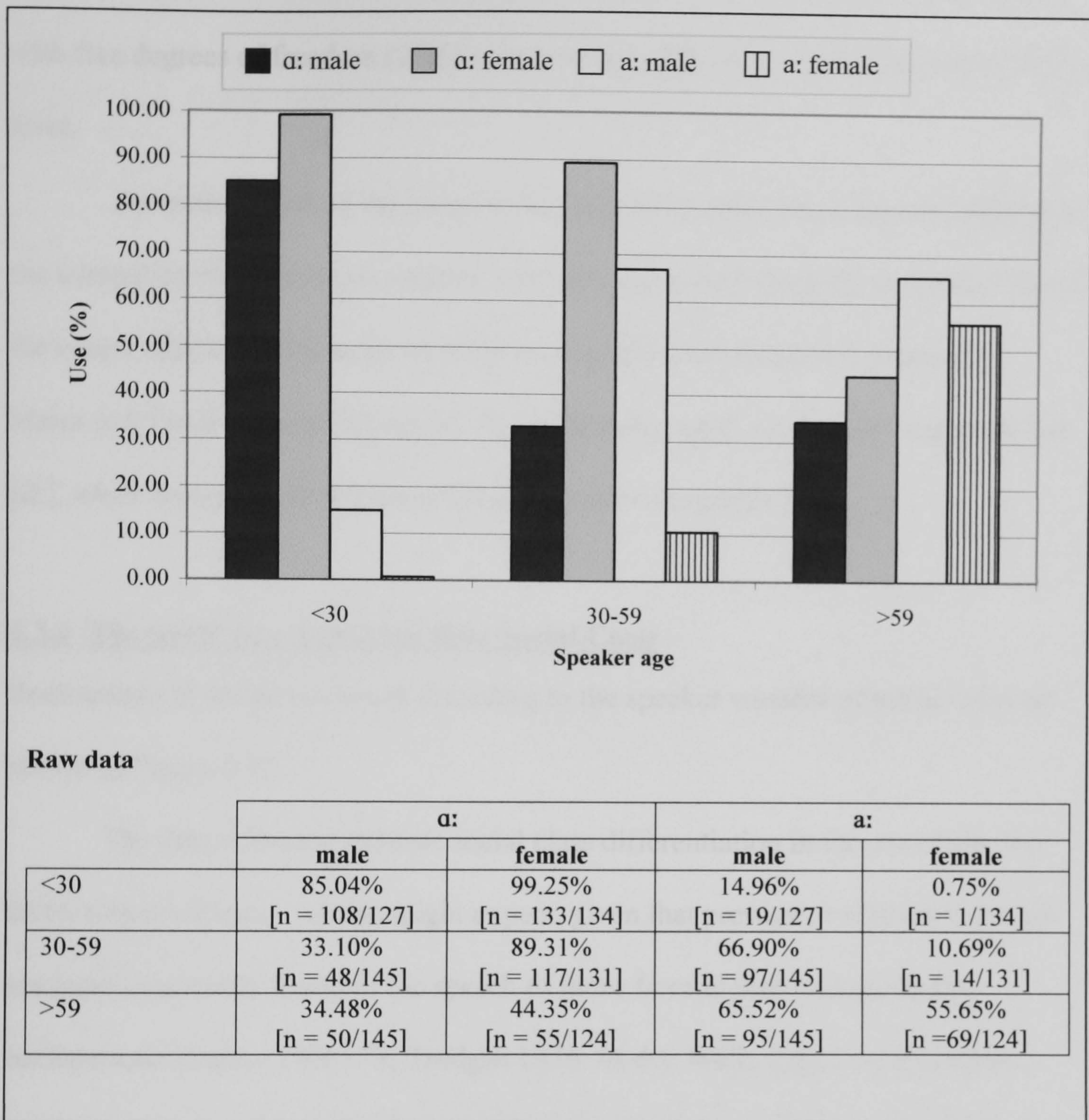
The data collected show a clear correlation between speaker age and the BATH vowel. Use of [ɑ:], associated with Received Pronunciation and with the South East (Altendorf and Watt 2004: 187; Hughes et al. 2005: 69; *OED*), decreases markedly as speaker age increases. Conversely, use of [a:], associated with the South West (Altendorf and Watt 2004: 199) increases with increasing speaker age. These differences in use between age groups are again highly statistically significant. The chi-squared value of 164.569 far exceeds the upper 0.1% point of a chi-squared distribution with two degrees of freedom (13.82), making the findings significant at the 99.9% level. Informants aged 60 and over are more likely to use [a:], while those under 60 are more likely to employ [ɑ:]. It appears that a change is in progress in the Southampton area, with the fronted realisation of the BATH vowel in decline.

6.3.3 The BATH Vowel and Speaker Sex and Age

Though sections 6.3.1 and 6.3.2 are of great interest, in that they show clear correlations between the BATH vowel and the speaker variables of sex and age, even more revealing is an analysis of the BATH vowel by both of these variables simultaneously. Figure 6.9 shows pronunciations of the BATH vowel in relation to the sex and age of speakers.

In section 6.3.1, [ɑ:] is shown to be the preferred variant for female informants. When sex and age are combined, however, it is clear that it is only the preferred option for women under the age of 60. Women aged 60 and over are more likely to employ the [a:] variant. In this instance, age appears to be a greater influence on use than sex.

Figure 6.9 Realisations of the BATH vowel according to the speaker variables of sex and age



Analysis in section 6.3.1 shows that there is little difference in the use of the two BATH variants by male informants. When combined with the speaker variable of age, however, it is apparent that [ɑ:] is in fact the preferred variant for males under the age of 30, accounting as it does for 85.04% of the tokens collected from this speaker group. Male informants over the age of 30 only employ the backed variant around a third of the time, preferring the [a:] realisation. Differences in

pronunciation between speaker groups are again highly statistically significant. The chi-squared value of 266.762 exceeds the 0.1% point of a chi-squared distribution with five degrees of freedom (20.52), making the differences significant at the 99.9% level.

An examination of the BATH vowel in relation to both age and sex highlights the interaction between these speaker variables and reveals patterns in speaker use of the vowel which are not apparent if the two variables are examined separately.

Males and females under the age of 30, and females aged 30-59 most frequently use [ɑ:], while men over 30 and women aged 60 and over prefer [a:].

6.3.4 The BATH Vowel and Speaker Social Class

Realisations of the BATH vowel according to the speaker variable of social class are shown in Figure 6.10.

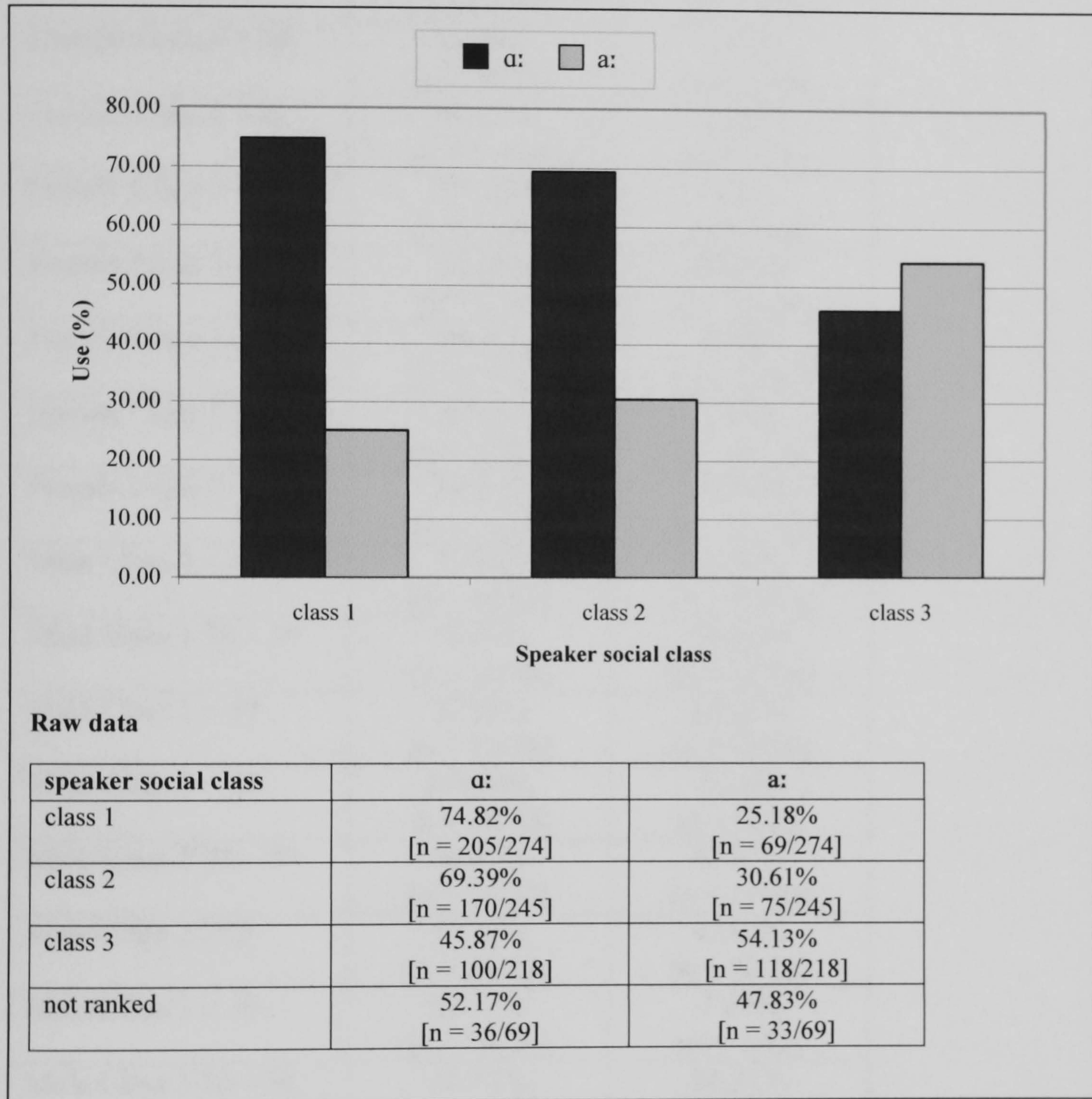
The data collected indicate social class differentiation in the use of the two BATH vowel variants. As one might expect, given that prestige or standard variants are more commonly found in the speech of those from higher socioeconomic backgrounds (Labov 1966: 73; Trudgill 1974: 61-63; Wells 1982: 14; Chambers 2003: 57-58), use of [ɑ:], the Received Pronunciation and south-eastern variant, is highest amongst those in the highest social class, and lowest amongst those in the lowest social class. Conversely, use of [a:] is lowest in class 1 and highest in class 3. This supports Wells' (1982: 346) assertion that, in Southampton, the front variant is preferred in working class speech and the back variant in middle class speech.

Again, these differences in use between classes are highly statistically significant.

The chi-squared value of 48.302 exceeds the 0.1% point of a chi-squared distribution with two degrees of freedom (13.82), making the differences significant at the 99.9%

level. The higher the social class of a speaker in the Southampton area, the more likely they are to use [ɑ:].

Figure 6.10 Realisations of the BATH vowel according to the speaker variable of social class



6.3.5 The BATH Vowel and Speaker Sex, Age and Social Class

In Table 6.1, realisations of the BATH vowel are shown according to the speaker variables of sex, age and social class.

Table 6.1 Realisations of the BATH vowel according to the speaker variables of sex, age and social class

	[ɑ:]	[a:]
Female Class 1 < 30	100.00% [n = 70/70]	0.00% [n = 0/70]
Female Class 1 30-59	100.00% [n = 21/21]	0.00% [n = 0/21]
Female Class 1 > 59	96.67% [n = 29/30]	3.33% [n = 1/30]
Female Class 2 <30	98.15% [n = 53/54]	1.85% [n = 1/54]
Female Class 2 30 - 59	100.00% [n = 44/44]	0.00% [n = 0/44]
Female Class 2 >59	10.26% [n = 4/39]	89.74% [n = 35/39]
Female Class 3 <30	100.00% [n = 10/10]	0.00% [n = 0/10]
Female Class 3 30 – 59	90.91% [n = 20/22]	9.09% [n = 2/22]
Female Class 3 >59	40.00% [n = 18/45]	60.00% [n = 27/45]
Male Class 1 < 30	73.85% [n = 48/65]	26.15% [n = 17/65]
Male Class 1 30 – 59	50.00% [n = 15/30]	50.00% [n = 15/30]
Male Class 1 > 59	37.93% [n = 22/58]	62.07% [n = 36/58]
Male Class 2 < 30	100.00% [n = 36/36]	0.00% [n = 0/36]
Male Class 2 30 – 59	36.67% [n = 11/30]	63.33% [n = 19/30]
Male Class 2 >59	52.38% [n = 22/42]	47.62% [n = 20/42]
Male Class 3 < 30	92.31% [n = 24/26]	7.69% [n = 2/26]
Male Class 3 30 – 59	31.43% [n = 22/70]	68.57% [n = 48/70]
Male Class 3 > 59	13.33% [n = 6/45]	86.67% [n = 39/45]

It is apparent from Table 6.1 that use of [ɑ:] is near categorical for class 1 females, regardless of age. For class 3 females, use of [ɑ:] declines with increasing speaker age. Class 3 females over 59 prefer [a:], whereas those under 60 prefer [ɑ:]. Realisations of the BATH vowel by class 2 females do not pattern as analysis

according to sex, age and social class separately suggest they will. Use of [ɑ:] is categorical or near categorical among those class 2 females under 60. However, [a:] is preferred by a large margin by those aged 60 and over. Use of this variant is higher among class 2 females over the age of 59 than it is among class 3 females aged 59 and over.

For male informants, who are responsible for approximately 75% of all tokens of rhoticity, there is much greater variation across age and class groups than for their female counterparts. The [a:] variant is used very little by men under 30, though it is noteworthy that the highest use of this form among men under 30 is in fact found among class 1 informants. Use of [a:] among males aged 30 to 59 increases as class lowers. Among those aged over 59, use of [a:] is higher among class 1 informants than it is among class 2 informants. It is however, highest among those in class 3. The [a:] variant is the preferred realisation for those males over 59 in classes 1 and 3, but not for those in class 2, who prefer [ɑ:]. Patterns are apparent in this analysis of the BATH vowel according to sex, age and social class which are not visible when this variable is analysed according to these extralinguistic variables separately.

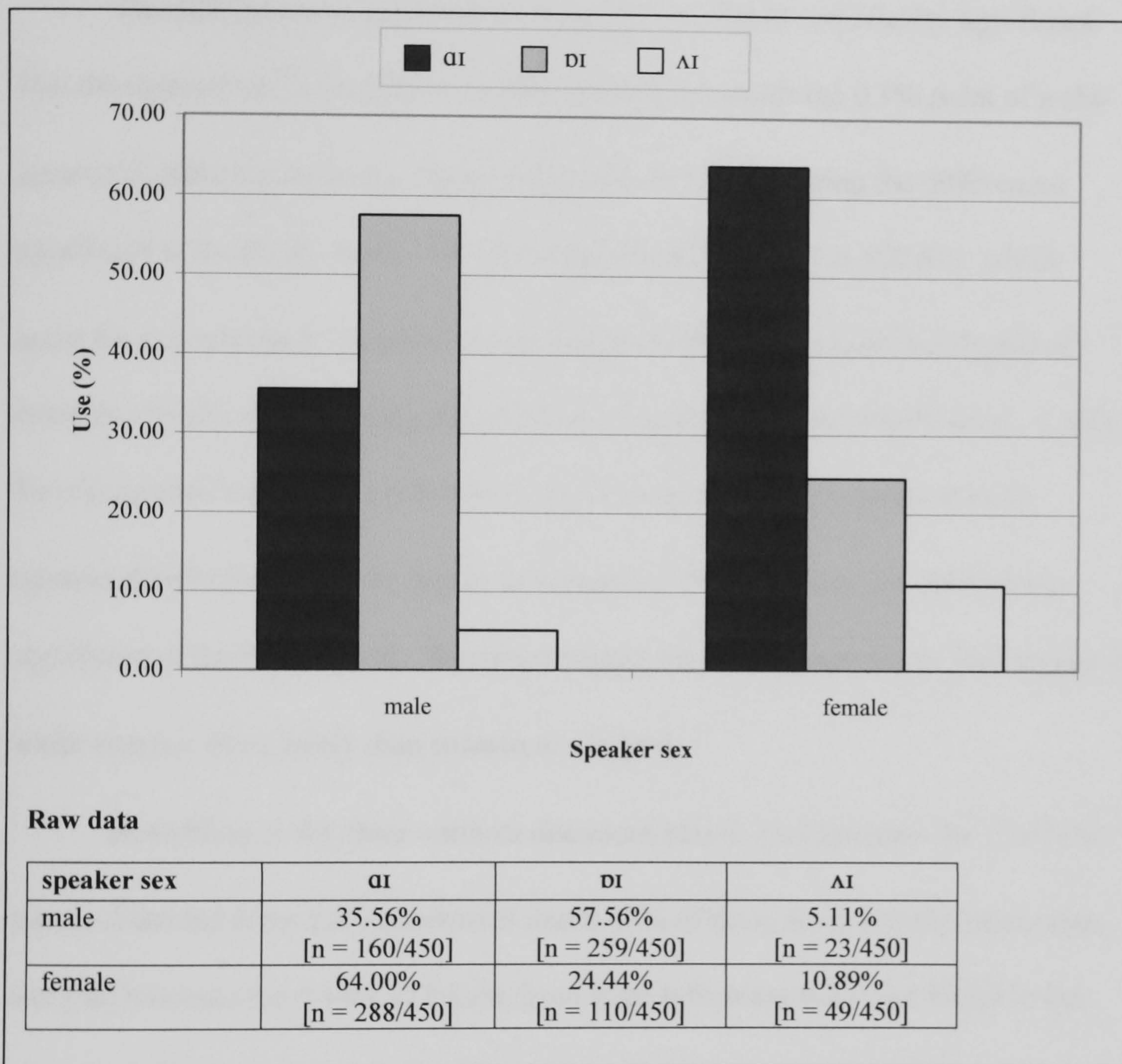
6.4 The PRICE Vowel

In the following sections, the PRICE vowel is analysed according to the speaker variables of sex, age and social class. In each section, the three most frequently occurring variants are analysed first, followed by the remaining two variants, which occur in very small quantities.

6.4.1 The PRICE Vowel and Speaker Sex

The three most frequently-occurring variants of the PRICE vowel are analysed in Figure 6.11 according to the speaker variable of sex.

Figure 6.11 Most frequently occurring realisations of the PRICE vowel according to the speaker variable of sex



For the female informants in the present study, [ɑI] is the preferred variant, while [ɒI] is most frequently used by the male informants. Since women tend to use more prestige or standard forms than men (Chambers 2003: 116), it is unsurprising that the Received Pronunciation variant [ʌI] is used more frequently by female speakers than it is by males. The higher use by women than men of [ɑI] also

suggests that this variant might be perceived as relatively prestigious in the Southampton area, or at least perhaps more prestigious than [ɒɪ]. Though it is not RP, it is closer to the RP [ʌɪ] realisation than the [ɒɪ] variant is, as both [ɑɪ] and [ʌɪ] are unrounded, while [ɒɪ] is rounded.

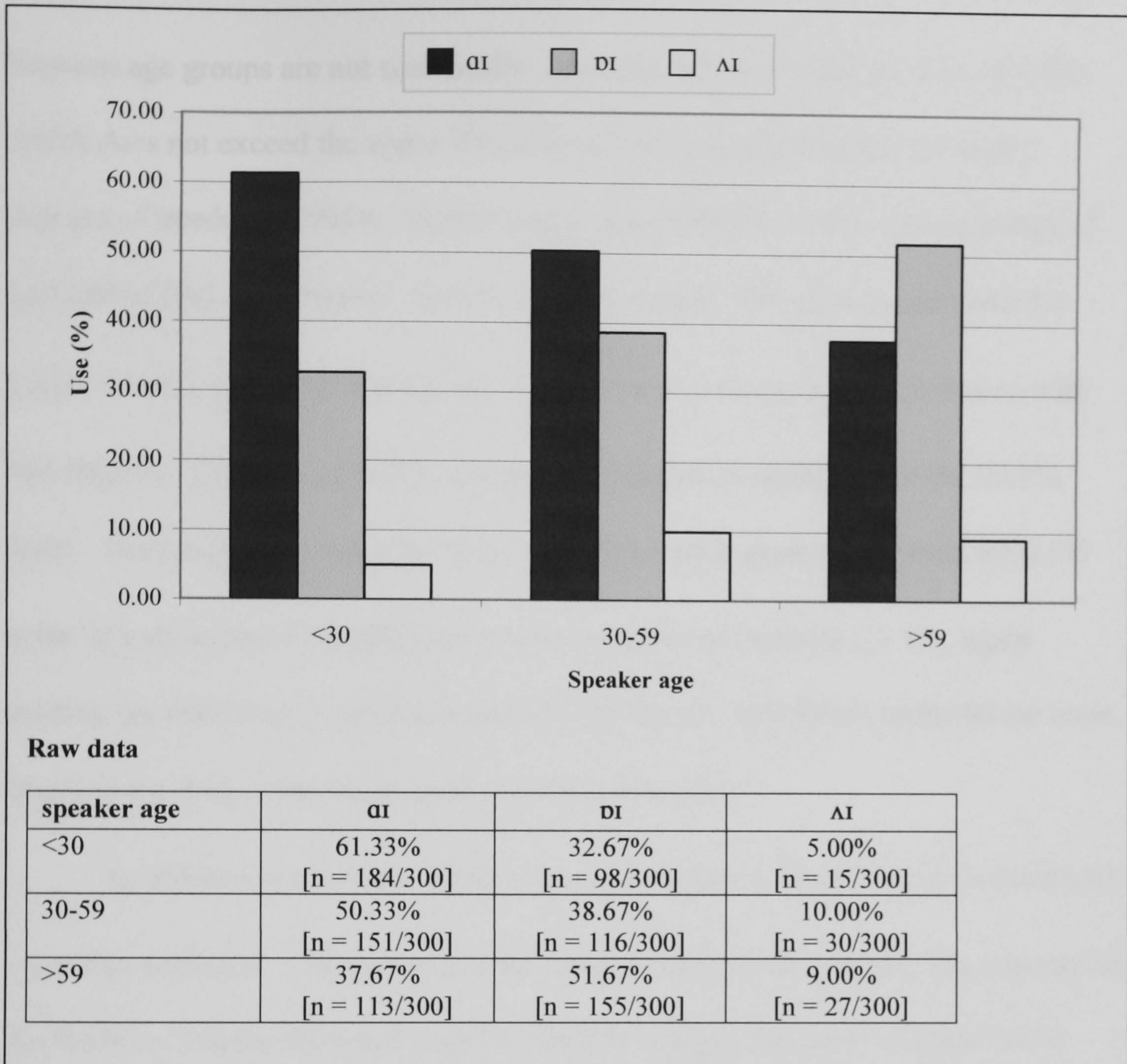
The differences in use between the sexes are highly statistically significant. The chi-squared value for [ɑɪ] is 71.686, which far exceeds the 0.1% point of a chi-squared distribution with one degree of freedom (10.83), making the differences significant at the 99.9% level. The chi-squared value for [ɒɪ] is 100.611, which again far exceeds the 0.1% point of a chi-squared distribution with one degree of freedom (10.83), again making the differences significant at the 99.9% level. Lastly, the chi-squared value for [ʌɪ] is 9.435, which exceeds the 0.5% point of a chi-squared distribution with one degree of freedom (7.879), making the differences significant at the 99.5% level. Women are more likely than men to use [ɑɪ] and [ʌɪ], while men are more likely than women to use [ɒɪ].

In addition to the three variants discussed above, [ɔɪ] accounts for 1.33% of tokens collected from male informants and 0.67% of those from female informants, and [əɪ] accounts for 0.44% of tokens from male informant but is not found in the speech of female informants. Levels of use of these two variants are too low, however, to draw conclusions regarding the effect of speaker sex on their use.

6.4.2 The PRICE Vowel and Speaker Age

The three most frequently occurring variants of the PRICE vowel are shown in Figure 6.12 according to speaker age.

Figure 6.12 Most frequently occurring realisations of the PRICE vowel according to the speaker variable of age



It is apparent from Figure 6.12 that there is variation in use of the PRICE vowel according to the age of the speaker. While use of [ɑI] decreases with increasing speaker age, use of [ɒI] increases with increasing age. The [ɑI] variant is preferred by informants under age of 60, while [ɒI] is preferred by those aged 60 and over. The [ʌI] variant is the least frequently employed of the three variants, regardless of speaker age. That it is used most frequently used by those aged 30 to 59 is perhaps to be expected, given that it is an RP variant and that the middle age group is most influenced by standard or prestige forms (Chambers and Trudgill

1998: 79). It should be noted, however, that the difference in use between the middle age group and those aged 60 and over is small, and that differences in use of [ʌɪ] between age groups are not statistically significant (the chi-squared value is 5.707 which does not exceed the upper 5% point of a chi-squared distribution with 2 degrees of freedom (5.991)). Differences in pronunciations between age groups of [ɑɪ] and of [ɒɪ] are, however, statistically significant. The chi-squared value for [ɑɪ] is 33.663, which far exceeds the 0.1% point of a chi-squared distribution with two degrees of freedom (13.82), making the differences significant at the 99.9% level. The chi-squared value for [ɒɪ] is 23.398, which again far exceeds the 0.1% point of a chi-squared distribution with two degrees of freedom (13.82), again making the differences significant at the 99.9% level. Informants under 60 are most likely to use [ɑɪ], while those aged over 60 favour [ɒɪ].

In addition to the three variants shown in Figure 6.12, tokens of [əɪ] and [ɔɪ] were also collected. The former is only found in the speech of under 30s, accounting for 0.67% of tokens collected from this speaker group. The latter is found in the speech of all three age groups, accounting for 0.33% of tokens collected from those aged under 30, 1.00% of those aged 30 to 59 and 1.67% of those collected from speakers aged 60 and over. While there is an increase in use of [ɔɪ] with increasing speaker age, numbers are too small to draw definitive conclusions regarding the effect of speaker age on this variant. The limited number of tokens also prevents us from doing so in the case of [əɪ].

In light of these data, and SED records for Hampshire which show [ɒɪ] and [ɔɪ] realisations of the PRICE vowel in the word *lights* (Orton and Wakelin 1967-68: 356), it would appear that a change is in progress in the Southampton area, with use

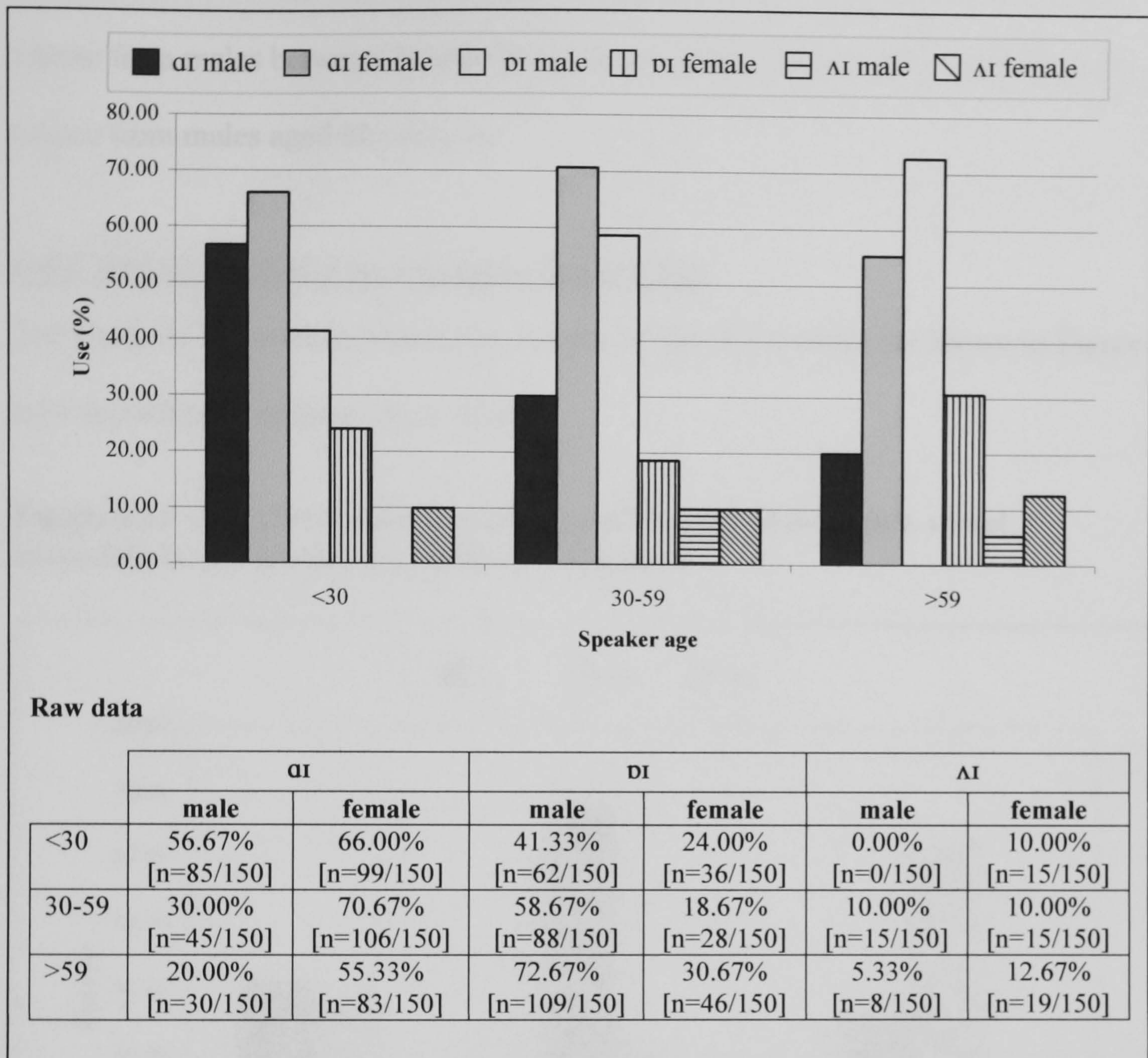
of [aɪ] on the increase and [ɒɪ] and [ɔɪ] decreasing, the latter more rapidly than the former.

6.4.3 The PRICE Vowel and Speaker Sex and Age

Although there are differences in pronunciations of the PRICE vowel between speaker groups when analysed by age and sex separately, of much greater interest are the differences which become apparent when PRICE is analysed by sex and age simultaneously. The three most frequently-occurring variants of the PRICE vowel are shown in Figure 6.13 according to speaker age and sex.

In the case of speakers under the age of 30, [aɪ] is the preferred variant, regardless of speaker sex. However, in the case of those speakers aged between 30 and 59 and those aged 60 and over, [aɪ] is the preferred variant for females only, with male informants more likely to use [ɒɪ] instead. Use of [aɪ] falls among male informants as age increases. Conversely, use of [ɒɪ] among male speakers increases as age increases. Use of [ɒɪ] by female informants is highest among those aged 60 and over, followed by those under the age of 30. It is lowest among those females aged between 30 and 59. In the under 30 and the over 59 age groups, use of [aɪ] is higher for females than it is for males. In the 30 to 59 age group, there is no such sex differentiation.

Figure 6.13 Most frequently occurring realisations of the PRICE vowel according to the speaker variables of sex and age



These differences in use of three most frequently occurring realisations of the PRICE vowel are highly statistically significant. The chi squared value for [ɑɪ] is 123.344, for [ɒɪ] it is 137.013, and for [Δɪ] it is 21.377. All of these exceed the upper 0.1% point of a chi-squared distribution with five degrees of freedom (20.52), making the differences significant at the 99.9% level.

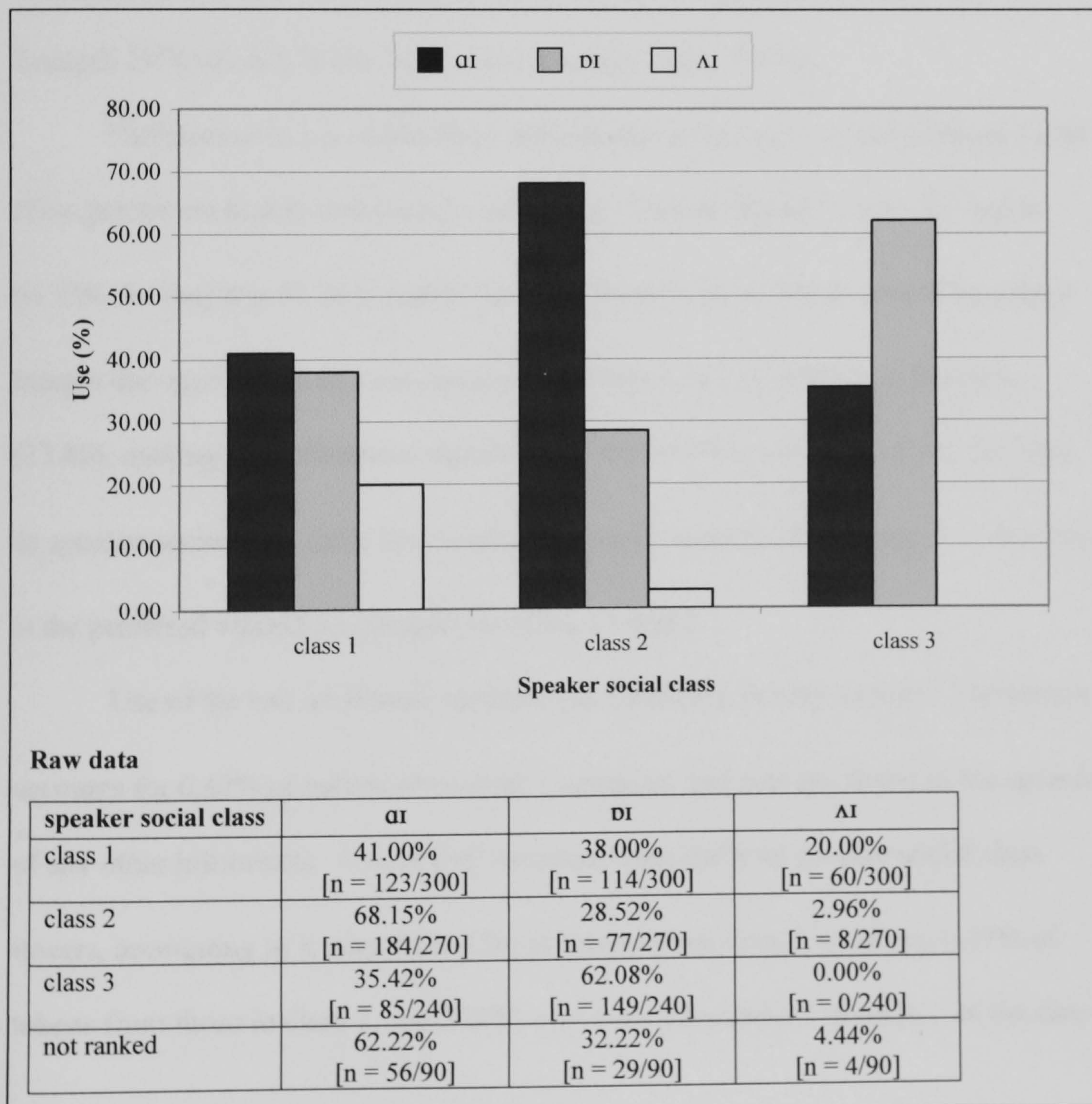
In the case of the remaining two realisations of PRICE, [əɪ] is found only in the speech of males under the age of 30, accounting for 1.33% of tokens from this speaker group. The other variant, [ɔɪ], is not found in the speech of females under

the age of 30 but was used by all other speaker groups, accounting for 0.67% of tokens from males under 30 and females between the ages of 30 and 59, 1.33% of tokens from males between 30 and 59 and females aged 60 and over, and 2.00% of tokens from males aged 60 and over.

6.4.4 The PRICE Vowel and Speaker Social Class

The three most frequently-occurring variants of the PRICE vowel are shown in Figure 6.14 according to speaker social class.

Figure 6.14 Most frequently occurring realisations of the PRICE vowel according to the speaker variable of social class



For speakers in classes 1 and 2, [ɑɪ] is the preferred variant, while for those in class 3, [ɒɪ] is most frequently used. In the case of class 1 speakers, however, use of [ɑɪ] and [ɒɪ] is very similar, each accounts for around 40.00% of the total number of tokens, while class 2 speakers are more than twice as likely to use [ɑɪ] as they are to use [ɒɪ].

Use of the [ʌɪ] variant shows clear social class differentiation, with use of the form increasing the higher the speaker social class. This is to be expected given that this is the RP variant, and that prestige or standard forms are more commonly found in the speech of those from a high socioeconomic background (Labov 1966: 73; Trudgill 1974: 61-63; Wells 1982: 14; Chambers 2003: 57-58).

Differences in use of the three realisations of the PRICE vowel between social class groups are highly statistically significant. The chi-squared value for [ɑɪ] is 64.939, for [ɒɪ] it is 61.863, and for [ʌɪ] it is 84.891, all of which exceed by a large margin the upper 0.1% of a chi-squared distribution with 2 degrees of freedom (13.82), making the differences significant at the 99.9% level. Use of [ʌɪ] declines as speaker social class falls, [ɒɪ] is most frequently used by those in class 3, and [ɑɪ] is the preferred variant for speakers in classes 1 and 2.

Use of the two additional variants, [əɪ] and [ɔɪ], is very limited. The former accounts for 0.67% of tokens from class 1 speakers, and was not found in the speech of any other informants. Use of [ɔɪ] increases marginally as speaker social class lowers, accounting as it does for 0.33% of tokens from class 1 speakers, 0.37% of tokens from those in class 2, and 2.50% of tokens from class 3 speakers. In the case

of both [əɪ] and [ɔɪ], however, use is too low to draw meaningful conclusions regarding the possible effect of speaker social class on these variants.

6.5 Phonology – A Summary

All of the phonological variables investigated are highly significant in the Southampton area dialect. In the case of L Vocalization, it is apparent that use of this form is on the increase in the Southampton area. L Vocalization does not vary significantly according to speaker sex or social class, indicating that it is not a stigmatised feature in the Southampton area. Rhoticity, by contrast, is declining in the Southampton area, being found most frequently in the speech of older males. Use of this variant does not vary according to speaker social class, indicating that sex and age are stronger influences on rhoticity than class. Realisations of the BATH vowel differ significantly according to speaker sex, age and social class. While [ɑ:] is the variant preferred by those men under 30, women of all ages, and those in classes 1 and 2, men over 30 and those in class 3 prefer [a:]. This indicates a change in progress away from the traditionally south-western [a:] to [ɑ:], led by the young, and by women. Realisations of the PRICE vowel also differ significantly according to sex, age and social class. Female informants of all ages, males under 30 and informants in classes 1 and 2 are more likely to use [ɑɪ] than their older male counterparts and those in class 3, who favour [ɒɪ]. The variant [ʌɪ] is most frequently used by class 1 informants and by women, indicating that it is a prestigious variant in the Southampton area.

Altendorf and Watt state:

[I]t has [...] been reported that parts of the Southwest are changing [...] and are [...] doing so under the influence of the expanding Home Counties Modern Dialect Area. Trudgill [...] claims that this is true for coastal cities such as Southampton, Portsmouth and Bournemouth. From a geographical point of view, these cities are part of the Central Southwest. From a dialectological point of view, they may not belong to this area any longer.

(Altendorf and Watt 2004: 197)

In many respects, this quotation is a contentious one. There is no consensus as to whether Southampton lies in the South East or the South West, either geographically or linguistically (see section 1.4 for linguistic perspectives on the location of the Southampton area, and section 7.6 for the opinions of the Southampton area's informants regarding the position of the city). Altendorf and Watt (2004: 196) themselves admit that '[t]he West Country is a region with imprecise boundaries'. However, in the present study, the limited use by informants under the age of 30 of phonological features typically associated with the South West, the [a:] realisation of the BATH vowel (see section 6.3.2) and rhoticity (see section 6.2.2), does indicate that a change is in progress in the Southampton area. Similarly, the higher use of L Vocalization by informants under 60 than by those aged 60 and over (see section 6.1.2) indicates an increasing use of a traditionally south-eastern feature. Whether these changes are under the influence of a 'Home Counties Modern Dialect area' is another matter. The data collected do suggest, however, that, at least as far as rhoticity and the BATH vowel are concerned, the accent of the Southampton area is moving away from that traditionally associated with the South West.

Chapter 7

7.0 The Speaker Variable of Identity in the Southampton Area

The present survey collected both quantitative and qualitative data concerning speaker identity in the Southampton area. Speakers firstly completed a questionnaire containing four questions, three of which are based on those used by Underwood (1988) in his study of Texan identity. Responses to these three questions were quantified. The questionnaire is discussed at length in section 3.5.1 and a copy of it can be found in 3.5.1.1. In addition to quantitative data, the Southampton area identity questionnaire was also used to collect qualitative data. Speakers were asked to expand on their identity questionnaire answers during the SuRE interview, and their responses were used as a starting point for a more general discussion about the local area. Further questions were posed concerning speakers' likes and dislikes about the local area and what they would like to see changed or improved.

7.1 Findings of the Southampton Area Identity Questionnaire

Responses to each of the four Southampton area identity questions are presented in the following sections, both as an overall result from the entire speaker sample and, in order to ascertain to what extent identity is influenced by other speaker variables, according to sex, age and social class. For overall results, Tables 7.1 to 7.4 show the number of responses to a particular question as a proportion of the total number of speakers. For example, in Table 7.1 twenty-nine speakers out of a possible sixty (29/60) say that they feel closer to local people than to others. Tables 7.1 to 7.4 then express these results according to the sex of the speakers (there are thirty male and thirty female informants), according to speaker age (there are twenty speakers in each of the three age groups), and according to speaker social class (there are 20

informants in class 1 (the highest social group), 18 in class 2, and 16 in class 3). In the case of speaker social class, class sizes are not equal, and results are given in percentages as well as tokens in order to make the comparison of findings across classes easier. For reasons stated in section 3.9.3, six informants are not allocated to a social class, so it is not possible to analyse their responses to the Southampton area identity questionnaire according to this speaker variable. Their responses are, however, analysed according to age and sex.

7.1.1 Question One

Question one asks informants to say whether they feel closer to local people than to others. The responses given to this question can be seen in Table 7.1.

Table 7.1 Responses to question one of the Southampton area identity questionnaire

Feel closer to local people			No closer to local people than to others			Don't know, can't say		
Overall result			Overall result			Overall result		
29/60			24/60			7/60		
Male		Female	Male		Female	Male		Female
18/30		11/30	10/30		14/30	2/30		5/30
<30	30-59	>59	<30	30-59	>59	<30	30-59	>59
10/20	10/20	9/20	5/20	10/20	9/20	5/20	0/20	2/20
Class 1	Class 2	Class 3	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
40%	61.11%	43.75%	30%	38.89%	50%	30%	0%	6.25%
[n = 8/20]	[n = 11/18]	[n = 7/16]	[n = 6/20]	[n = 7/18]	[n = 8/16]	[n = 6/20]	[n = 0/18]	[n = 1/16]

Many of the twenty-nine informants who claim to feel closer to local people than to others cite having something in common with other locals as the reason for this. JH, a 73-year-old class 2 male from Eastleigh Borough, comments:

Oh yes, you feel closer to local people. [He is asked why that is]. I suppose familiarity. You have something in common. An area in common, haven't you?

It is a sentiment echoed by DD, an 18-year-old class 1 male from Southampton, who states:

I feel close to local people because if you're on holiday and someone says, "Oh, I'm from Southampton", all of a sudden you've got something to talk about, something in common.

When asked why she feels closer to local people than to others, JH, a 27-year-old class 2 female from Eastleigh Borough, comments:

You always feel that you might know someone through someone else. [...] I always feel that [...] somehow you're all connected by the people you know.

Other informants cite a shared way of speaking as the reason for feeling closer to local people than to others. NC, a 51-year-old class 2 male from Southampton, states:

I do find it easier, maybe just because we have the same accent, to get on with somebody like BO [the other informant in the interview], who I must confess we disagree a lot about politics, but I can get on with BO because he's just a mush, that's what I would say, he's a mush, he's a local. I feel comfortable when I'm talking to him.

JS, a 56-year-old class 3 male from Eastleigh Borough, echoes this sentiment:

JS: Feel closer to local people myself, yes.

KW: Why do you think that is?

JS: [...] You can understand what they're saying. The number of [...] times you phone up [...] a company and someone Scotch answers [...] and you've got a terrible job to understand what they're saying. [...] I just find you can communicate to local people rather than some northerners [...]. The accent is so, especially round Durham way, the accent is so, so fierce, if I can use that word [...] it's so difficult to actually relate to these people as well.

AH, a 41-year-old class 3 male from Eastleigh Borough, makes a similar comment:

You've just got a common bond with people from your type of area rather than [...] you don't feel so close to people that maybe don't talk the same as you.

These comments appear to support Tabouret-Keller's statement that:

linguistic items are not just attributes of groups or communities, they are themselves the means by which individuals both identify themselves and identify with others.

(Le Page and Tabouret-Keller 1985: 5)

The speakers quoted claim to identify with other local people as a result of a shared dialect.

In contrast to those who express feelings of closeness to other locals, twenty-four speakers say that they feel no closer to local people than to others. TS, a 20-year-old class 1 male from Southampton, states:

To be honest, I don't feel any closer to people from the area [...]. Comparing it, say, when I go on holiday and you meet people from all over Britain [...] I get on as well with them as I do from people round here.

A similar sentiment is expressed by GL, a 48-year-old class 1 male from Southampton:

I know I've lived in Southampton all my life really but [...] I don't feel a huge affinity towards it. [...] The only affinity would be if I'm in the other side of the world or something and I met somebody from Southampton I think.

Like GL, MP, a 65-year-old female from Southampton, unclassified according to social class, feels no closer to locals than those from other areas. She states:

I take people individually. I tend not to judge them from where they come from or even what relation they are to me. I take people as individuals.

More male informants than female informants claim to feel close to local people (eighteen males compared to eleven females). The difference between the two sexes is not statistically significant, however, as the chi-squared value, 2.403, does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). There is little variation between the number of speakers in each of the three age groups who report feeling closer to local people than to others (ten in

both the youngest and middle groups and nine in the oldest). With a chi-squared value of 0.134, again this difference is not statistically significant, as it does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). A higher proportion of class 2 informants report feeling closer to local people than do class 1 or class 3 informants. As in the case of differences according to speaker sex and age, however, the differences between social classes are not statistically significant, as the chi-squared value of 1.867 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

7.1.2 Question Two

Question two is designed to measure ‘in-group preference’ (Reed, cited in Underwood 1988: 410), and asks informants to imagine that they are the manager of a company that has to hire a scientist. Two equally qualified people apply for the job, one local and one from elsewhere. The question asks speakers to say who they would choose. The responses given to this question can be seen in Table 7.2.

Question two elicits fewer feelings of local identity than question one.

Seventeen informants say that they would employ the person from Southampton or Eastleigh Borough. RS, a 21-year-old class 1 male from Eastleigh Borough, argues that by giving the job to the candidate from Southampton/Eastleigh Borough ‘you’d be helping out someone local’. JH, a 27-year-old class 2 female from Eastleigh Borough, echoes this sentiment when she says ‘I feel you should stick together if you’re from the same area [...] give somebody else a chance who’s local’. Several informants cite the importance of local knowledge, including the geography of the area, as the reason for choosing a local applicant.

Table 7.2 Responses to question two of the Southampton area identity questionnaire

Person from Southampton/Eastleigh Borough			Person from another area			It depends, don't know		
Overall result			Overall result			Overall result		
17/60			3/60			40/60		
Male		Female	Male		Female	Male		Female
11/30		6/30	1/30		2/30	18/30		22/30
<30	30-59	>59	<30	30-59	>59	<30	30-59	>59
6/20	6/20	5/20	2/20	0/20	1/20	12/20	14/20	14/20
Class 1	Class 2	Class 3	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
15%	38.89%	43.75%	10%	0%	6.25%	75%	61.11%	50%
[n = 3/20]	[n = 7/18]	[n = 7/16]	[n = 2/20]	[n = 0/18]	[n = 1/16]	[n = 15/20]	[n = 11/18]	[n = 8/16]

Forty informants, two-thirds of those interviewed, reply that their decision as to who to appoint would depend on factors other than the geographical origins of the candidates, or that they do not know who they would choose. MW, a 70-year-old class 2 female from Eastleigh Borough, states that 'ability and how they get on with the rest of the staff' is more important than where the job applicant is born and raised. SH, a 41-year-old class 2 female from Eastleigh Borough, says that her decision as to who to appoint would 'come down to the person that you liked'. The personality of the candidates is cited by a number of informants as being an important part of the decision-making process.

RD, a 24-year-old class 1 male from Eastleigh Borough and one of the three informants who say that they would choose someone from another area, justifies his decision by saying that 'they'd bring different ideas and experiences and ways of working rather than having someone that thinks the same way as you'.

Over a third of male informants, eleven out of thirty, compared with only a fifth of female informants, six out of thirty, claim that they would employ a local person for the job. RW, a 69-year-old class 2 male from Eastleigh Borough, openly

admits his decision to hire a local person for the job could be construed as 'favouritism' but is nevertheless willing to admit that this is what he would do. The difference between male and female responses is not statistically significant, as the chi-squared value of 1.313 does not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). However, more men than women in this instance report having strong feelings of local identity.

Like question one, responses to question two do not appear to vary greatly with speaker age. Six informants in both the under 30 and 30 to 59 age groups say that they would choose a local candidate, compared to five speakers aged over 59. This small difference is not statistically significant as the chi-squared value of 0.164 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

More informants from classes 2 and 3 claim they would hire a local person than do class 1 informants. The differences between classes are again not statistically significant, however, as the chi-squared value of 4.093 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

7.1.3 Question Three

Like question two, question three is also designed to measure 'in-group preference' (Reed, cited in Underwood 1988: 410), and asks speakers who they would favour for their local Member of Parliament: a person born and brought up locally, or someone from outside the area. The responses given to this question can be seen in Table 7.3.

Table 7.3 Responses to question three of the Southampton area identity questionnaire

Person from Southampton/Eastleigh Borough			Person from another area			It depends, don't know		
Overall result			Overall result			Overall result		
55/60			0/60			5/60		
Male		Female	Male		Female	Male		Female
28/30		27/30	0/30		0/30	2/30		3/30
<30	30-59	>59	<30	30-59	>59	<30	30-59	>59
19/20	19/20	17/20	0/20	0/20	0/20	1/20	1/20	3/20
Class 1	Class 2	Class 3	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
85%	94.44%	100%	0%	0%	0%	15%	5.56%	0%
[n = 17/20]	[n = 17/18]	[n = 16/16]	[n = 0/20]	[n = 0/18]	[n = 0/16]	[n = 3/20]	[n = 1/18]	[n = 0/16]

Question three evokes very intense feelings of local identity in informants, far more so than question two. One possible explanation for this is that politics is a subject about which people often have very strong opinions. Another possible explanation is that respondents are perhaps concerned that they might look less than even-handed, particularly in light of anti-discrimination laws, if they were to reply that they would definitely select a local person over a person from elsewhere for a job, thus producing a marked difference in the responses to the two questions.

Fifty-five out of sixty speakers report that they would choose a local person as their Member of Parliament. Most cite a local person's knowledge of local issues as the reason for their response. When questioned about his reasons for choosing a Member of Parliament who has been born and raised locally, RU, a 64-year-old class 1 male from Southampton, states:

They know the problems, they know the area, they know the people, they know the facilities and if they live in Southampton, they've been there long enough, they'll know what's required.

A shared way of speaking is also cited as a factor in the decision to choose a local person to act as a local Member of Parliament. GM, a 53-year-old class 1 male from Eastleigh Borough, states of a Member of Parliament:

I think they've got to be able to relate to people and regardless of whether [...] you like it or not, the average middle class and working class person relates to somebody who has the same dialect as them and can also talk about where they come from.

A similar view is expressed by ME, an 83-year-old class 2 female from Southampton:

How can a man [...] that's been brought up with a Scottish background [...] come down and be an M[ember of] P[arliament] for, say, Southampton, right down here in the South? I mean, the morals are different, the language is different [...]. I do think, on balance, a local man is better to represent.

As in the case of the comments made in response to question one, these statements appear to support Tabouret-Keller's notion that linguistic features function as a means by which speakers identify with others (Le Page and Tabouret-Keller 1985: 5).

The remaining five informants say that they do not know who they would choose, or that their decision would depend on factors other than the geographical origins of the candidates. The ability of the two candidates to do the job, and their personalities, are cited among these factors. No informants claim that they would choose a person from another area to be their Member of Parliament.

There is little sex-based variation in the responses to question three, and what difference there is, is not statistically significant. Similarly, there is no statistically significant age-based or class-based variation in the responses to question three. For those speakers who claim that they would choose a local Member of Parliament, the chi-squared values are as follows: 0.000 according to speaker sex; 0.491 according to speaker age; and 1.207 according to speaker social class. None of these values are

statistically significant, as they do not exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841) in the case of analysis by speaker sex, or two degrees of freedom (5.991) in the case of analysis by speaker age and by social class. That a local Member of Parliament has been born and raised in the Southampton area appears to be of great importance to informants, regardless of the informant's sex, age or social class.

7.1.4 Question Four

Question four asks informants whether they feel closer to people from the South East or the South West of England. The responses given to this question can be seen in Table 7.4.

Table 7.4 Responses to question four of the Southampton area identity questionnaire

South East			South West			Neither		
Overall result			Overall result			Overall result		
7/60			18/60			35/60		
Male	Female		Male	Female		Male	Female	
4/30	3/30		10/30	8/30		16/30	19/30	
<30	30-59	>59	<30	30-59	>59	<30	30-59	>59
4/20	2/20	1/20	4/20	9/20	5/20	12/20	9/20	14/20
Class 1	Class 2	Class 3	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
5%	16.67%	12.50%	25%	38.89%	25%	70%	44.44%	62.50%
[n = 1/20]	[n = 3/18]	[n = 2/16]	[n = 5/20]	[n = 7/18]	[n = 4/16]	[n = 14/20]	[n = 8/18]	[n = 10/16]

'South East' is the least popular response to question four, with only seven speakers saying that they feel closer to people from this area. Eighteen respondents say that they feel closer to people from the South West. However, the majority of speakers, thirty-five out of sixty, claim that they do not feel close to people from either the South East or the South West. When questioned about this, many commented that they felt Southampton to be separate from both the South East and

the South West. East-west identity in the Southampton area is discussed at greater length in section 7.6.

There is no statistically significant sex-based variation in responses to question four. The chi-squared value for the response 'South East' according to speaker sex is 0.000, and for 'South West', it is 0.079. Neither values exceed the upper 5% point of a chi-squared distribution with one degree of freedom (3.841). Age-based variation is, however, apparent. It is perhaps unsurprising, given the stereotypical association of the South East (encompassing London) with fashion and progress and the South West with a more rural way of life, that the age group with the highest number of speakers who feel close to people from the South East is the under 30s. JS, an 18-year-old class 2 female from Eastleigh Borough, states that she feels closer to people from the South East 'because I think we're more like London [...] whereas the South West is a bit more agricultural'. DW, a 25-year-old class 3 male from Southampton cites 'the way of talking' as part of the reason that he feels closer to people from the South East. Though he does not claim to feel close to people from either the South East or the South West, RD, a 24-year-old class 1 male from Eastleigh Borough, states:

In terms of lifestyle, I think we're closer to the South East, in terms of how we live, the money and education and everything else.

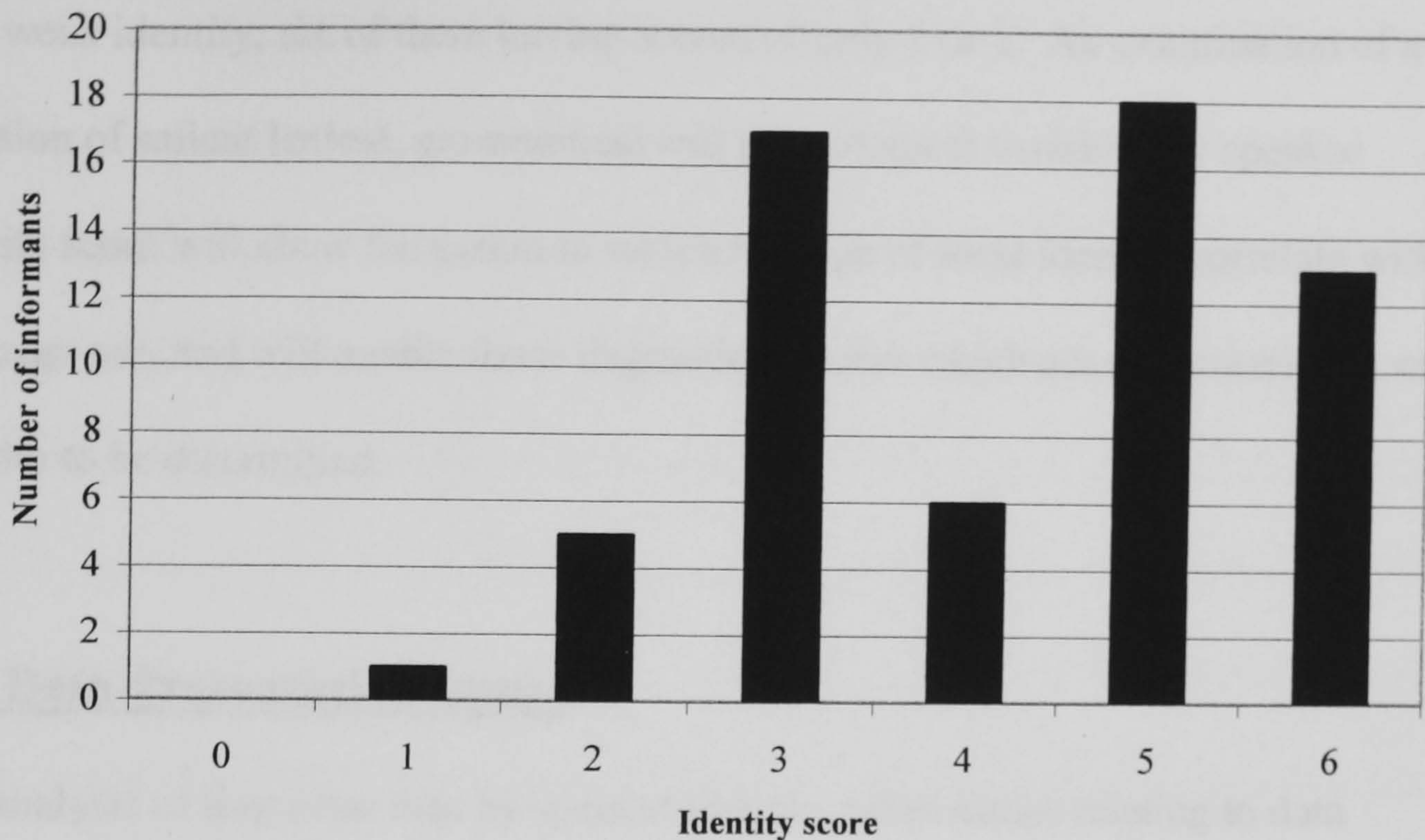
It can be seen from Table 7.4 that identification with people from the South East decreases steadily with increasing age. It should be noted, however, that the differences between age groups in the number of speakers claiming to feel closer to people from the South East are not statistically significant. The chi-squared value is 1.011 which does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Table 7.4 also shows that closeness to people from the South West is lowest among the under 30s. DD, an 18-year-old class 1 male from Southampton says of the South West that ‘you start thinking of carrot crunchers’. In contrast to this, RA, a 57-year-old male from Southampton who is unclassified by social class says that ‘from the point of view of accent and feeling [...] the farther east I go, the less comfortable I feel with people’. Again, however, the differences between age groups in the number of speakers claiming to feel closer to people from the South West are not statistically significant. The chi-squared value is 3.333 which does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). Similarly, the differences between social classes in the number of speakers claiming to feel closer to people from the South East or the South West are not statistically significant. The chi-squared value for the response ‘South East’ according to speaker social class is 0.454, and the chi-squared value for the response ‘South West’ according to speaker social class is 0.424, neither of which exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

7.1.5 Southampton Area Identity Scores

Speaker responses to the first three questions of the Southampton area identity questionnaire have been used to calculate an identity score for each informant. A positive answer (one favouring local people) has received a score of 2, a neutral response (answering *don't know*, *can't say* or *it depends*) a score of 1, and a negative response (one favouring people from outside the area) a score of 0. These scores have then been used to calculate a Southampton area identity score for each speaker of between 0 (low identification with the Southampton area) and 6 (high identification with the Southampton area). The results can be seen in Figure 7.1.

Figure 7.1 Southampton area identity scores



Raw data

Southampton area identity score	Informants with that score		
0	Overall result: 0/60		
	Male: 0/30		Female: 0/30
	<30: 0/20	30-59: 0/20	>59: 0/20
	Class 1: 0/20	Class 2: 0/18	Class 3: 0/16
1	Overall result: 1/60		
	Male: 1/30		Female: 0/30
	<30: 1/20	30-59: 0/20	>59: 0/20
	Class 1: 1/20	Class 2: 0/18	Class 3: 0/16
2	Overall result: 5/60		
	Male: 1/30		Female: 4/30
	<30: 0/20	30-59: 1/20	>59: 4/20
	Class 1: 2/20	Class 2: 1/18	Class 3: 1/16
3	Overall result: 17/60		
	Male: 6/30		Female: 11/30
	<30: 5/20	30-59: 8/20	>59: 4/20
	Class 1: 4/20	Class 2: 5/18	Class 3: 6/16
4	Overall result: 6/60		
	Male: 3/30		Female: 3/30
	<30: 3/20	30-59: 1/20	>59: 2/20
	Class 1: 4/20	Class 2: 1/18	Class 3: 1/16
5	Overall result: 18/60		
	Male: 11/30		Female: 7/30
	<30: 6/20	30-59: 5/20	>59: 7/20
	Class 1: 7/20	Class 2: 5/18	Class 3: 3/16
6	Overall result: 13/60		
	Male: 8/30		Female: 5/30
	<30: 5/20	30-59: 5/20	>59: 3/20
	Class 1: 2/20	Class 2: 6/18	Class 3: 5/16

While the majority of respondents identify very strongly with other local people, thirty-one out of sixty having identity scores of 5 or 6, other speakers have very weak identity, six of them having scores of only 1 or 2. An examination of a selection of salient lexical, grammatical and phonological variables by speaker identity score will show the extent to which feelings of local identity correlate with language use, and will enable those linguistic features which act as markers of local identity to be determined.

7.2 Data Presentation Issues

The analysis of linguistic data by speaker identity raises issues relating to data presentation not as relevant to the analysis of that same linguistic data by the speaker variables of sex, age and social class. When analysing linguistic data by the extralinguistic variable of sex, the Southampton area speaker sample of sixty is divided into two groups, male and female. In the case of the speaker variable of age, the speaker sample is divided into three groups: under 30; 30 to 59; and 60 and over. In the case of the speaker variable of social class, the speaker sample is again divided into three groups, though in this instance, of uneven sizes: class 1 (20 informants); class 2 (18 informants); and class 3 (16 informants). These subdivisions are relatively unproblematic, as the resultant groups are sufficiently large for the conclusions drawn from them to be meaningful. Identity, however, is measured in the present study on a six point scale, resulting in a potential seven groups (0 to 6, although in fact only values of 1 to 6 arise). To complicate this matter further, the numbers of speakers in each of these groups are unequal, as in the case of the social class groupings. Results drawn from these six groups might not be reliable, not only

because some categories might be very small but also because the categories are not directly comparable, being comprised of differing numbers of speakers.

The subdivision of the speaker sample by the six point identity scale is a particular issue as regards the analysis of grammatical data. Very few of the grammatical features investigated by the Southampton area language questionnaire are reported in high frequencies. The decision has been made to consider for examination by extralinguistic variable only those features reported by ten or more speakers (see section 3.10.2). Even when the cut-off point of ten speakers is applied to the selection of grammatical variables, to subdivide those informants who report using a particular variable into groups according to their identity scores would result in very small numbers of informants in each of the six categories. Conclusions drawn from such small numbers would be unreliable.

The decision has therefore been taken to group the identity scores according to the low, medium and high levels of identification which they represent (identity scores of 1 and 2 corresponding to low identification, scores of 3 and 4 to medium identification and scores of 5 and 6 to high levels of identification). A selection of lexical, grammatical and phonological features are analysed according to these levels of speaker identification, and possible correlation between reported use of these forms and speaker identification is more clearly seen than if each of the six identity scores was used.

Lexical data are presented, as in previous chapters on sex and age, as a percentage of the total number of responses given according to extralinguistic variable, in this case, level of speaker identity. Taking the example of the notion word *pregnant* and the variant *preggers*, the percentage use according to level of speaker identity is calculated as follows:

$$\left(\frac{\text{tokens of } \textit{preggers} \text{ reported by speaker group sharing same level of speaker identity}}{\text{total number of tokens offered for notion word } \textit{pregnant} \text{ by the same speaker group}} \right) \times 100$$

This method of data presentation does not need to be altered to account for the differing number of speakers in each speaker identity level group, as what is of interest is not the number of respondents but rather the number of responses, each notion word potentially having multiple responses. The method accounts for the differing number of responses offered by informants, allowing for the comparison of reported use of variants across identity levels (see section 3.10.1 for a full explanation of this data presentation method).

Reported use of a grammatical form by a particular identity group is presented as a percentage of the total number of informants having that particular level of identity. For example, *that ain't a bird* is reported by 5 speakers with medium identity and 6 with high identity. It is not reported by anyone with low identity. The percentage use according to speaker identity for this variable is calculated as follows:

$$\left(\frac{\text{number of speakers reporting } \textit{that ain't a bird} \text{ in a particular identity group}}{\text{total number of speakers in that same identity group}} \right) \times 100$$

In the case of speakers with medium identity, use of *that ain't a bird* is 21.74% ((5/23) x 100), whereas for those with high identity, it is 19.35% ((6/31) x 100), despite more speakers from the latter group reporting their use of the variable. (*Ain't* and speaker identity is discussed at greater length in section 5.4.1.) Grammatical data are not presented by number of informants reporting a particular form as, unlike in the case of sex and age, the number of informants in each identity category is not

comparable. Presented by number of informants reporting the form, it appears that *that ain't a bird* is more likely to be used by a speaker with high identity than one with medium identity as it is reported by 5 speakers with medium identity and 6 with high identity. However, it could be the case that *that ain't a bird* is reported by more speakers with high identity than those with medium identity simply because there are more speakers in the former group than in the latter (31 compared to 23) rather than because of the degree of local identity felt by the speakers.

Phonological data are shown, as in chapter 6, as a percentage of the total possible number of tokens of a particular variable according to speaker group, in this case low, medium or high identity. This percentage is calculated for each identity group by dividing the number of tokens of a particular variant offered by the total possible number of tokens, and multiplying by 100. This method of data presentation is discussed fully in section 3.10.3.

7.3 Identity and Lexis

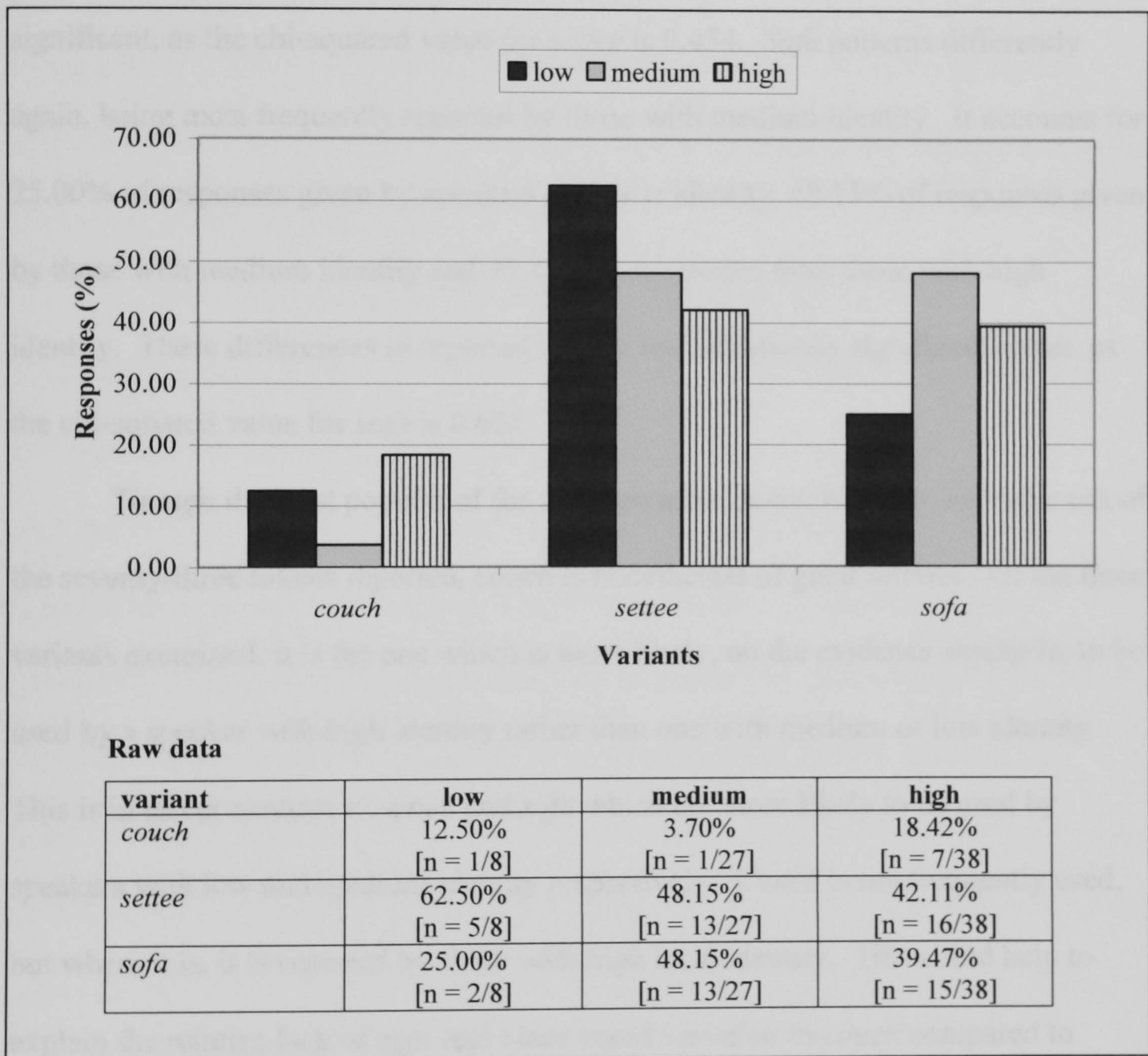
As in chapter 4, in this examination of the possible relationship between lexical use and speaker identity, for each variable the most frequently-occurring variants which permit straightforward quantitative analysis are analysed first, and then the remaining tail of the data, those variants which occur in very low frequencies, is investigated.

7.3.1 Long seat in main room of house

Four variants were elicited for the notion phrase *long seat in the main room of the house*, of which three were reported by more than one speaker. Examined in sections 4.4.1 and 4.4.2 in relation to the speaker variables of age and social class respectively, *long seat in main room of house* is investigated here to ascertain if

reported use of any of the three most frequently occurring variants, *couch*, *settee* and *sofa*, correlates with speaker identity. A fourth variant, *bench*, was elicited but is judged to be the result of a misunderstanding, and so is not further analysed. Figure 7.2 shows the most frequently reported variants for the notion phrase *long seat in main room of house* according to level of speaker identity.

Figure 7.2 Most frequently reported variants for the notion phrase *long seat in main room of house* according to level of speaker identity



Couch accounts for 12.50% of responses from those with low identity, dropping to 3.70% of responses from those with medium identity. It is most frequently reported by those with high identity, accounting for 18.42% of responses from this group. Differences in reported use between identity groups are not

statistically significant, however. The chi-squared value for *couch* according to speaker identity is 2.222 which does not exceed the upper 5% point for a chi-squared distribution with 2 degrees of freedom (5.991). Reported use of *settee*, by contrast, decreases with increasing level of speaker identity. It accounts for 62.50% of responses from those with low identity, compared with 48.15% of responses from those with medium identity and 42.11% of responses from speakers with high identity. Again, however, these differences in reported use are not statistically significant, as the chi-squared value for *settee* is 0.454. *Sofa* patterns differently again, being most frequently reported by those with medium identity. It accounts for 25.00% of responses given by speakers with low identity, 48.15% of responses given by those with medium identity and 39.47% of responses from those with high identity. These differences in reported use are not statistically significant either, as the chi-squared value for *sofa* is 0.624.

Though the least popular of the three variants, accounting for only nine out of the seventy-three tokens reported, *couch* is nonetheless of great interest. Of the three variants examined, it is the one which is more likely, on the evidence available, to be used by a speaker with high identity rather than one with medium or low identity. This is in direct contrast to *settee* and *sofa* which are more likely to be used by speakers with low and medium identity respectively. *Couch* is not frequently used, but where it is, it is reported by those with high local identity. This could help to explain the relative lack of age- and class-based variation in *couch* compared to *settee* and *sofa* seen in sections 4.4.1 and 4.4.2. It would appear that strength of local identity is a stronger factor in the use of this variant than speaker age or social class.

7.3.2 Grandmother and Grandfather

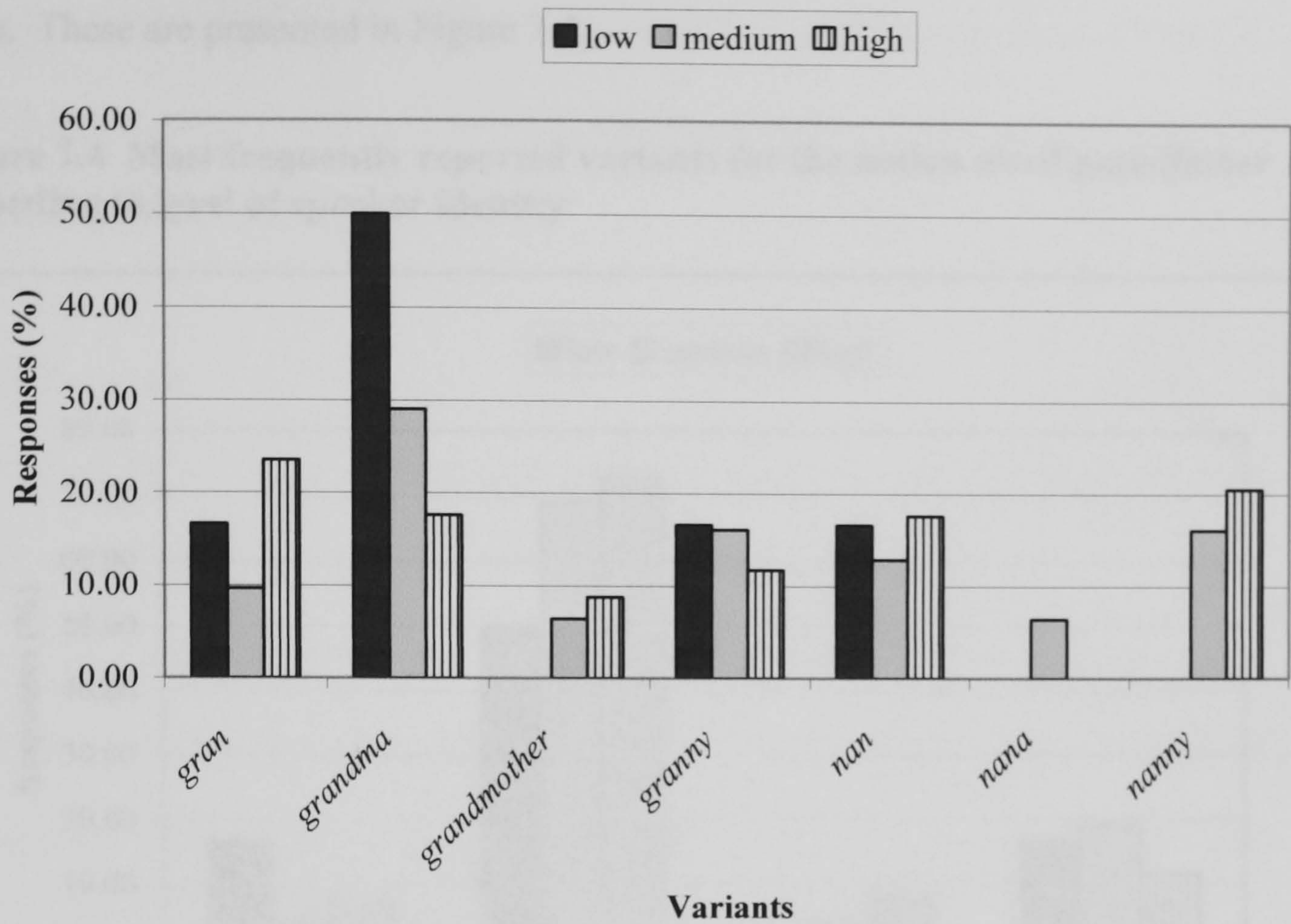
Seventeen different variants were elicited from the notion words *grandmother* and *grandfather*. Three of these (*aged people*, *grandparent*, *grandparents*) are not specific to either sex and so are discounted from analysis. Of the remaining fourteen variants, eight pertain to the notion word *grandmother* and six to the word *grandfather*.

7.3.2.1 Grandmother

In the case of *grandmother*, seven of the eight variants elicited were reported by more than one informant, and it is these seven which are presented according to level of speaker identity in Figure 7.3.

Four variants, *gran*, *grandmother*, *nan* and *nanny*, are reported most frequently by those informants with a high identity score. Two variants, *grandma* and *granny*, are reported more by informants with low identity than those with medium or high identity. Only *nana* is most frequently reported by those with medium or high identity than by those with low or high identity. In summary, *gran*, *grandmother*, *nan* and *nanny* are more likely to be used by people with a high level of Southampton area identity than are *nana*, *grandma* and *granny*. It should be noted, however, that none of the differences in reported use between identity groups of these lexical variants are statistically significant. The chi-squared values for these items are as follows: *gran* (1.620); *grandma* (1.612); *grandmother* (0.070); *granny* (0.189); *nan* (0.270); *nana* (1.355); and *nanny* (0.448). *Grandmum* is the only variant for *grandmother* reported once, and was used by an informant with medium identity. There are insufficient tokens of this word to draw any conclusions regarding its use in relation to speaker identity.

Figure 7.3 Most frequently reported variants for the notion word *grandmother* according to level of speaker identity



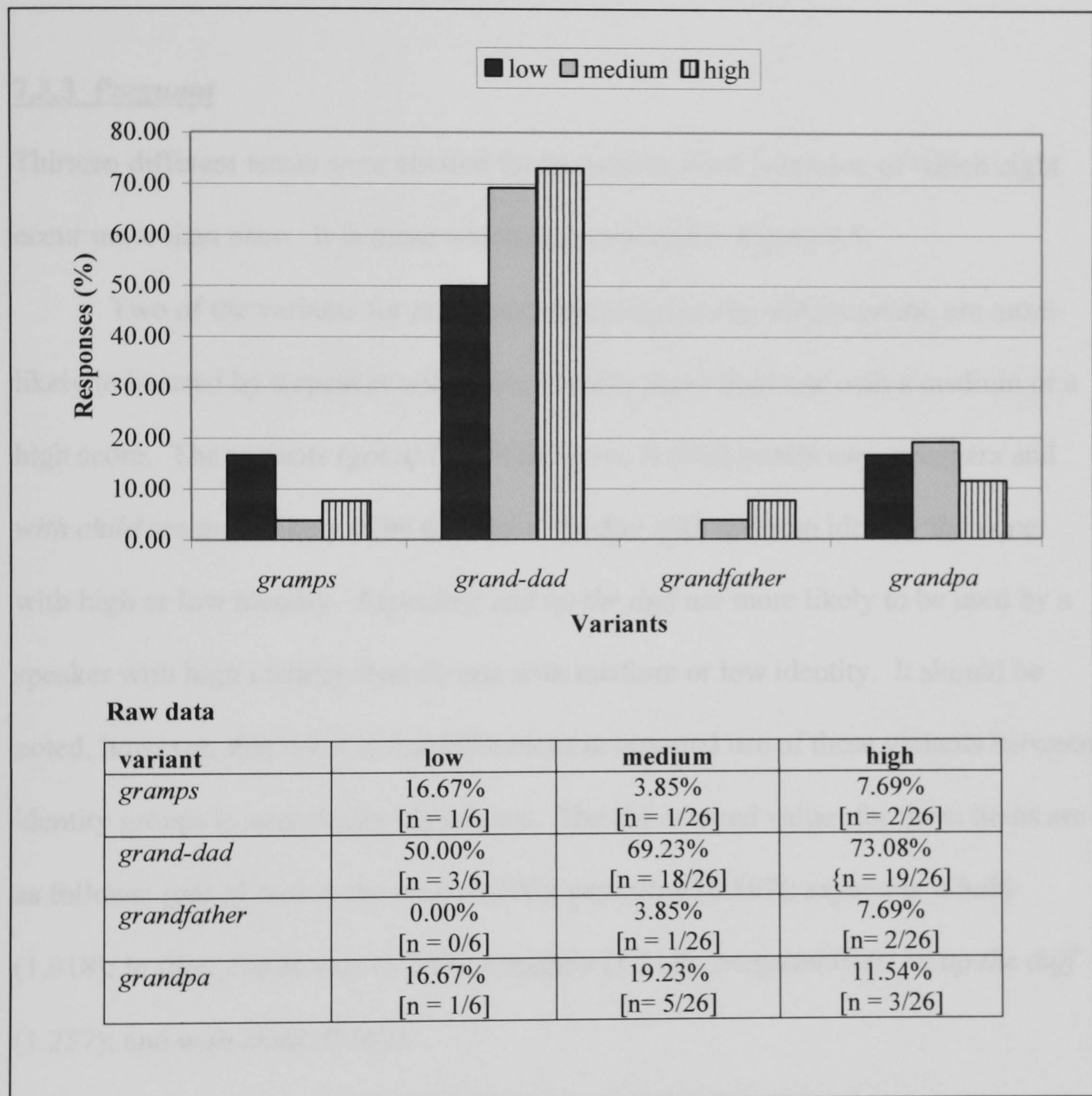
Raw data

variant	low	medium	high
<i>gran</i>	16.67% [n = 1/6]	9.68% [n = 3/31]	23.53% [n = 8/34]
<i>grandma</i>	50.00% [n = 3/6]	29.03% [n = 9/31]	17.65% [n = 6/34]
<i>grandmother</i>	0.00% [n = 0/6]	6.45% [n = 2/31]	8.82% [n = 3/34]
<i>granny</i>	16.67% [n = 1/6]	16.13% [n = 5/31]	11.76% [n = 4/34]
<i>nan</i>	16.67% [n = 1/6]	12.90% [n = 4/31]	17.65% [n = 6/34]
<i>nana</i>	0.00% [n = 0/6]	6.45% [n = 2/31]	0.00% [n = 0/34]
<i>nanny</i>	0.00% [n = 0/6]	16.13% [n = 5/31]	20.59% [n = 7/31]

7.3.2.2 Grandfather

In the case of *grandfather*, four of the six variants elicited were reported more than once. These are presented in Figure 7.4.

Figure 7.4 Most frequently reported variants for the notion word *grandfather* according to level of speaker identity



As in the case of *grandmother*, use of terms for *grandfather* appears to vary according to level of speaker identity. Both *grand-dad* and *grandfather* are reported most frequently by speakers with high identity. *Grandpa* is most frequently reported by speakers with medium identity and *gramps* most frequently reported by people with low identity. However, none of the differences in reported use of these variants

between identity groups is statistically significant. The chi-squared values for these items are as follows: *gramps* (0.122); *grand-dad* (0.409); *grandfather* (0.160); and *grandpa* (0.384). Two variants for *grandfather* are reported only once. *Gramp* is reported by a speaker with low identity and *grandpop* is reported by a speaker with medium identity.

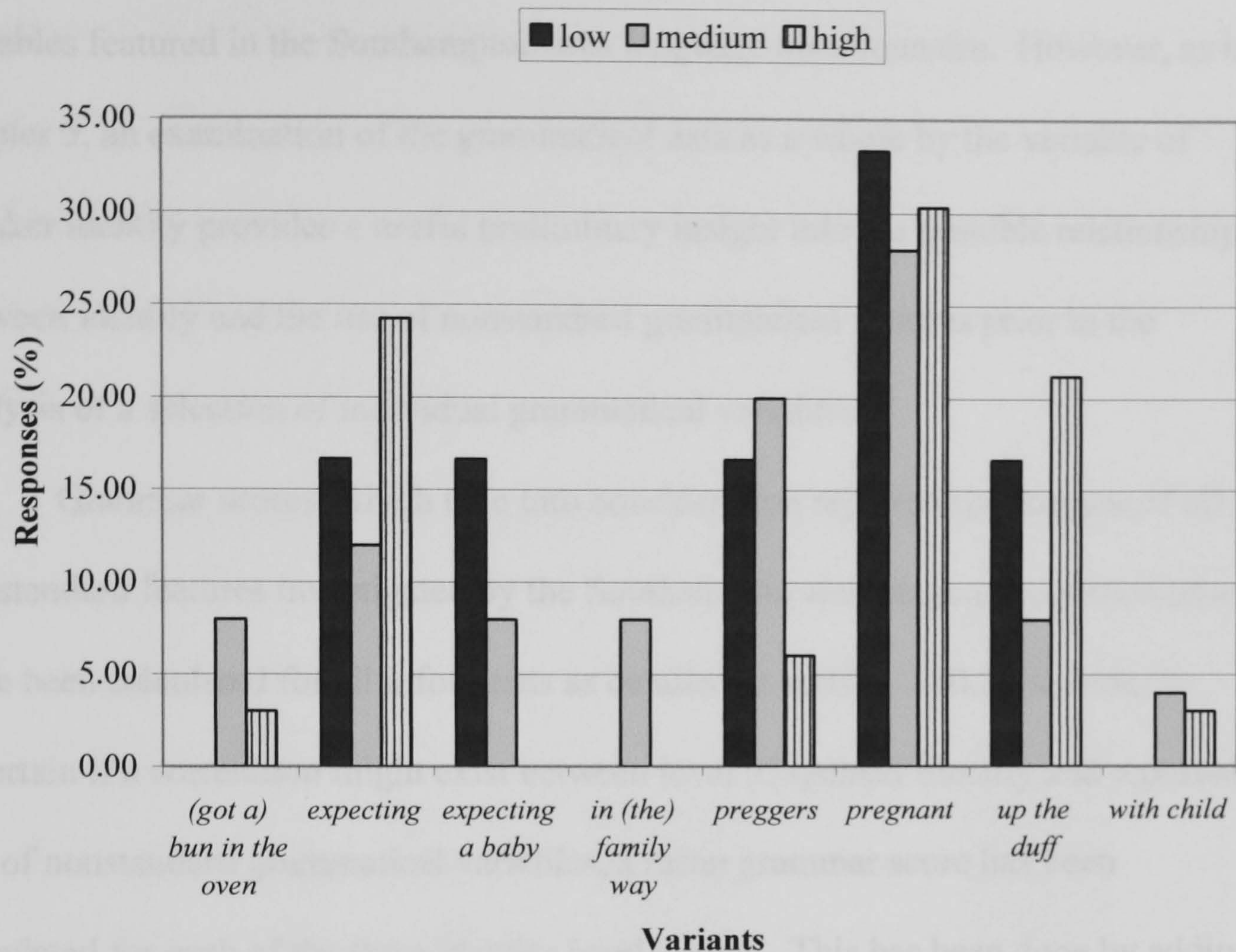
7.3.3 Pregnant

Thirteen different terms were elicited for the notion word *pregnant*, of which eight occur more than once. It is these which are presented in Figure 7.5.

Two of the variants for *pregnant*, *expecting a baby* and *pregnant*, are more likely to be used by a speaker with a low identity score than one with a medium or a high score. The variants *(got a) bun in the oven*, *in (the) family way*, *preggers* and *with child* are more likely to be used by a speaker with medium identity than one with high or low identity. *Expecting* and *up the duff* are more likely to be used by a speaker with high identity than by one with medium or low identity. It should be noted, however, that none of the differences in reported use of these variants between identity groups is statistically significant. The chi-squared values for these items are as follows: *(got a) bun in the oven* (0.276); *expecting* (0.867); *expecting a baby* (1.018); *in (the) family way* (1.503); *preggers* (1.518); *pregnant* (0.077); *up the duff* (1.257); and *with child* (0.862).

Five variants were reported only once. Of these, four were reported by speakers with high local identity (*bundle of joy*, *fall pregnant*, *having baby* and *up the gut*) and one by a speaker with medium identity (*in the club*). *Bundle of joy* is judged to be an anomalous variant, being used usually as a term for a baby rather than for pregnancy.

Figure 7.5 Most frequently reported variants for the notion word *pregnant* according to level of speaker identity



Raw data

variant	low	medium	high
<i>(got a) bun in the oven</i>	0.00% [n = 0/6]	8.00% [n = 2/25]	3.03% [n = 1/33]
<i>expecting</i>	16.67% [n = 1/6]	12.00% [n = 3/25]	24.24% [n = 8/33]
<i>expecting a baby</i>	16.67% [n = 1/6]	8.00% [n = 2/25]	0.00% [n = 0/33]
<i>in (the) family way</i>	0.00% [n = 0/6]	8.00% [n = 2/25]	0.00% [n = 0/33]
<i>preggers</i>	16.67% [n = 1/6]	20.00% [n = 5/25]	6.06% [n = 2/33]
<i>pregnant</i>	33.33% [n = 2/6]	28.00% [n = 7/25]	30.30% [n = 10/33]
<i>up the duff</i>	16.67% [n = 1/6]	8.00% [n = 2/25]	21.21% [n = 7/33]
<i>with child</i>	0.00% [n = 0/6]	4.00% [n = 1/25]	3.03% [n = 1/33]

7.4 Identity and Grammar

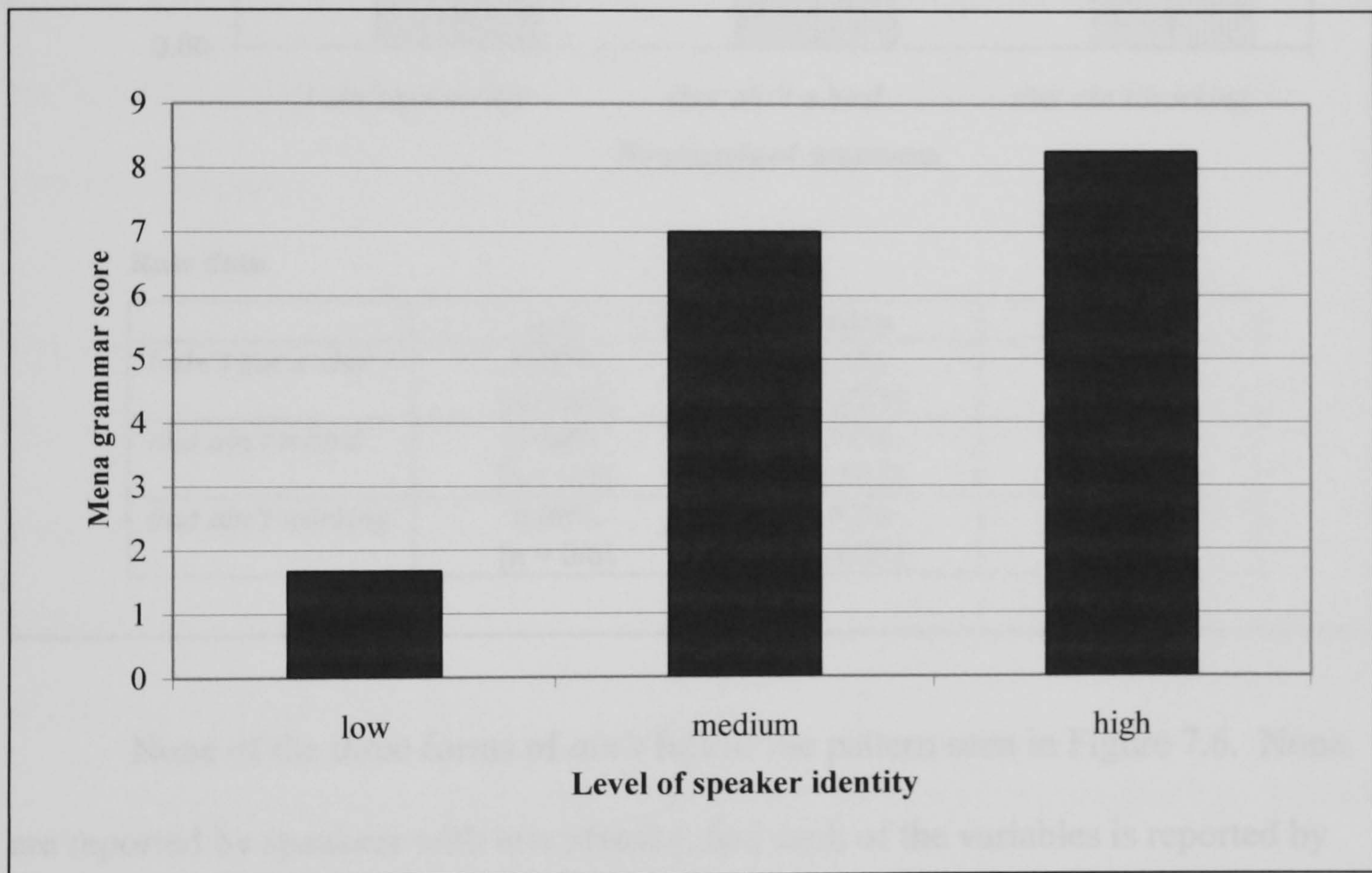
It is impracticable to analyse individually by speaker identity all of the grammatical variables featured in the Southampton area language questionnaire. However, as in chapter 5, an examination of the grammatical data as a whole by the variable of speaker identity provides a useful preliminary insight into the possible relationship between identity and the use of nonstandard grammatical features prior to the analysis of a selection of individual grammatical variables.

Grammar scores, which take into consideration reported speaker use of all the nonstandard features investigated by the Southampton area language questionnaire, have been calculated for all informants as detailed in section 3.10.2. In order to ascertain if a correlation might exist between level of speaker identity and reported use of nonstandard grammatical variables, a mean grammar score has been calculated for each of the three identity level groups. This has been done by adding together the grammar scores of all informants who fall into a particular identity level category and dividing this number by the total number of informants in that category. The results are shown in Figure 7.6. A high mean grammar score corresponds to high reported use of nonstandard grammatical variables, while a low mean grammar score corresponds to low reported use of nonstandard grammatical forms which, by extension, equates to a claim to use standard grammatical forms.

Figure 7.6 shows correlation between speaker identity and reported use of nonstandard grammatical variables. Reported use of nonstandard grammatical features increases as the level of speaker identity increases. Speakers with a low level of identification have a mean grammar score of 1.67. This mean score increases steeply to 7 for those with a medium level of identification, before peaking at 8.26 for those with a high level of identification. Given that the maximum mean

grammar score achievable is 45, as the Southampton area language questionnaire features 45 nonstandard grammatical features, these are not high mean scores. This is perhaps to be expected given the stigmatised nature of many of the features under examination. There is, however, a marked difference in the mean grammar scores of those with high and low identity. An examination of a selection of nonstandard grammatical variables will show whether this pattern is consistent for individual features.

Figure 7.6 Mean grammar scores according to level of speaker identity

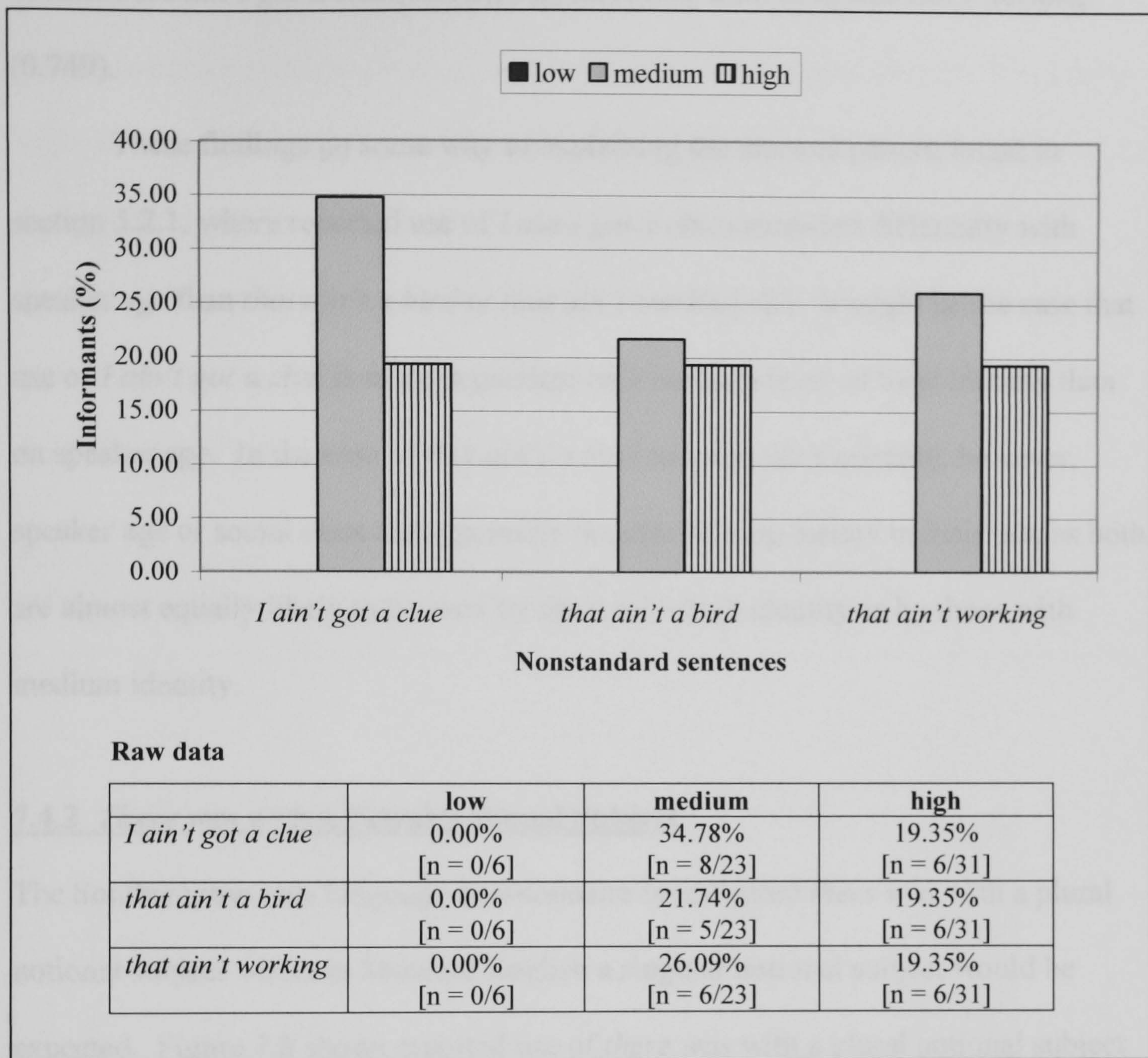


7.4.1 Ain't

Ain't is investigated in the Southampton area language questionnaire as a negative present tense contracted form of the verbs *be*, both as the copula (*that ain't a bird*) and the auxiliary (*that ain't working*), and *have*, as the auxiliary (*I ain't got a clue*).

Figure 7.7 shows reported use of the three forms of *ain't* according to level of speaker identity.

Figure 7.7 Reported use of *ain't* when talking to a friend according to level of speaker identity



None of the three forms of *ain't* follow the pattern seen in Figure 7.6. None are reported by speakers with low identity, and each of the variables is reported by 19.35% of speakers with high identity. They are, however, most frequently reported by informants with a medium level of identity. Of the three forms, however, it is *I ain't got a clue* which correlates most strongly with a medium identity score, being nearly twice as likely to be reported by a speaker with medium identity as one with high identity, compared to *that ain't a bird* and *that ain't working*, which are only marginally more likely to be reported by a speaker with medium identity than one with high identity. The differences in reported use of these variables between

identity groups are not statistically significant, however. The chi-squared values are as follows: *I ain't got a clue* (1.958); *that ain't a bird* (0.431); *that ain't working* (0.749).

These findings go some way to explaining the unusual pattern found in section 5.2.1, where reported use of *I ain't got a clue* correlated differently with speaker age than *that ain't a bird* or *that ain't working* did. It might be the case that use of *I ain't got a clue* is more dependent on a medium level of local identity than on speaker age. In the case of *that ain't a bird* and *that ain't working*, however, speaker age or social class are apparently the determining factors in their use, as both are almost equally likely to be used by those with high identity as by those with medium identity.

7.4.2 There was with a Plural Notional Subject

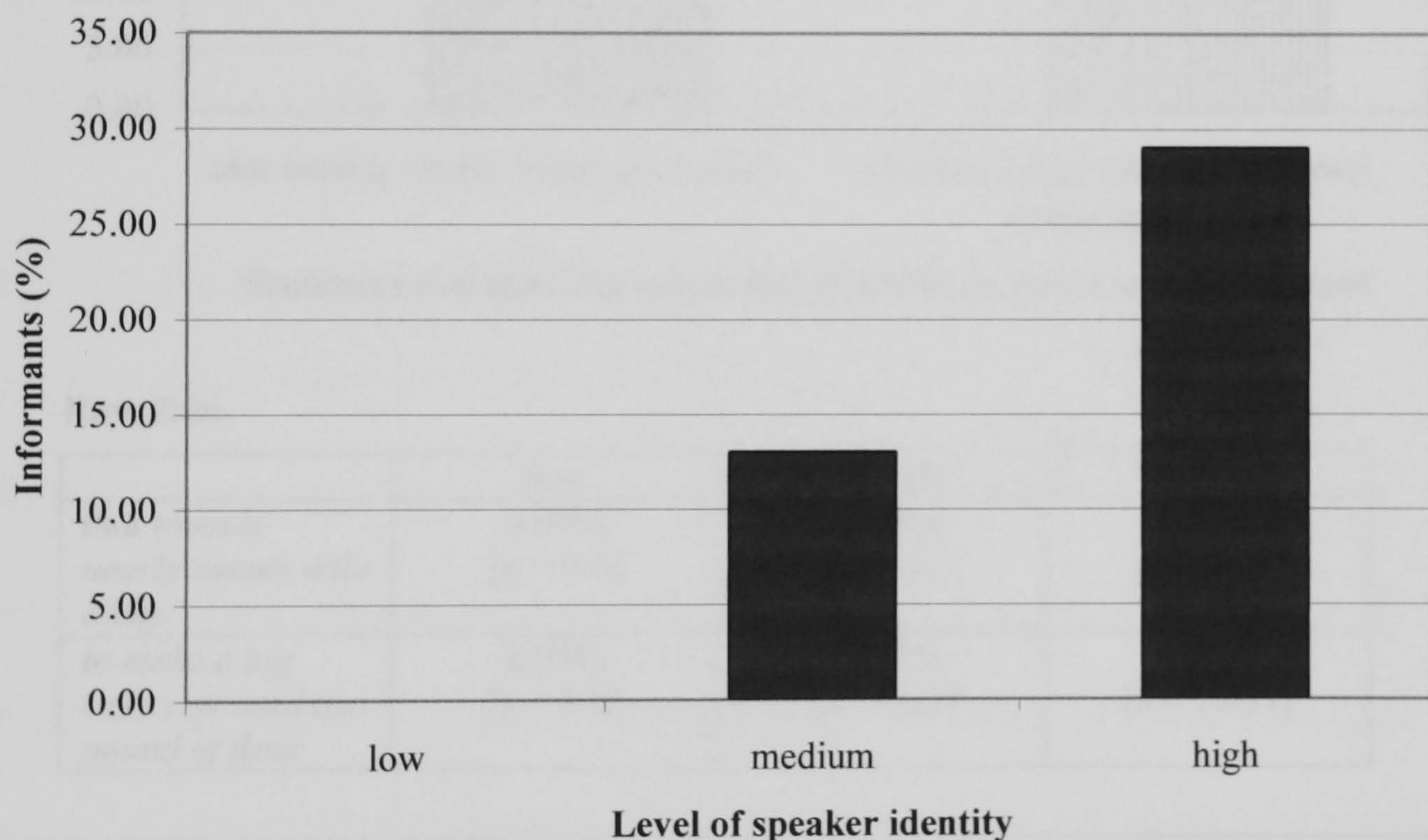
The Southampton area language questionnaire investigated *there was* with a plural notional subject where in Standard English a singular notional subject would be expected. Figure 7.8 shows reported use of *there was* with a plural notional subject according to speaker identity.

There was with a plural subject follows the pattern seen in Figure 7.6. Use of this variable increases with increasing levels of speaker identity. *There was some singers here a minute ago* is reported by 13.04% of speakers with medium identity and 29.03% of speakers with high identity. It is not reported by anyone with a low level of identity. As in the case of *ain't*, speakers who do not identify with the Southampton area do not report using *there was* with a plural notional subject. The higher the identity score, the more likely *there was* with a plural notional subject is to be reported, with more than double the percentage of informants with high identity reporting it compared with their counterparts with medium identity. This might go

some way to explaining the lack of sex- and class-based variation in reported use of *there was some singers here a minute ago* in sections 5.5.1 and 5.5.2 respectively.

On the evidence available, it appears to be the case that speaker identity is a greater influence than the speaker variables of sex and social class on the reported use of this feature. It should be noted, however, that differences in reported use between identity groups of *there was* are not statistically significant, as the chi-squared value of 1.906 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Figure 7.8 Reported use of *there was* with a plural subject when talking to a friend according to level of speaker identity



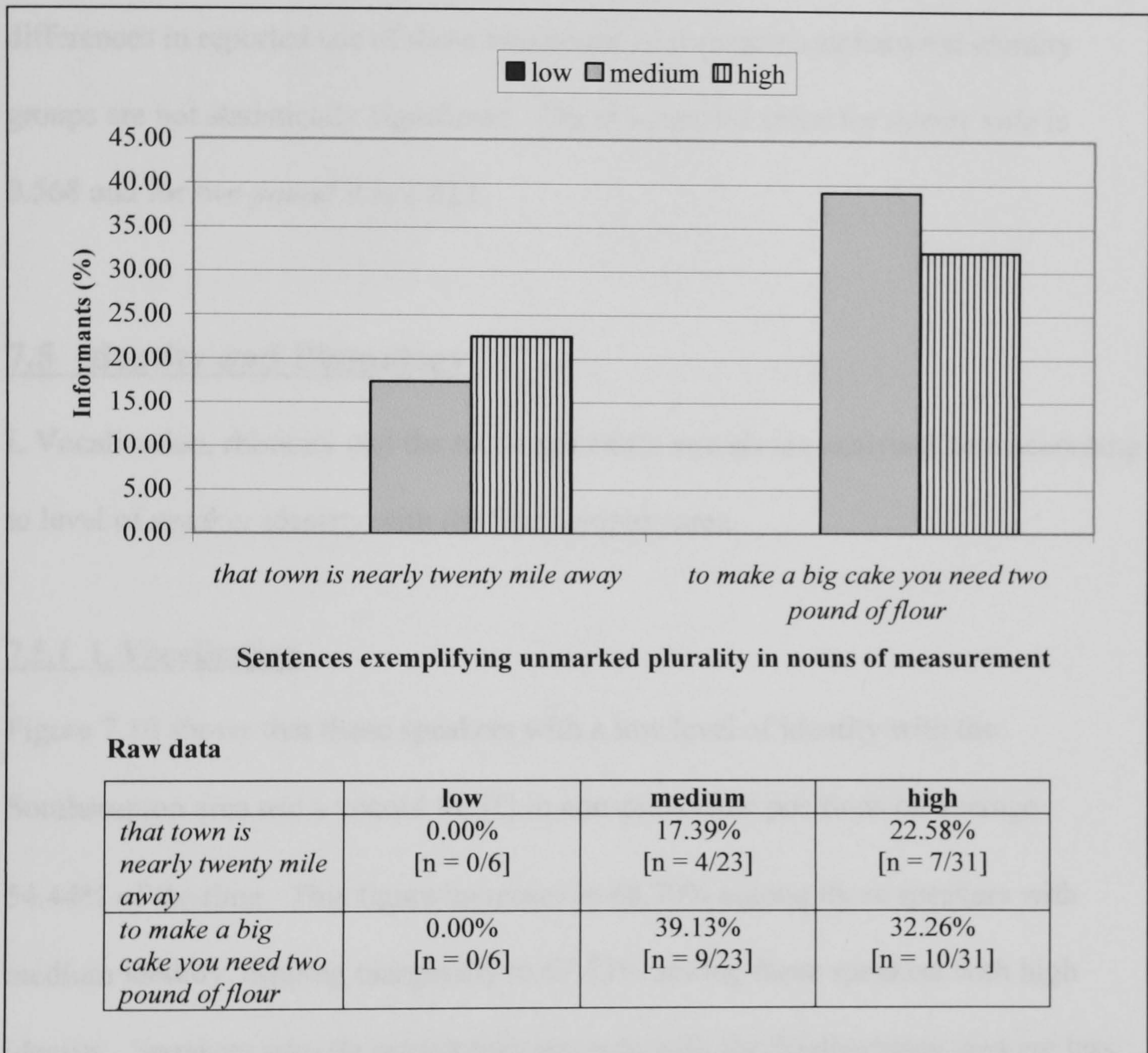
Raw data

low	medium	high
0.00%	13.04%	29.03%
[n = 0/6]	[n = 3/23]	[n = 9/31]

7.4.3 Nouns of Measurement with Zero Plurals

Figure 7.9 shows reported use of *twenty mile* and *two pound* according to level of speaker identity.

Figure 7.9 Reported use of nouns of measurement with zero plurals when talking to a friend according to level of speaker identity



As in the case of both *ain't* and *there was* with a plural subject, neither *twenty mile* nor *two pound* is reported by anyone with low identity. *Twenty mile* is reported by more speakers with high identity than those with medium identity, while the reverse is true of *two pound*. *That town is nearly twenty mile away* is reported by 17.39% of speakers with medium identity and 22.58% of speakers with high identity. *To make a big cake you need two pound of flour* is reported by 39.13% of speakers

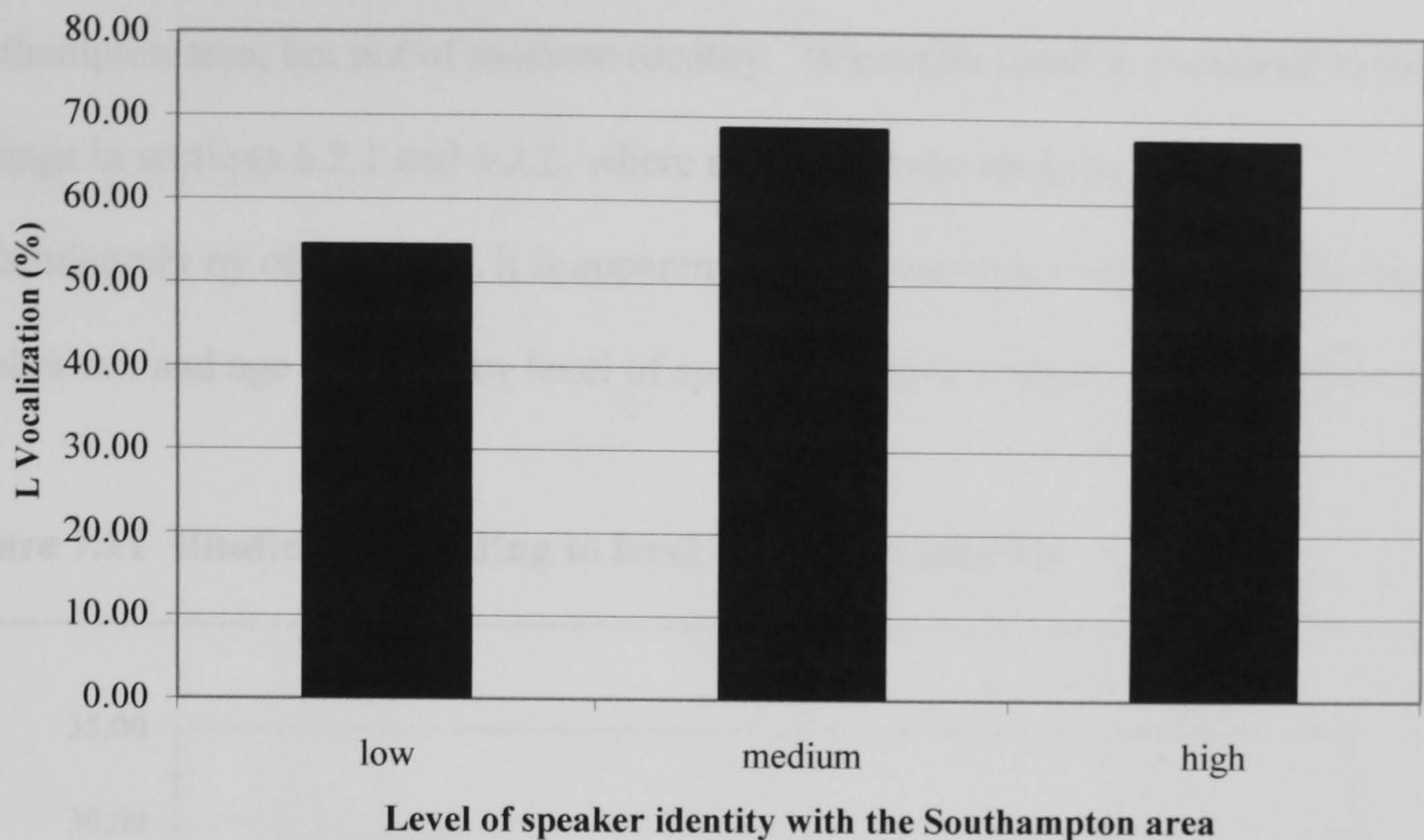
with medium identity and 32.26% of speakers with high identity. Though *two pound* is the more widely reported of the two variables, being reported by nineteen speakers while *twenty mile* is reported by eleven speakers, on the evidence available it is *twenty mile* which is more likely to be used by someone with high identity than by someone with medium or low identity. Again, however, it should be noted that differences in reported use of these two nouns of measurement between identity groups are not statistically significant. The chi-squared value for *twenty mile* is 0.568 and for *two pound* it is 1.822.

7.5 Identity and Phonology

L Vocalization, rhoticity and the BATH and PRICE vowels are analysed here according to level of speaker identity with the Southampton area.

7.5.1 L Vocalization

Figure 7.10 shows that those speakers with a low level of identity with the Southampton area use a vocoid for [ɫ] in non-prevocalic positions on average 54.44% of the time. This figure increases to 68.70% among those speakers with medium identity, dipping marginally to 67.53% among those speakers with high identity. Speakers who do not identify strongly with the Southampton area are less likely to use a vocoid for [ɫ] in non-prevocalic positions than are those who have a stronger sense of local identity. The chi-squared value for L Vocalization according to level of speaker identity is 6.844, which indicates a strong relationship between speaker identity and L Vocalization. As level of identity with the Southampton area increases, so too does L Vocalization. L Vocalization could therefore be argued to function as a marker of local identity in the Southampton area.

Figure 7.10 L Vocalization according to level of speaker identity**Raw data**

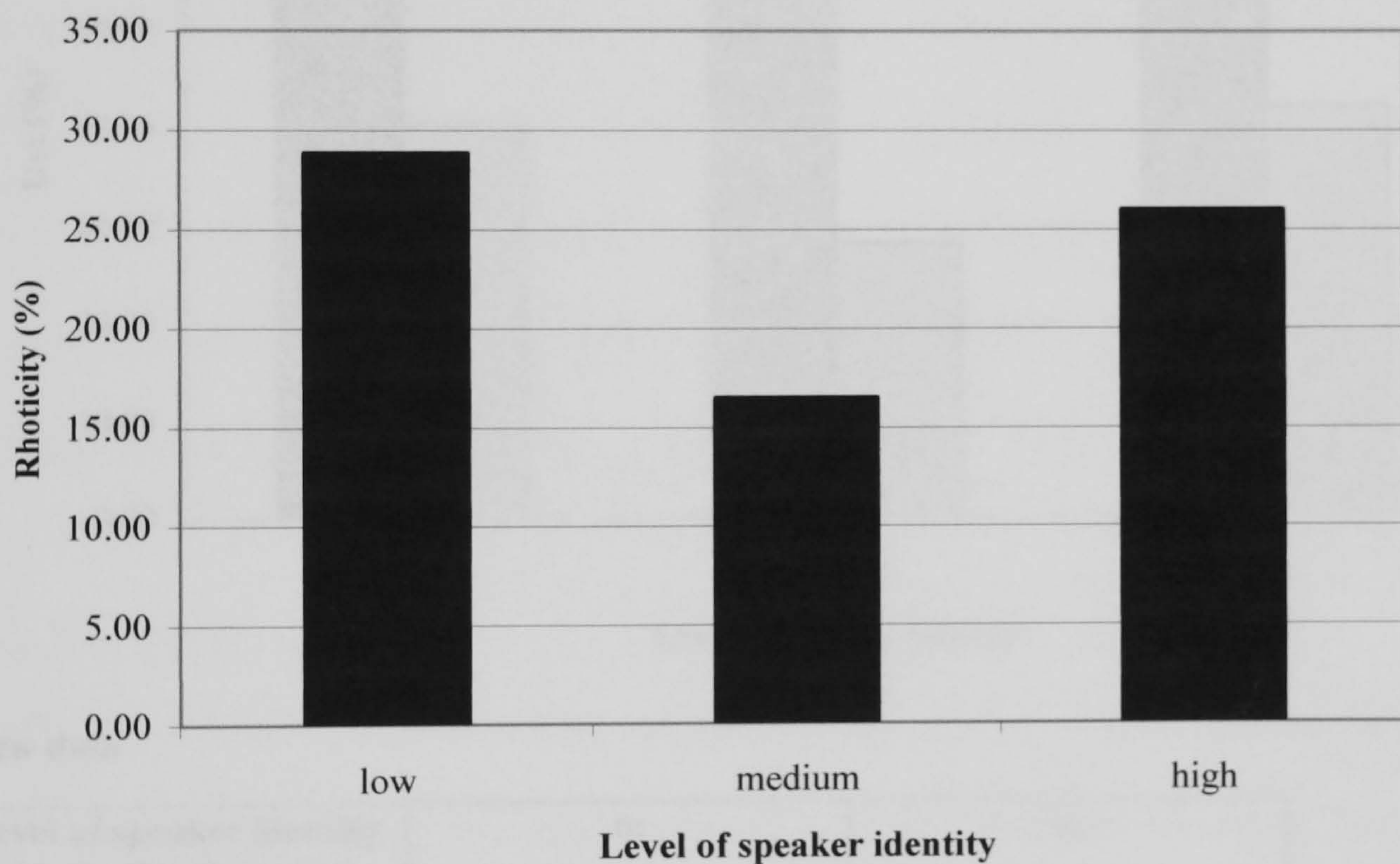
level of speaker identity with the Southampton area	L Vocalization
low	54.44% [n = 49/90]
medium	68.70% [n = 237/345]
high	67.53% [n = 314/465]

7.5.2 Rhoticity

The relationship between level of identity with the Southampton area and rhoticity is different to that between identity and L Vocalization. Figure 7.11 shows that rhoticity is highest among speakers with a low level of local identity, decreasing among those with medium identity before rising again among speakers with high identity. With a chi-squared value of 12.405, this difference in rhoticity among the identity groups is highly statistically significant. Rhoticity is most likely to be found in the speech of those with high or low identity with the Southampton area, and less

likely to be found in the speech of those with a medium level of identity. Rhoticity can be argued to function as a marker of both high and low identity with the Southampton area, but not of medium identity. When this result is compared to the findings in sections 6.2.1 and 6.2.2, where rhoticity is shown to be used predominantly by older males, it is apparent that this variable is more influenced by speaker sex and age than it is by level of speaker identity with the Southampton area.

Figure 7.11 Rhoticity according to level of speaker identity



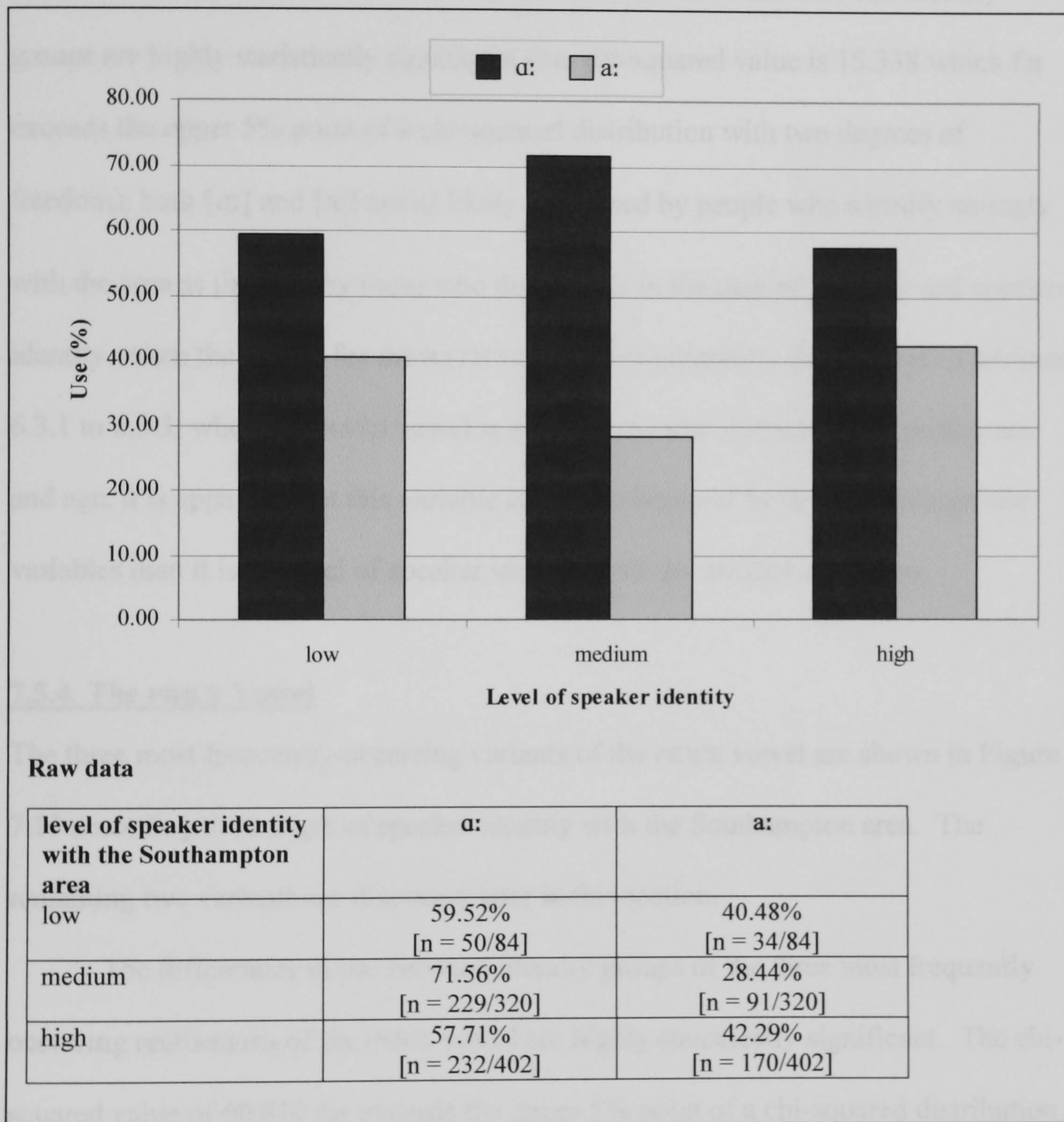
Raw data

level of speaker identity with the Southampton area	rhoticity
low	28.89% [n = 26/90]
medium	16.52% [n = 57/345]
high	26.02% [n = 121/465]

7.5.3 The BATH Vowel

Figure 7.12 shows realisations of the BATH vowel in relation to strength of speaker identity with the local area.

Figure 7.12 Realisations of the BATH vowel according to level of speaker identity



It is clear from Figure 7.12 that the [ɑ:] variant is the favoured realisation of the BATH vowel, regardless of level of speaker identity with the Southampton area.

This variant is most frequently used, however, by those who have a medium level of

identity. Conversely, [ɑ:] is least likely to be used by people with a medium level of identity. Use of the two variants by speakers with high identity with the Southampton area is very similar to that of speakers with low identity. In both groups, the backed form accounts for nearly 60% of tokens, while the fronted variant is used in around 40% of cases. Though the differences in use between identity groups are highly statistically significant (the chi-squared value is 15.338 which far exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom), both [ɑ:] and [a:] are as likely to be used by people who identify strongly with the area as they are by those who do not. As in the case of rhoticity and speaker identity, when the results for the BATH vowel are compared to the findings in sections 6.3.1 to 6.3.3, where the BATH vowel is shown to pattern strongly with speaker sex and age, it is apparent that this variable is more influenced by these extralinguistic variables than it is by level of speaker identity with the Southampton area.

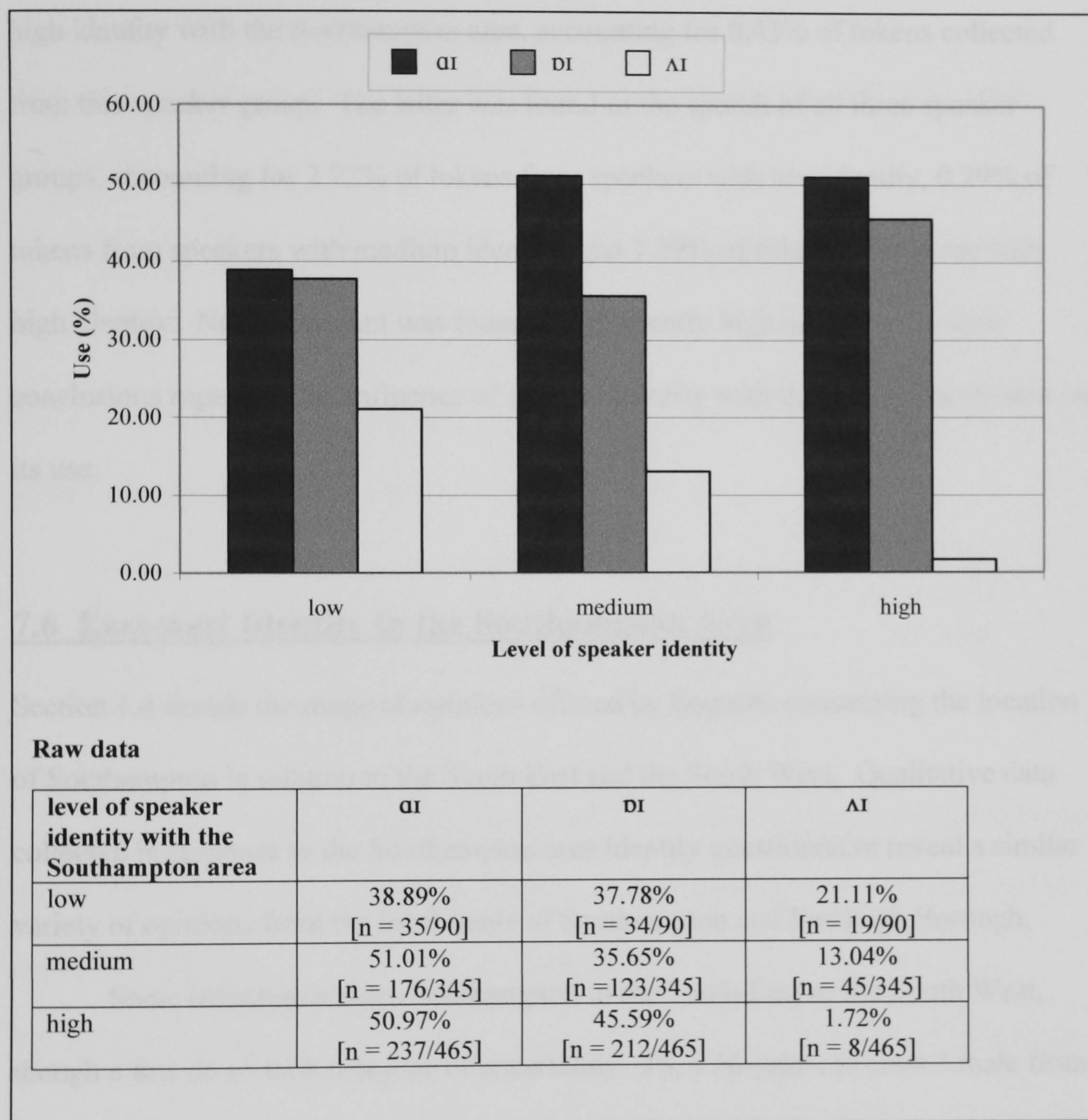
7.5.4 The PRICE Vowel

The three most frequently-occurring variants of the PRICE vowel are shown in Figure 7.13 according to strength of speaker identity with the Southampton area. The remaining two variants are discussed later in this section.

The differences in use between identity groups of the three most frequently occurring realisations of the PRICE vowel are highly statistically significant. The chi-squared value of 60.810 far exceeds the upper 5% point of a chi-squared distribution with four degrees of freedom (9.488). Though [ɑɪ] is the preferred variant for all three identity groups, it is more likely to be used by informants with medium or high levels of identity with the Southampton area than by those with low identity. As such, it can be argued to be a marker of local identity in the Southampton area. Use

of [ɒɪ] is highest among those with a high level of identity with the Southampton area, and can, therefore, also be argued to function as a marker of local identity.

Figure 7.13 Most frequently occurring realisations of the PRICE vowel according to level of speaker identity



Most frequently employed by those with low identity with the Southampton area, use of [ʌɪ] declines markedly as level of speaker identity with the Southampton area increases. It is clear that this variant is not a marker of strong identity with the

Southampton area, and suggests that the PRICE vowel is a significant non-RP feature in Southampton.

In addition to the three variants shown in Figure 7.13, tokens of [əɪ] and [ɔɪ] were also collected. The former was found only in the speech of informants with high identity with the Southampton area, accounting for 0.43% of tokens collected from this speaker group. The latter was found in the speech of all three speaker groups, accounting for 2.22% of tokens from speakers with low identity, 0.29% of tokens from speakers with medium identity, and 1.29% of tokens from those with high identity. Neither variant was found in sufficiently high quantities to draw conclusions regarding the influence of speaker identity with the Southampton area on its use.

7.6 East-west Identity in the Southampton Area

Section 1.4 details the range of opinions offered by linguists concerning the location of Southampton in relation to the South East and the South West. Qualitative data collected in response to the Southampton area identity questionnaire reveal a similar variety of opinions from the inhabitants of Southampton and Eastleigh Borough.

Some informants place Southampton in the South East or the South West, though a few do so with a degree of uncertainty. JS, a 56-year-old class 3 male from Eastleigh Borough says, 'I think we just come in the South East don't we round here'. Similarly, as discussed in 7.1.4, twenty-five informants out of sixty express feelings of closeness to people from the South East or the South West. However, there is apparent confusion expressed by some informants as to where Southampton is located geographically. RS, a 21-year-old class 1 male from Eastleigh Borough, says:

You see sometimes on the internet and stuff like that, they say [Southampton is in the] South East or West. I don't know really which one we're in. I thought we were bang in the middle. Central South.

GL, a 48-year-old class 1 male from Southampton states:

I'm never sure where Southampton lies. It's right on the border I would say because some maps it's classed as South West, other times it's South East. [...] I don't think there's any great affinity to either.

This is a sentiment echoed by JI, a 23-year-old class 1 female from Southampton who states:

I see Southampton as being quite unique really. If somebody said "Are you from the South West or the South East", and often you have to fill in those sorts of things on internet sites [...], I never know which one I'm on. I just see it as slap-bang southerners.

When asked to say where Southampton is geographically, ME, an 83-year-old class 2 female from Southampton, replies:

South. I'm a bit confused about that because sometimes we're put in the South West and sometimes we're put in the South East. We're basically bang in the middle [...]. Because I don't really feel aligned to the South West, although I think it's a lovely part of the country, because it is more rural basically. It's different altogether to the South East which is just South of London with all that [sic] problems.

Though informants appear to be unsure as to whether Southampton lies in the South East or the South West ('I don't know really which one we're in'), the reality seems to be that they do not believe the Southampton area to be in either. The Southampton area is 'central South', 'bang in the middle', its inhabitants 'slap-bang southerners' without 'any great affinity to either [the South East or the South West]'.
RW, a 69-year-old class 2 male from Eastleigh Borough argues that 'we live in the

South, not in the South East or the South West, because we are virtually in the middle'. These informants do not appear to identify with the South East or the South West. MG, a 22-year-old class 2 female from Eastleigh Borough, states, 'I don't really class myself as South East or South West [...] I'm southern'. When asked

why she does not feel close to people from either the East or the West. NW, a 19-year-old class 1 female from Southampton, states:

The West is viewed as more sort of country bumpkins [...] and the East is sort of very much... you just think of London [...]. It's just hard to [...] associate yourself with either East or West when I don't particularly geographically class myself as either.

DP, a 17-year-old class 1 male from Southampton, states of the Southampton area, 'I don't think we're as urbanised as London and I don't think we're as sparsely populated as [...] the West'.

Among some speakers, there appears to be a desire for the Southampton area and the wider central southern region to be recognised as separate from the South East and the South West. JS, a 56-year-old class 3 male from Eastleigh Borough states:

I think [...] that they [people from the Southampton area] would prefer to feel that they're central South, not to be put with one or the other [the South East or the South West].

This is a sentiment echoed by GH, a 60-year-old class 1 female from Southampton:

I think we're right in the middle here and they've put us in the South-Eastern area for the regions that they've made up [the Regional Assemblies (South East England Regional Assembly 2004)] but I think there needs to be a central region [...]. I think they ought to have had three regions in the South.

Llamas (2000: 123), discussing Berwick-upon-Tweed and Middlesbrough, comments that 'a 'border town' status [...] can entail a problematic construction of identity'. It is clear that this comment could well be applied to the Southampton area. Identity questionnaire results and comments offered by informants indicate that east-west identity in the Southampton area is a complex matter about which not even its inhabitants agree.

7.6.1 East-west Identity and Phonology

Le Page and Tabouret-Keller (1985: 18) argue that individuals adopt particular patterns of linguistic behaviour in order to resemble the group or groups with which they wish to be identified. Since such groups are often geographically based, one would expect speakers who identify with a particular area to use linguistic features typical of that area. Though one might speculate on the relationship between lexis and grammar and east-west identity, it is the relationship between phonology and this identity which is perhaps most productively examined, owing to the relative wealth of phonological data available for comparative purposes (Cheshire et al. 1989: 185; Johnson 1996: 83). L Vocalization, a feature traditionally associated with the South East (Altendorf and Watt 2004: 195-96), rhoticity, a feature traditionally associated with the South West (Altendorf and Watt 2004: 200-01), and the BATH vowel, a variable with clearly marked south-eastern and south-western realisations (Hughes et al. 2005: 69), are analysed according to speaker responses to question four of the Southampton area identity questionnaire (which asks speakers to say whether they feel closer to people from the South East, the South West or neither) in order to ascertain if there is a correlation between the employment of these variables and feelings of east-west identity.

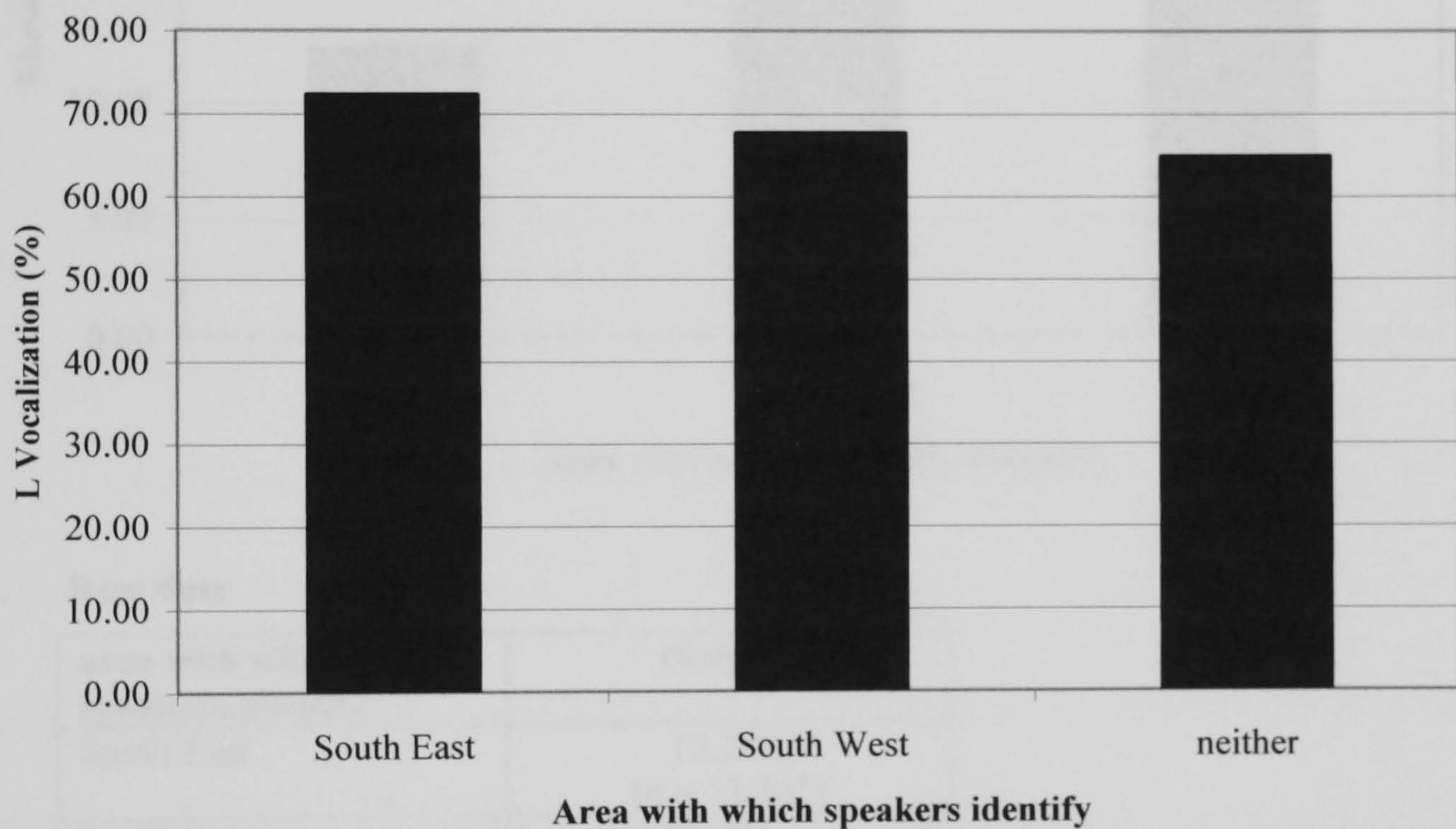
7.6.1.1 L Vocalization

Figure 7.14 shows L Vocalization according to speaker feelings of east-west identity.

As one might expect given that it is a feature typically associated with the South East (Altendorf and Watt 2004: 195-96), L Vocalization is greatest among those who claim to feel closer to people from that area. The speakers who claim to feel close to people from the East use a vocoid for [t] in non-prevocalic positions on average 72.38% of the time. Speakers who claim to feel closer to people from the

South West use a vocoid on average 67.78% of the time. Those who claim not to feel close to people from either the East or the West have the lowest level of L Vocalization, using a vocoid on average 64.95% of the time. It should be noted, however, that differences in L Vocalization between the identity groups are not statistically significant, as the chi-squared value of 2.387 does not exceed the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991).

Figure 7.14 L Vocalization according to speaker feelings of east-west identity



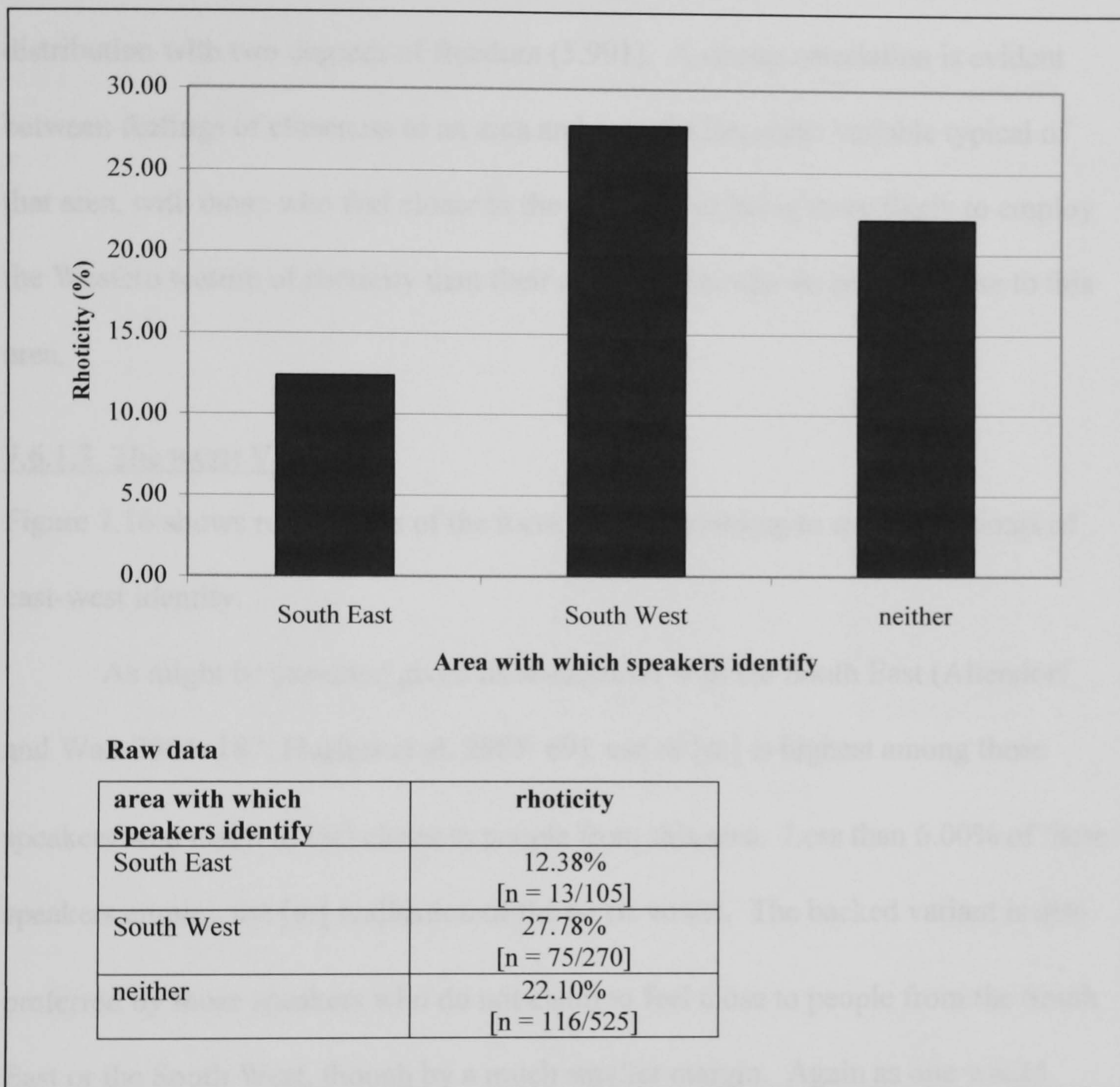
Raw data

area with which speakers identify	L Vocalization
South East	72.38% [n = 76/105]
South West	67.78% [n = 183/270]
neither	64.95% [n = 341/525]

7.6.1.2 Rhoticity

Figure 7.15 shows rhoticity according to speaker feelings of east-west identity.

Figure 7.15 Rhoticity according to speaker feelings of east-west identity



As one might expect given its association with the West (Altendorf and Watt 2004: 200-01), levels of rhoticity are highest for those who claim to feel closer to people from the South West, and lowest for those who claim to feel closer to people from the South East. Those who report feeling closer to people from the West are rhotic in word-final or preconsonantal position on average 27.78% of the time while informants who claim to feel closer to people from the East are rhotic on average

12.38% of the time. Informants who claim that they do not feel close to people from either the East or the West are rhotic on average 22.10% of the time. The differences in levels of rhoticity between identity groups are highly statistically significant. The chi-squared value of 10.459 far exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). A strong correlation is evident between feelings of closeness to an area and use of a linguistic variable typical of that area, with those who feel closer to the South West being more likely to employ the Western feature of rhoticity than their counterparts who do not feel close to this area.

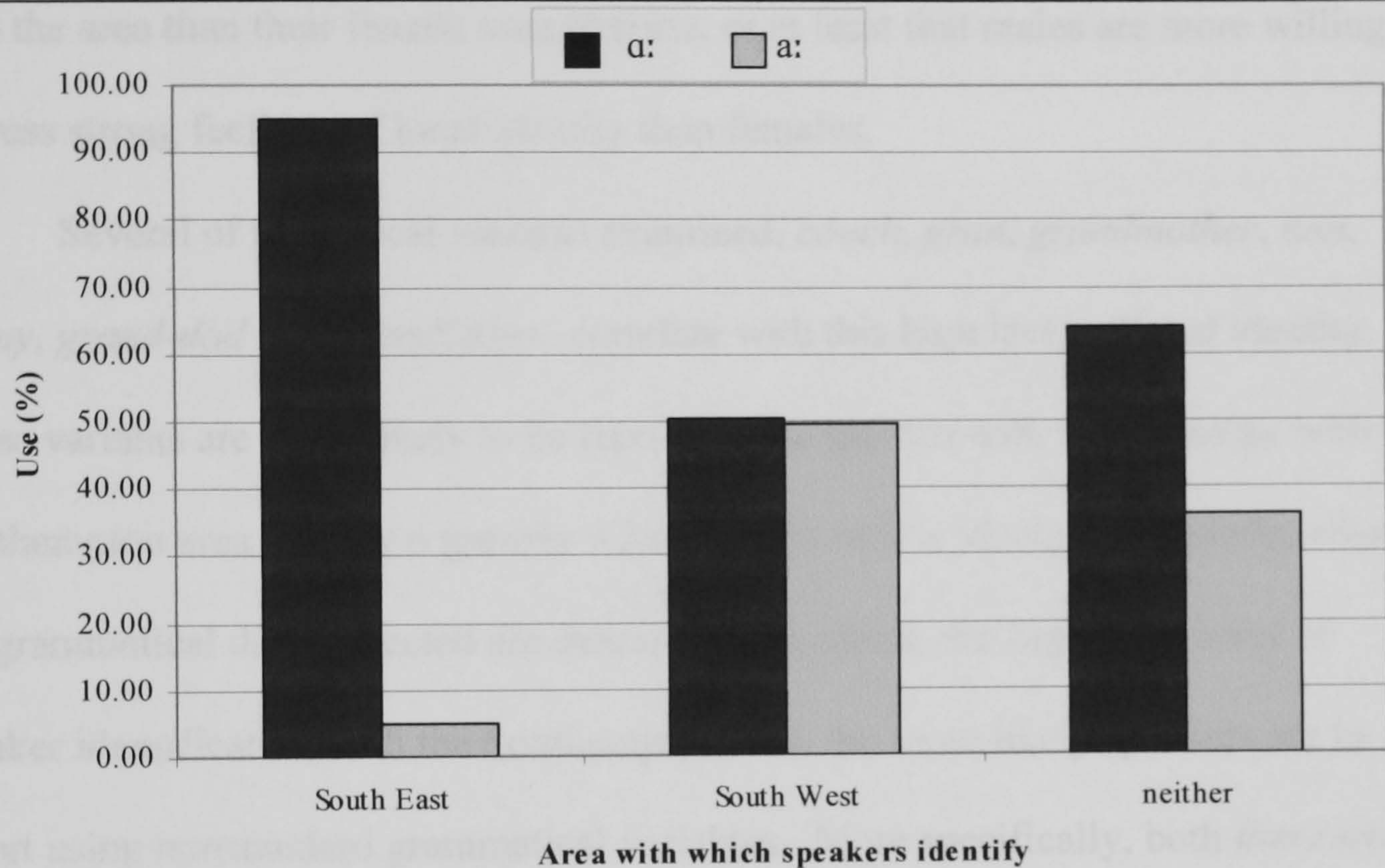
7.6.1.3 The BATH Vowel

Figure 7.16 shows realisations of the BATH vowel according to speaker feelings of east-west identity.

As might be expected given its association with the South East (Altendorf and Watt 2004: 187; Hughes et al. 2005: 69), use of [ɑ:] is highest among those speakers who claim to feel closer to people from this area. Less than 6.00% of these speakers employ the [a:] realisation of the BATH vowel. The backed variant is also preferred by those speakers who do not claim to feel close to people from the South East or the South West, though by a much smaller margin. Again as one would expect, use of [a:], a variant associated with the South West (Altendorf and Watt 2004: 199; Hughes et al. 2005: 69), is highest among those who claim to feel close to people from that area. The differences in realisations of the BATH vowel between the identity groups are remarkable, and are highly statistically significant. The chi-squared value of 106.380 far exceeds the upper 5% point of a chi-squared distribution with two degrees of freedom (5.991). So significant is this chi-squared value, that it also exceeds the upper 0.1% point of a chi-squared distribution with two

degrees of freedom (13.82). As Le Page and Tabouret-Keller's Acts of Identity Theory (1985:18) suggests will be the case, correlation between identity with an area and its people, and use of linguistic features typical of that area is apparent in the present study in the case of the BATH vowel.

Figure 7.16 Realisations of the BATH vowel according to speaker feelings of east-west identity



Raw data

area with which speakers identify	ɑ:	a:
South East	94.85% [n = 92/97]	5.15% [n = 5/97]
South West	50.58% [n = 130/257]	49.42% [n = 127/257]
neither	63.94% [n = 289/452]	36.06% [n = 163/452]

7.7 The Speaker Variable of Identity in the Southampton Area – A Summary

In general, responses to the Southampton area identity questionnaire show high levels of local identity. When responses to the identity questionnaire are quantified, thirty-one out of sixty informants have high identity scores compared to only six who have low identity scores. Results show that male informants identify more strongly with the area than their female counterparts, or at least that males are more willing to express strong feelings of local identity than females.

Several of the lexical variants examined, *couch*, *gran*, *grandmother*, *nan*, *nanny*, *grand-dad* and *grandfather*, correlate with this high level of local identity. These variants are more likely to be reported by a speaker with high identity with the Southampton area than by a speaker with medium or low identity. Similarly, when the grammatical data collected are examined as a whole, the higher the level of speaker identification with the Southampton area, the more likely speakers are to report using nonstandard grammatical variables. More specifically, both *there was* with a plural notional subject and *twenty mile* correlate with high speaker identity. Though rhoticity and the BATH vowel do not correlate with high identity with the Southampton area, L Vocalization and the [ɑɪ] and [ɒɪ] realisations of the PRICE vowel do. It could be argued that the lexical, grammatical and phonological features mentioned serve as markers of Southampton area identity. At the very least, they might be said to be the favoured variants in terms of expressing being from the Southampton area.

There are several lexical and phonological features which correlate with low feelings of local identity. Interestingly, none of the grammatical variables investigated (*ain't*, *there was* with a plural subject and unmarked plurality in nouns of measurement) correlate with low feelings of local identity; indeed, no informants

with low identity report using these forms. Features that do correlate with low local identity are: *settee*; *grandma*; *granny*; *gramps*; *expecting a baby*; *pregnant*, rhoticity and the [ʌɪ] realisation of the PRICE vowel. Use of these features does not appear to be a means of expressing feelings of local identity.

The majority of speakers do not report feeling close to people from the South East or the South West. Instead, many argue that they do not believe the Southampton area to be part of the East or the West. They consequently find it hard to identify with people from either of these areas. Where a claim to feel closer to one or other of these regions is made, east-west identity appears to correlate with speaker age, with younger informants claiming to feel closer to people from the South East and speakers in the middle and oldest age groups being more likely to identify with people from the South West. L Vocalization is, as might be expected, more frequent among speakers who identify with the South East and rhoticity more frequent among those who identify with the South West. The [a:] realisation of the BATH vowel is most frequently found in the speech of those who identify with the South West, while the [ɑ:] realisation is favoured by a very large margin by those who associate with the South East. When examined according to speaker feelings of east-west identity, L Vocalization, rhoticity and the BATH vowel would appear to support Tabouret-Keller's assertion that linguistic items are the way in which individuals create their own identity and identify with others (Le Page and Tabouret-Keller 1985: 5).

Chapter 8

8.0 Conclusion

This chapter provides an overview of the main findings of the Southampton area study. It critiques the Survey of Regional English methodology, and provides suggestions for further work.

8.1 A Critique of the Adapted SuRE Methodology

Every means of data elicitation has its advantages and its disadvantages. The SuRE methodology is no exception. In sections 8.1.1 and 8.1.2, these advantages and disadvantages are discussed with specific reference to the lexical and grammatical data that the adapted SuRE methodology has been employed to collect.

8.1.1 Lexical Data

The lexical data collected allow us to critique the adapted SuRE methodology; to highlight the ways in which it is an effective means of data elicitation but also to identify ways in which it might be improved. Relatively little is written in variationist literature about lexis in comparison to that written about the subjects of phonology and grammar (Johnson 1996: 83). The relative lack of interest in lexis is not due to the fact that it is 'unworthy' of study, but that the collection of lexical data is problematical and time-consuming, and, once collected, the resultant lexis is not always readily analysed. As an example of the former, the SED questionnaire was so long that fieldworkers sometimes started the questioning with one informant in a given locality before moving on to another informant to finish it in order to save speakers' time (Orton 1962: 16-17). Of the elicitation of material for her work on lexical variation in the South Eastern United States, Johnson states that:

Sixty-one items, especially adjectives, that were included in the original plan proved in three trial interviews to be difficult to elicit without breaking up the flow of the questioning by focusing on words rather than content or by using an awkward ‘‘fill the blank’’ type of strategy. Even without these items, the interviews often averaged three hours in length, and the volunteer informants often ran out of patience by the end of the sessions.

(Johnson 1996: 6-7)

By comparison, the SuRE SRNs are a means of quickly eliciting appreciable amounts of lexical data. The volume and range of lexis collected indicates that the SuRE Sense Relation Networks were very successful in eliciting lexical variation in the Southampton area. Most informants were quick to grasp the concept behind the SRNs, and many provided more than one variant per notion word.

However, some responses offered by informants indicate that there is a disparity between the response the notion word or phrase had been intended to elicit and that which was offered by the informant. For example, it was the intention that the notion phrase *types of bread* should elicit variants for the Standard English words *roll* and *loaf*. Instead, quite reasonably, informants often responded with varieties of bread, for example *granary* and *white*. Though speakers were questioned about the notion phrase in the subsequent interview, and the words *roll* and *loaf* and their variants elicited, in hindsight the notion phrase as it appeared on the SRN was not completely clear. Similarly, the phrase *non-alcoholic drinks* was included to elicit generic terms for carbonated and still drinks, for example *pop* and *squash*. Though it was to an extent successful in doing this, it also elicited specific varieties of non-alcoholic drinks such as *water*, *tea* and *barley water* and soft drink brand names. The notion word *sweets* collected terms such as *sweeties* but also *pudding*, as it was not clear which was the intended referent. (It had been included to elicit words for confectionery rather than desserts.) In both cases, the notion phrase could perhaps have been more specific, or the expectations regarding the margin for ‘error’ inherent

in a methodology such as the SuRE SRNs, completed without influence from the fieldworker and therefore open to interpretation, might have been more realistic.

Another matter the lexical data raise is that of speaker self-reporting. When speakers are asked to self-report their language use, there can be a disparity between what they report they say and what they actually say, with speakers often claiming to use a prestige form when they actually use a nonstandard variant, or vice-versa (Newbrook 1999; Labov 2001: 194). The only way to establish whether speakers are accurately reporting their own use is, as Johnson (1996: 35) notes, to analyse an extended recording of an individual's speech. Since the completion of the SRNs was followed by an interview, it is possible, to a certain extent, to compare a speaker's reported use of particular lexical variants with their actual use. This cannot be taken far, however, given the relatively short length of the interview.

In many respects, speaker self-reporting is a considerable strength of the SuRE methodology, since it allows large quantities of comparable lexical data to be collected without either the time-consuming questioning employed by studies such as the SED, or the element of chance and likely lack of comparability that are inherent in the analysis of unstructured conversation. Johnson (1996: 81) questions whether it might be the case that lexical items are less subject to stigmatisation than grammar or phonology. To the extent that this is true, the problems associated with self-reporting of lexical use might not be as serious as for grammar or phonology, since speakers would have less cause falsely to report their use for reasons of overt or covert prestige. At the same time, Johnson (1996: 81) does, however, acknowledge the need for further research on lexical stigmatisation. Provided one is aware of the possibility of disparity between reported and actual lexical use, there is no reason

why the SuRE SRNs should not continue to be regarded as a highly successful means of data elicitation.

8.1.2 Grammatical Data

When completing the Southampton area language questionnaire, it was assumed that informants would select either the first option (that the informant would use the sentence themselves when talking to a friend) or the second (that they would not use the sentence themselves but have heard it used by others in the local area), since there is what was believed to be a clear distinction in these responses between the reporting of use or non-use of a particular construction. Alternatively, they could leave both boxes blank, indicating that neither option was relevant to them.

However, a number of informants ticked both option one and option two, making analysis of their data difficult. One could assume that, in ticking both boxes, informants are reporting both their own use of the construction and the use of the sentence by others in the local area. Speakers might have ticked one of the boxes in error, however, so it could be misleading to make such assumptions, particularly when the two statements contradict each other in part. These so-called 'double responses', given by six informants are discounted from analysis. Twenty-three sentences are affected in total, seven of these having two pairs of double responses and the remaining sixteen sentences one pair each. Altogether, 60 out of 5760 responses have been discounted, around 1% of the total number of responses.

With hindsight, ease of use of the grammar questionnaire for informants might have been improved by a slight rewording of one of the sentence responses. The options *I'd use this when I was talking to a friend* and *I'd use this when I was writing to a friend* would have remained the same in order to continue to collect information concerning the linguistic spheres in which informants consider the use of

a particular construction acceptable, the written word normally being more formal than spoken language. However, rather than the response *I wouldn't use this but some people do use it in Southampton/Eastleigh Borough*, the revised option might have instead read *I hear this used in Southampton/Eastleigh Borough*. In this way, there is perhaps a clearer distinction between the individual responses. At present, the second option (*I wouldn't use this but some people do use it in Southampton/Eastleigh Borough*) contains two statements and appears to be the source of some confusion. This proposed revision might allow speakers to report instances when they believe their own use of a particular sentence to be outside the norms of their speech community by ticking the *I'd use this when I was talking to a friend* option but not the box to report hearing the use of a construction locally. How likely this particular combination of responses would be is unclear though. Despite this minor problem, the Southampton area language questionnaire was nonetheless a success. The majority of informants completed the grammar questionnaire without difficulty and it proved a quick and effective means of eliciting comparable data.

The claim for the efficacy of the language questionnaire as a means of data elicitation is further supported when the control sentences featured in the questionnaire are examined. The inclusion of these sentences was designed to ensure that informants took the time to read each construction in turn and to assess the form on its own merit, rather than making assumptions about the sentences in the questionnaire as a whole and reducing the reliability of the resultant data by answering in the same way for every sentence if this was not appropriate. If we look at each of the control sentences in turn, we can see that the majority of informants in each case reported their own use of these forms. *She likes toffees* is reported by 58.33% of informants, *she laughs really loudly* by 76.67% and *I've got four dozen*

eggs by 75%. With an average reported use of 70%, the control sentences are a useful and successful means by which to gauge the accuracy with which speakers have completed the language questionnaire.

A further issue arising from an analysis of the grammatical data elicited is the matter of speaker self-reporting, a subject which is of particular interest when discussing the collection of grammatical data. As Johnson (1996: 81) suggests, it could be the case that lexis is inherently less subject to stigmatisation than grammar and phonology, since words are rarely stigmatised based on an association with the speech of lower status social groups. The same cannot be said, however, of nonstandard grammatical features, which are often heavily and overtly stigmatised. In many cases, speakers not only distanced themselves from the nonstandard features by reporting them on their questionnaires under the heading *I wouldn't use this but some people do use it in Southampton/Eastleigh Borough* but, when asked what type of people they heard using the forms, they suggested the sentences were used by what they perceived to be low status members of the speech community, such as those with little education or those who lived in less desirable parts of the area. In some cases, respondents then proceeded to employ that particular variable later in the interview. The stigma attached to these nonstandard forms is apparently so great that they did not, either consciously or subconsciously, accurately report their own use.

This potential for disparity between reported and actual use of both nonstandard and standard forms must be borne in mind when analysing data collected using the speaker self-reporting technique. While the usual trend in reporting the use of nonstandard or stigmatised forms is to under-report (claim to use, but not actually use, high prestige features, in this case, Standard English forms), some informants over-report their use of these features (claim to use, but not actually

use, low-prestige forms) (see, for example, Labov 1966: 455; Newbrook 1999: 100). For certain people, there is covert prestige in the use of nonstandard or stigmatised forms, be they grammatical, phonological or lexical. Trudgill states:

Women use linguistic forms associated with the prestige standard more frequently than men. One reason for this is that working-class speech has favourable connotations for male speakers. Favourable attitudes to non-standard speech are not normally expressed, however, and emerge only in inaccurate self-evaluation test responses.

(Trudgill 1972: 179)

Whether the issue is under- or over-reporting, one must retain an awareness of the potential pitfalls of relying on data elicited by speaker self-reporting.

In order to ascertain the accuracy of a speaker's self-reporting, their questionnaire responses can in part be compared to their actual use of nonstandard forms in the interview which follows the completion of the SuRE pack. However, to listen to an informant's speech as a means of verifying, or indeed contradicting, their questionnaire responses is not without its flaws. The first issue of relevance in this regard is the observer's paradox, a problem faced by every fieldworker aiming to collect vernacular linguistic data. Given the stigmatised nature of many of the nonstandard grammatical features under examination, and the relatively formal nature of the interview context, it is unlikely that informants will speak exactly as they would in everyday life. Instead, their speech might tend towards the formal end of the style spectrum. A second issue concerns the lack of comparability and the time-consuming nature of free conversation as a method of grammatical data collection. It would not be possible to verify the reliability of each informant's questionnaire responses by comparing these with their speech in the SuRE interview, since it is highly unlikely that they will have employed constructions comparable to the forty-eight sentences featured in the Southampton area language questionnaire during the course of the interview.

That said, the comparison undertaken in the present study of reported and actual speaker use of the present tense verb morphology and adverbs is of great interest (see sections 5.6.1 and 5.6.2, respectively). These variables were selected as they occurred in high frequencies in the SuRE interviews, making a comparison of reported and actual use possible. Accuracy of speaker reporting of personal use of these forms is generally very high, supporting the use of a grammar questionnaire as an effective means to collect large quantities of comparable data in a short space of time. In cases where speaker self-reporting is not as accurate, important issues are raised regarding the best way in which to collect grammatical data. Similarly, levels of under- and over-reporting are very telling with regard to the information they provide us about speaker attitudes to the variables under examination.

Speaker self-reporting can be an efficient method of data elicitation, capable of producing very telling results. It has certainly been used to great effect by, among others, Cheshire et al. (1989) in the Survey of British Dialect Grammar, which provides information about the use and geographical distribution of a wide range of grammatical features (see also Johnson 1996 on lexical self-reporting, and Newbrook 1999 on phonological self-reporting). The use of a structured questionnaire allows the speedy collection of large quantities of comparable data that it might not otherwise have been possible to collect given the time-consuming and problematical nature of grammatical data elicitation outlined in 3.5.2. It could be argued that no methodology is without its flaws and that, provided one is aware of the potential pitfalls associated with the use of speaker self-reporting, there is no reason why it should not be successfully employed to collect grammatical data, particularly since this material is so difficult to elicit using other methodologies.

8.2 The Speaker Variables of Sex, Age and Social Class

Analysis of lexical, grammatical and phonological data shows language use in the Southampton area to differ according to the speaker variables of sex, age and social class. As expected in light of the findings of many of the social dialectological studies which have preceded it (Labov 1966; Trudgill 1974), the current study shows that, in general, women tend to use standard or prestige forms. Not all variables analysed were subject to linguistic sex differentiation, however. This is of interest, since it suggests that certain features might function to a greater extent than others as markers of male and female identity.

The findings of the present study also support previous research into linguistic age differentiation (Trudgill 1974; Macaulay 1977) which shows that it is often the youngest informants who most frequently employ or report employing stigmatised variants, and those in the middle age group who, in general, do so least frequently. In this respect, the Southampton area is again shown to be a conventional speech community. Where variants have not patterned as expected, it is possible to speculate that certain features function more strongly as markers of membership of a particular age group than others.

Several changes in the use of lexical, grammatical and phonological forms appear to be in progress. The Southampton area study is an apparent time one. Data from differing age groups have been compared and used to infer temporal developments. Where possible changes in progress have been identified, whether a variant is suspected of being on the increase or on the decline, real-time study would be needed to ascertain for certain whether this is actually the case and not a result of age-grading. In some cases, data from the SED are available which allow a real-time comparison to be made.

Class-based variation is also evident in the Southampton area data. Such class differences provide us with information regarding the status of particular forms in the Southampton area, as linguistic features, arguably grammatical and phonological variables more so than lexical ones, are often stigmatised as a result of their association with lower status social groups.

8.3 The Speaker Variable of Identity

The present study examines the issue of identity from two perspectives. The first of these is the identity of informants with the Southampton area itself. The second is the identity of those same speakers with the South East and the South West of England. As predicted in chapter 1, the matter of identity has proved to be of great significance in this survey. The majority of speakers reported strong feelings of identity with the Southampton area, and very few speakers had low levels of local identity. These differing levels of local identity in turn correlate to varying extents with use of particular lexical, grammatical and phonological features. Certain linguistic features appear to function as stronger markers of Southampton area identity than others.

Informants for the present study have, in general, very definite ideas about the relationship between the Southampton area and the South East and South West of England. Qualitative data collected show that the majority of those interviewed do not believe the area to be in either the East or the West, and informants appear to resent the attempts of others outside the area to assign the Southampton area to one of these regions.

That is not to say, however, that many of the informants do not identify in some way with people from the South East and the South West. Where informants

do report feeling close to people from the East or the West, these feelings are reflected in their language. L Vocalization, a feature traditionally associated with the East, is found more frequently in the speech of informants who claim to feel close to people from the East than by those who do not report feeling close to people from that area. Similarly, rhoticity and the [a:] realisation of the BATH vowel, features traditionally associated with the South West, are found more frequently in the speech of informants who claim to feel close to people from the West than by those who do not report feeling close to people from that area. The Southampton area study supports the assertion that individuals create their patterns of linguistic behaviour in order to identify themselves with particular groups (Le Page and Tabouret-Keller 1985: 18).

The Southampton area accent contains a mixture of traditionally south-eastern and traditionally south-western forms. In many ways, it is difficult to describe a typical Southampton accent, as some speakers sound south-eastern, others south-western. It is perhaps this mixture of forms which are associated with one or other of these areas that characterises the Southampton area. Speaker comments expressing confusion regarding the geographical position of the Southampton area (see section 7.6) could be argued to suggest that Southampton is somehow lacking its own identity, its inhabitants unsure as to where they belong. However, in many respects this liminality is in fact key to Southampton's identity. The Southampton area inhabitants' sense of identity appears to be borne out of not wishing to be labelled as south-eastern or as south-western, but rather to be acknowledged as Southern. The mixture of south-eastern and south-western forms in the Southampton area accent is used to express this Southernness.

The changes which appear to be in progress with regard to rhoticity and the BATH vowel, however, suggest a move away from phonological features traditionally associated with the South West. Whether this is a move towards the South East, however, is a different matter. Though L Vocalization, a traditionally south-eastern form, appears to be on the increase in the Southampton area, it has, however, been present in Hampshire for many years, being found in the SED Incidental Material. The issue of identity in the Southampton area is undoubtedly a complex one, but one which is nonetheless of great significance.

8.4 Suggestions for Further Work

It has been impracticable to examine by the speaker variables of sex, age, social class and identity all of the linguistic material collected. The findings of the present study suggest, however, that there would be merit in further analysing these data, particularly given that so little research has been done on the speech of the Southampton area. A description of the accent of the area might be usefully constructed using data collected from the Southampton area word list, supplemented by phonological data from the interviews.

8.5 Closing Comments

The Southampton area study is a pioneering one in which lexical, grammatical and phonological data have been examined according to the speaker variables of sex, age, social class and identity, in a region previously neglected by dialectologists. There is scope for further investigation of the data collected, but analysis already undertaken shows clear differences in the use of particular linguistic items according to the sex, age, social class and feelings of identity of the speakers interviewed.

Appendix 1 – The original core Survey of Regional English methodology

BIOGRAPHICAL INFORMATION

name

sex

age (dob).....

place of birth.....

birth place of mother.....

birth place of father

birth place of grandmothers

birth place of grandfathers

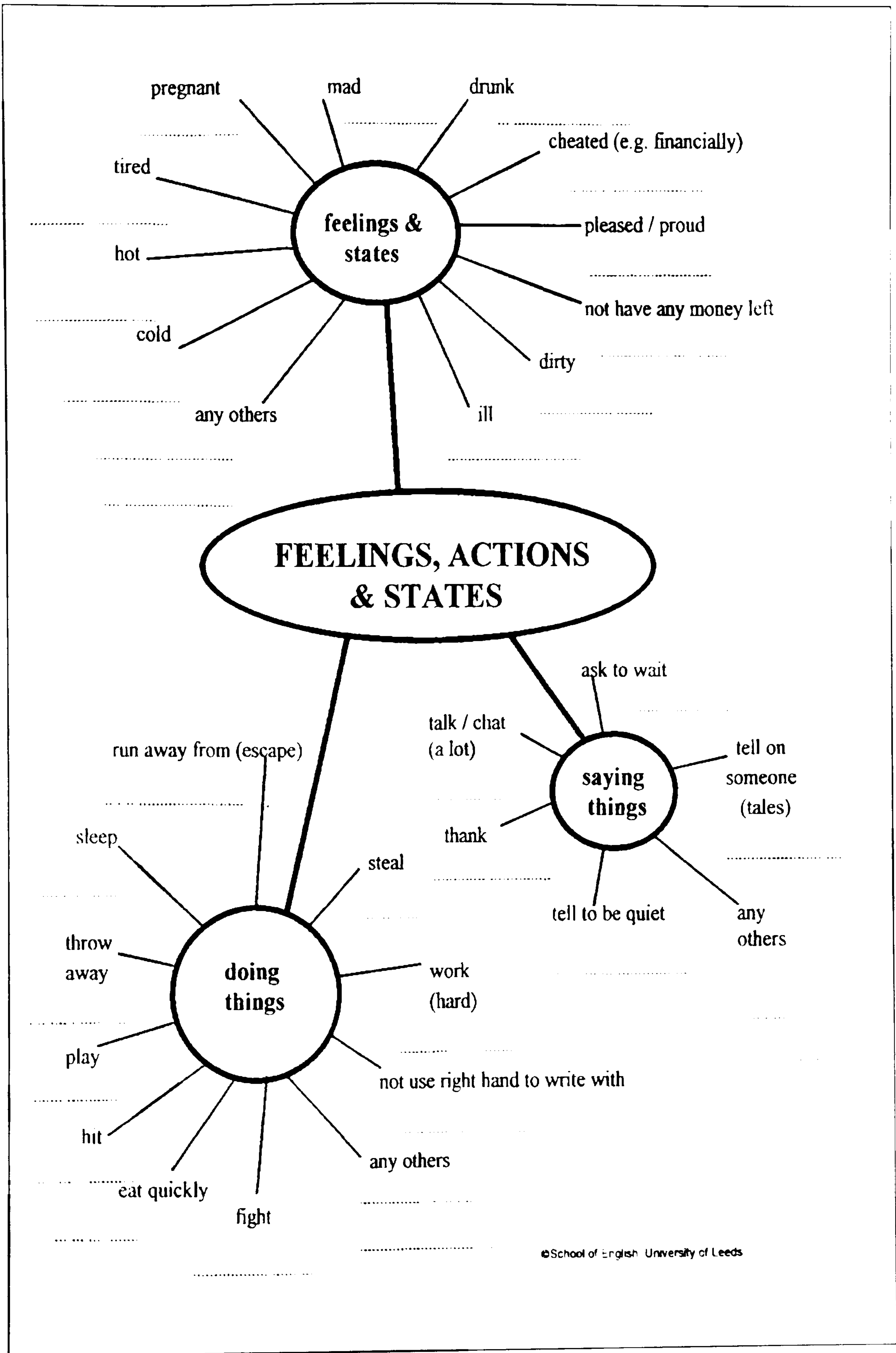
ethnic group.....

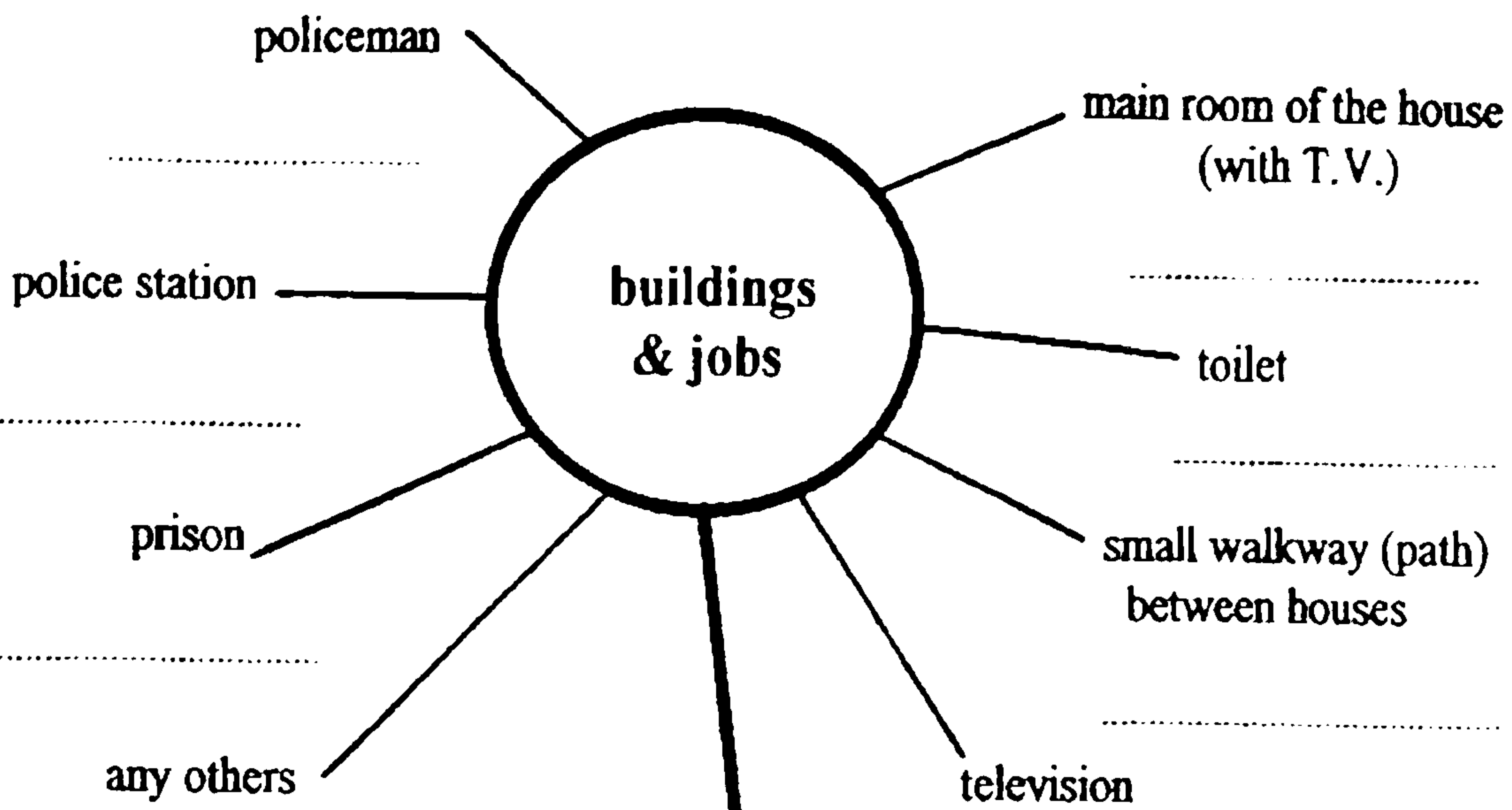
occupation (current / usual).....

assessment of social class

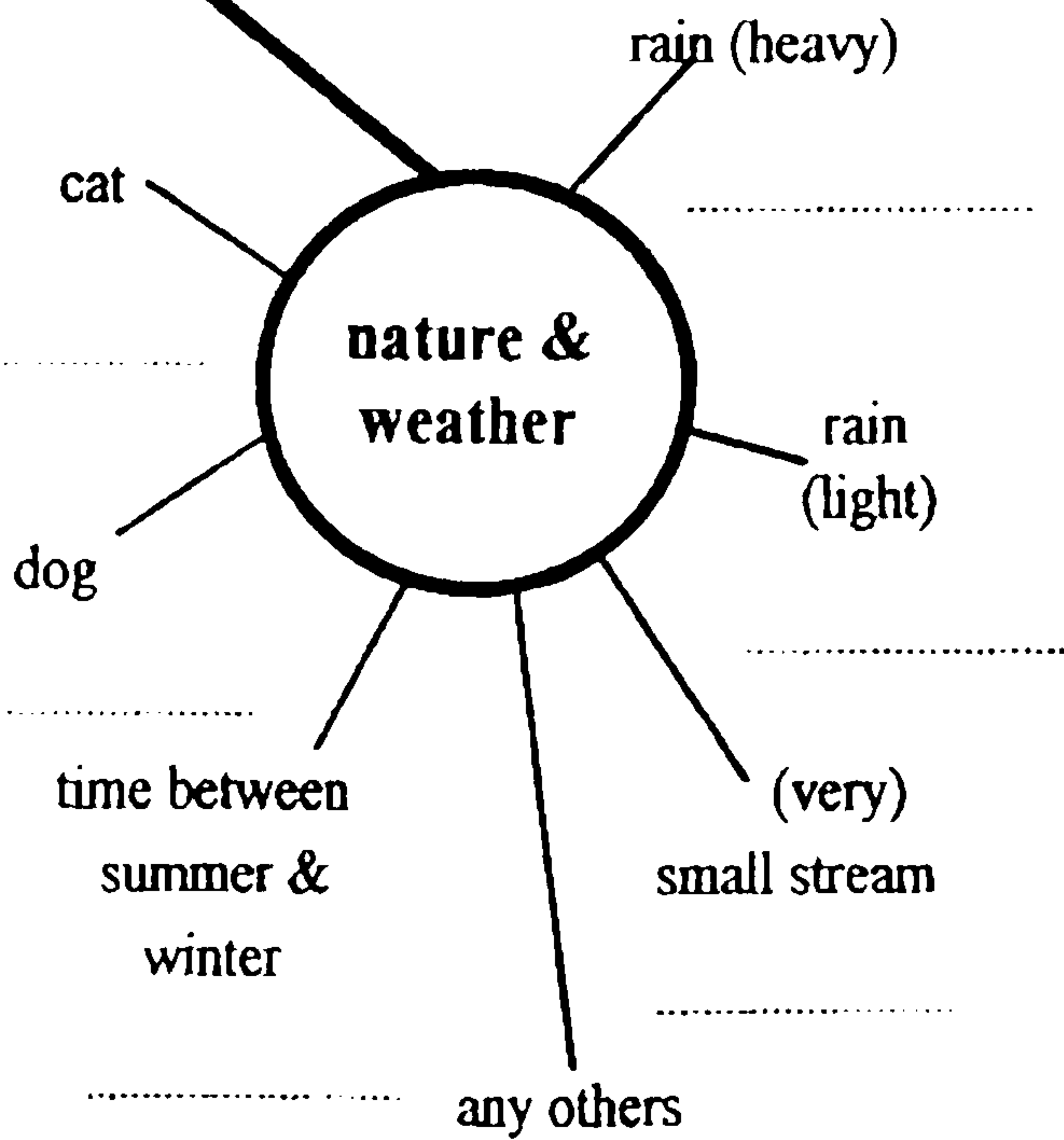
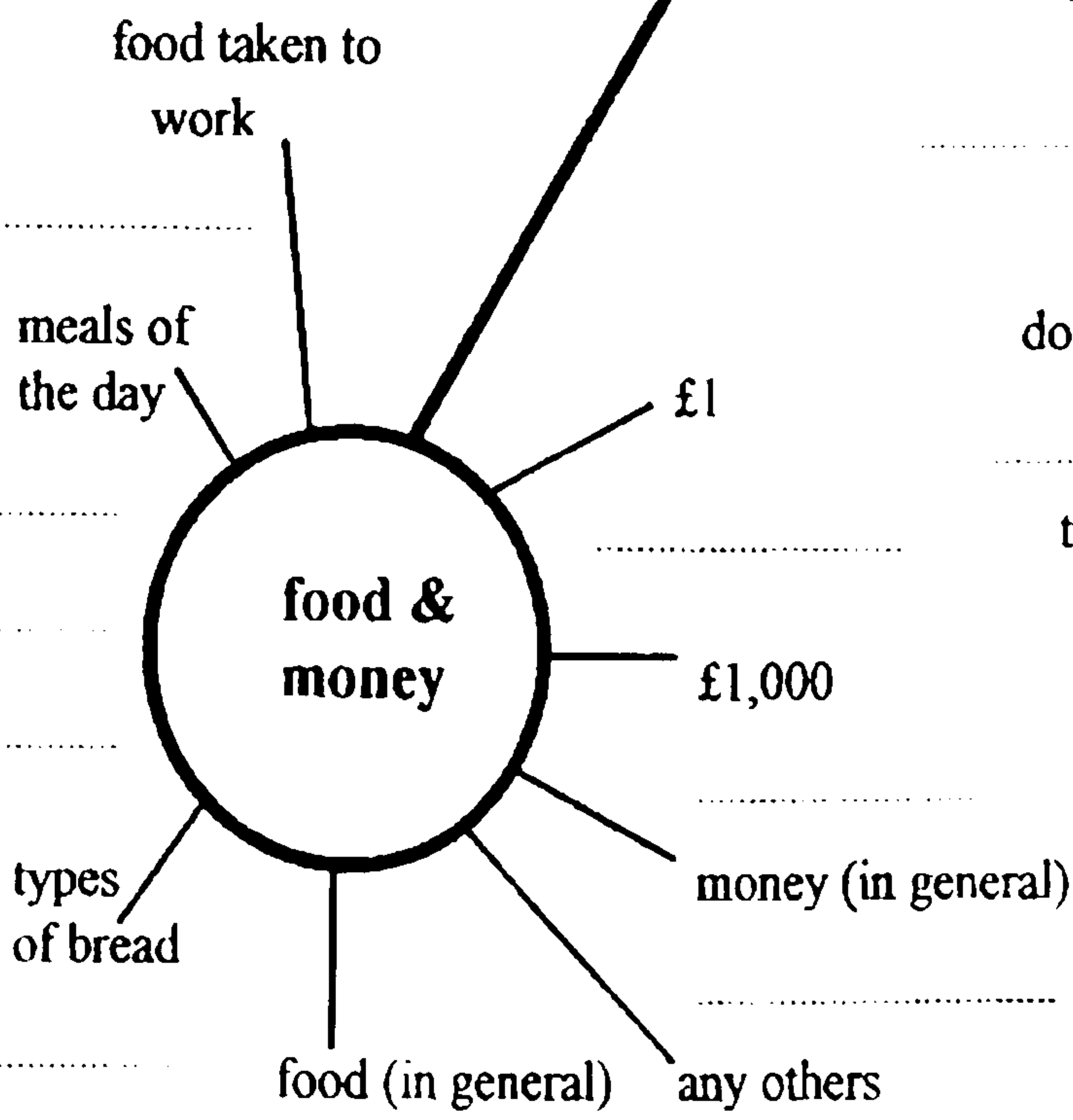
housing.....

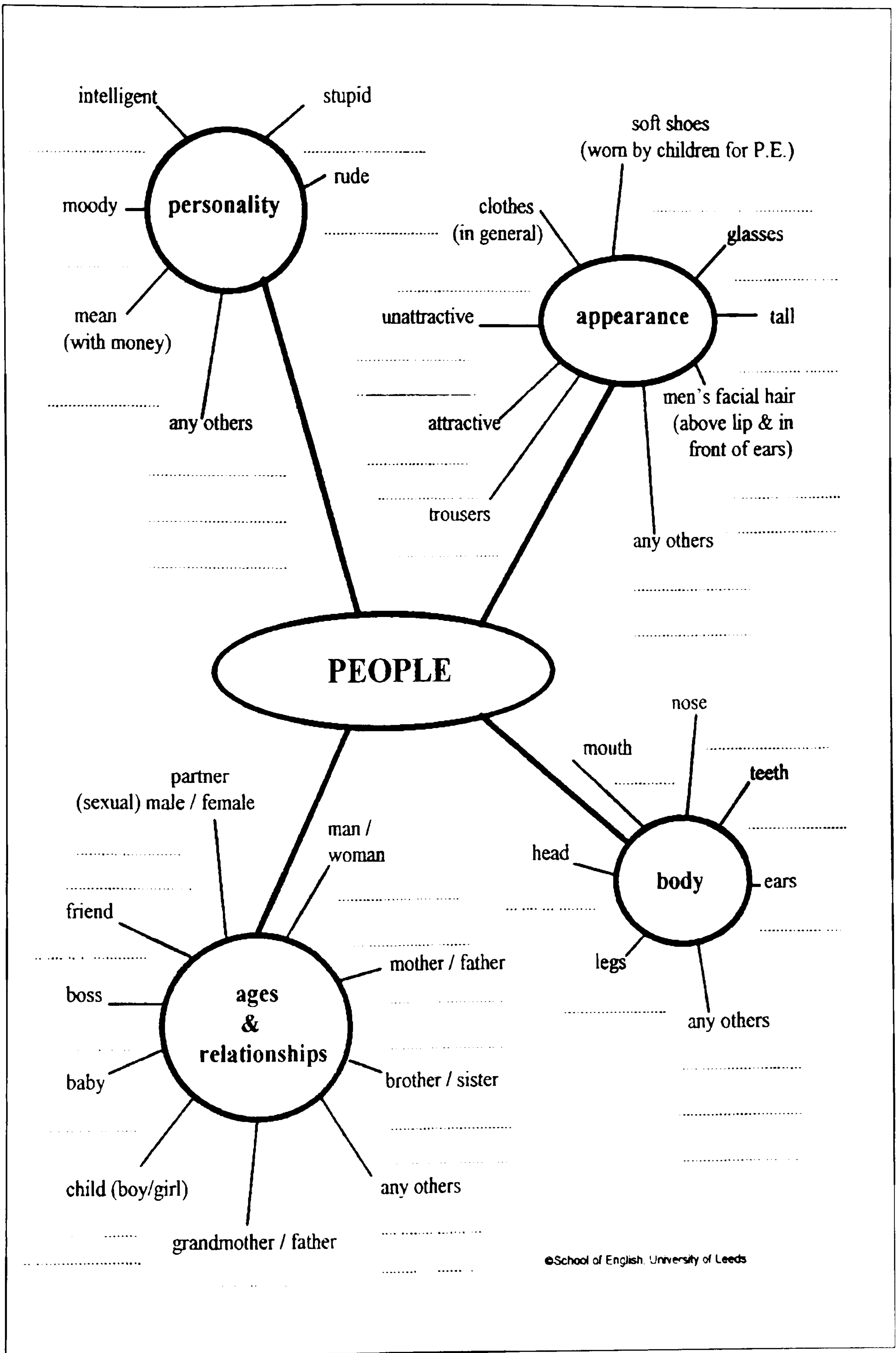
education.....





THE OUTSIDE WORLD





Appendix 2 - Variants recorded by Southampton area speakers on the SuRE Sense Relation Network sheets, listed in decreasing order of frequency

Being, Saying and Doing	
Saying	
Notion word or phrase	Variants
to talk (a lot)	chatter (12), chatter box (9), natter (7), rabbit (4), chat (3), N/R (3), talkative (3), chatty (2), gassing (2), go on (2), (to) rabbit on (2), verbal diarrhoea (2), can't get a word in edgeway (1), chat shit (1), chatters (1), gab (1), gas bag (1), gases a lot (1), goes on (1), goes on a bit (1), goes on a lot (1), going on (1), going on and on (1), gossip (1), to gush (1), has a lot to say (1), he/she doesn't stop (1), hold on a bit (1), mouth in gear (1), mouthy (1), natters away (1), noisy (1), to spew words (1), talk hind leg off a donkey (1), talks a lot (1), talks non-stop (1)
tell to be quiet	(tell to) shut up (31), (tell to) be quiet (10), shush (9), (to) shut it (6), sh (6), hush (2), N/R (2), quiet (2), quiet please (2), be quiet please (1), desist (1), give it a rest (1), leave it out (1), less noise (1), order (1), put a sock in it (1), silence (1), steady on old chap (1), tone down a bit (1)
ask to wait	hang on (15), (to) hold on (13), wait (8), N/R (5), hang on a minute (2), hang on a sec (2), just a minute (2), please wait (2), stay (2), stop (2), wait a mo (2), could you wait a moment please (1), hang about (1), hang on a bit (1), hang on a mo (1), hold on a moment (1), I won't be a minute (1), just a moment please (1), please stay (1), queue (1), stand by (1), stay there (1), stop there (1), wait a minute (1), wait a sec (1), wait here (1), wait please (1), wait there (1), whoa (1)
hello	hi (34), hello (22), all right (11), hiya (11), afternoon (2), hey (2), morning (2), all right mate (1), evening (1), good day (1), hello + name (1), N/R (1), oh hello (1), ok (1), wotcher (1)
how are you?	how are you? (20), all right? (12), how's it going? (5), how's you? (4) ok? (4), are you ok? (3), how are you doing? (3), you all right? (3), enquire (2), how you doing? (2), N/R (2), all ok? (1), are you all right? (1), are you well? (1), everything all right? (1), fine (1), hiya (1), hi, how are you doing? (1), hope you are well (1), how are you feeling? (1), how are you keeping? (1), how are you today? (1), you all right, mate? (1), what you been up to? (1), what you up to? (1), you are looking well (1), you ok? (1)

Notion word or phrase	Variants
goodbye	see you (21), cheerio (20). bye (19), see you later (10). goodbye (6), cheers (5). bye bye (2). later (2). take care (2), adios (2), adieu (1). all the best (1). bye. have a good day (1), catch you later (1). cher-ah (1). farewell (1), in a bit mush (1), later (1). N/R (1). see you in a bit (1), see you later then (1), see you later on (1). so long (1), ta-ta (1)
tell tales on someone	(to) grass (14), snitch (11), N/R (5). dob in (4), (to) gossip (4), tell tales (4). split (3), tell-tale (2). creep (1). dob (1), dobbing (1), don't tell tales (1). drop in it (1). grass them up (1), grassing (1), have you heard (1). inform (1), clip (1), liar (1). load of rubbish (1). nark (1), porky-pies (1), sell out (1). sneak (1). snivel (1). stir shit (1), stirring it (1), tale-teller (1). talk behind one's back (1), talk behind your back (1). tell tales on someone (1), tell-tit (1)
Doing	
play truant	skive (17), bunk off (10), bunk (8), hookey (7), play truant (4), N/R (3), skive off (3). bunking (2). miss school (2), run away (2), skip (2). away (1). A.W.O.L. (1), bunked off (1), bunking off (1), dust (1). off on a skive (1). play hookey (1), skip off (1). skip school (1). skiver (1), skiving (1), slopes off (1). that's a lot of bunk (1), truancy (1), truant (1), wag (1)
to play	(to) play (30), N/R (12). muck about (6), have fun (4). mess around (2), muck around (2), playing (2). gambol (1). games (1), mess (1), mucking about (1), piss around (1), rejoice (1)
to fight	(to) fight (23). (to) scrap (17), N/R (6). punch-up (6), bundle (3), beat (2), brawl (2), fisticuffs (2), kick off (2), rumble (2), argue (1). batter (1). box (1), destroy (1), fighting (1), have a set-to (1), hurt (1), knock out (1). put 'em up (1), resist (1), ruck (1), ruckus (1). to spat (1), struggle (1)
to hit	(to) hit (15), punch (11), (to) smack (11). (to) thump (8), slap (7), (to) whack (5), bash (4). N/R (3), strike (3). wallop (3), (to) clobber (2), smash (2), bashed (1). batter (1), beat (1), clout (1). destroy (1), knock out (1). smacked (1), snuff (1), whacked (1)
to sleep	(to) kip (23). (to) sleep (20), (to) snooze (6). N/R (4). doze (3), nod off (3), have a kip (2). slumber (2). bed (1). bobos (1). forty winks (1). have forty winks (1). have a nap (1). like a log (1). nap (1). right out (1). siesta (1), zzzz (1)
to throw	(to) chuck (29). (to) throw (18). lob (10). (to) toss (7). bung (3). fling (2). hurl (2). N/R (2). he flung it (1). launch (1). lop (1). lug (1). luz (1). sling (1)

Notion word or phrase	Variants
to kiss	(to) kiss (29), (to) snog (26), peck (5), N/R (4), embrace (3), smooch (3), (to) pull (2), get in (1), get off (1), get off with (1), lips together (1), pash (1), peck (1), smacker (1), steam (1)
work hard	(to) work hard (18), (to) graft (15), N/R (7), slog (5), toil (3), slave (2), beaver (1), caning it (1), get stuck in (1), hardworking (1), keep at it (1), peddling (1), put back into (1), slog away (1), sweat your cobs off (1), swot up (1), worked my butt off (1), worked my socks off (1), worker (1), work like a black man (1)
to vomit	(to) be sick (17), (to) throw up (16), puke (12), sick (7), spew (5), chuck up (4), chunder (2), vomit (2), being sick (1), blowing chunks (1), chuck (1), going to be sick (1), green (1), hurl (1), N/R (1), puked (1), retch (1), sickness (1)
run away from	(to) leg it (19), (to) run away from (8), scarper (8), N/R (7), run away (3), abscond (2), escape (2), flee (2), run (2), avoid (1), chicken out (1), to desert (1), gel on (1), high tail it (1), ran off (1), run off (1), scarpered (1), shirk (1), skedaddle (1)
Being	
ill	sick (26), poorly (12), ill (8), unwell (8), rough (7), not well (6), under the weather (3), dodgy (2), N/R (2), ailing (1), bad (1), iffy (1), not too good (1), not very well (1), poorly service (1), shit (1), sicky (1)
pleased	happy (22), chuffed (17), pleased (10), delighted (4), glad (3), N/R (3), over the moon (3), great (2), well chuffed (2), chuffed to bits (1), grateful (1), marvellous (1), ok (1), sound (1), wonderful (1)
insane	mad (27), mental (13), nuts (7), crazy (4), bonkers (3), stupid (3), crackers (2), loony (2), loopy (2), nutter (2), nutty (2), barmy (1), certified (1), completely lost it (1), daft (1), doolally (1), dopey (1), dotty (1), fucked up (1), he's just not with it (1), insane (1), mentally sick (1), N/R (1), not all there (1), not got all his marbles (1), psycho (1), screw loose (1), simple (1), unbalanced (1), weird (1)
pregnant	pregnant (19), expecting (12), up the duff (10), preppers (8), N/R (6), expecting a baby (3), (got a) bun in the oven (3), in (the) family way (2), with child (2), bundle of joy (1), fall pregnant (1), having baby (1), in the club (1), up the gut (1)

Notion word or phrase	Variants
drunk	pissed (24), drunk (12), hammered (4), legless (4), tipsy (4), wasted (4), sloshed (3), smashed (3), fucked (2), mullered (2), shit-faced (2), sozzled (2), steaming (2), wankered (2), battered (1), buckled (1), drop too much (1), he's had one over the eight (1), intoxicated (1), messy (1), N/R (1), off his face (1), one too many (1), out of it (1), over the limit (1), paralytic (1), piddled (1), pie-eyed (1), pissed as a rat (1), plastered (1), rat-arsed (1), skinful (1), stemmed (1), stewed (1), three sheets to the wind (1), tiddly (1), under influence (1), under the weather (1)
to be annoyed by someone	pissed off (16), irritated (5), N/R (5), aggravated (3), cross (3), hacked off (3), upset (3), annoyed (2), fed up with (2), fucked off (2), getting on my tits (2), irked (2), pissing me off (2), wound up (2), annoyed with (1), baked out (1), be aggravated by someone (1), bothered (1), bugging (1), doing my head in (1), drives one bonkers (1), frustrated (1), fuming (1), furious (1), gets my goat (1), gets on one's nerves (1), gets up my nose (1), gets up your backside (1), getting to me (1), get up my back (1), get up my nose (1), got on my nerves (1), ignore (1), pestered (1), pissed off by (1), pissed off with (1), riled (1), rubbed (1), teed off (1), wind me up (1)
cheated (e.g. financially)	robbed (11), ripped off (9), swindled (7), done (6), conned (5), fiddled (5), cheated (3), N/R (3), stitched up (3), deceived (2), screwed (2), screwed over (2), taken for a ride (2), twisted (2), been done (1), bumped (1), diddled (1), done out of money (1), done over (1), duped (1), embezzlement (1), fleeced (1), fraud (1), had (1), mislead (1), rooked (1), skanked (1), swizzed (1), swizzled (1)
dirty	filthy (26), dirty (11), grubby (5), messy (5), mucky (5), N/R (3), scruffy (3), minging (2), rank (2), unclean (2), cacky (1), grim (1), gross (1), grotty (1), gungy (1), mank (1), manky (1), neglected (1), not fit to live in (1), rude (1), scummy (1), shitty (1), smelly (1), soiled (1), unkempt (1)
left-handed	left-handed (28), cack-handed (14), N/R (9), lefty (4), awkward (3), southpaw (3), backhanded (1), cack (1), left-hooker (1)
tired	knackered (37), tired (9), shattered (7), sleepy (5), worn out (5), exhausted (4), weary (3), all in (1), battered (1), beat (1), fit to drop (1), had it (1), N/R (1), nodding off (1), shagged out (1), whacked (1)

Notion word or phrase	Variants
cold	freezing (34), chilly (12), cold (9), nippy (6). shrammed (3), bitter (2), brass monkey (1). brass monkey weather (1), cold enough to freeze the balls off a brass monkey (1). freezing my bollocks off (1), frozen (1), I ought to be an eskimo it's that cold (1). icy (1). it's a bit nippy (1)
hot	boiling (30), hot (15). baking (5), roasting (5), warm (5), close (2), sweating (2), boiling hot (1). cold (1). heated (1), N/R (1), pretty warm (1), scorched (1). scorching (1), sticky (1), sweaty (1), sweltering (1)
Everyday Life	
Nature and Weather	
cat	cat (30), moggie (12), N/R (7), puss (6), feline (4), pussy (4), furry shit machine (1), kitty (1), puddie (1), shit machines (1)
dog	dog (37), N/R (10), mutt (7), canine (3), hound (2), pooch (2), animal (1), mongrel (1), muttlow (1), woofers (1)
running water smaller than river	stream (59), brook (7), ditch (1), rivulet (1)
time between summer and winter	autumn (56), N/R (3), fall (2), spring (1)
to rain lightly	(to) drizzle (34), spitting (9), shower (6), spit (5), drizzling (4), in the air (1), mist (1), N/R (1), raining (1), showering (1), spitting down (1), steady (1)
to rain heavily	pouring (13), pour (9), pouring down (6), chucking it down (4), downpour (4), pissed down (4), piss down (3), raining cats and dogs (3), cats and dogs (2), chucking (2), pissing it down (2), pour down (2), (to) pour with rain (2), stair-rods (2), tipping down (2), belt down with rain (1), bucket down (1), bucketing down (1), chuck it down (1), hammering down (1), heavy rain (1), like stair-rods (1), N/R (1), p down (1), to piss it down (1), piss it off (1), rain cats and dogs (1), storm (1), throwing it down (1), tipping (1), torrential rain (1)
The Home	
main room of the house (with television)	lounge (35), living room (15), front room (11), sitting room (8)
toilet	loo (33), toilet (19), bog (18), lavatory (6), bathroom (3), gents (1), khazi (1), little room (1), N/R (1), shit house (1), shitter (1), small room (1), trap (1), W.C. (1)
small walkway between houses	alley (23), alleyway (12), cut (6), cut way (6), sideway (6), path (5), passage (3), N/R (2), passageway (2), back path (1), cut through (1), hall (1), paths (1), sidewalk (1), walkway (1)

Notion word or phrase	Variants
long seat in main room of house	settee (34), sofa (30), couch (9), bench (1)
television	TV (46), telly (15), (the) box (9), television (2), viewing (2), goggle box (1)
Crime and the Law	
to be in prison	banged up (18), inside (10), locked up (8), in jail (5), in prison (4), in the nick (4), (to be) in clink (3), N/R (3), behind bars (2), clink (2), gone down (2), in nick (2), jailed (2), prisoner (2), put away (2), away (1), cells (1), doing bird (1), in custody (1), in the clink (1), jail (1), nick (1), nicked (1), sent down (1)
the police	(the) police (23), cops (10), coppers (8), (the) old bill (5), (the) fuzz (4), pigs (4), bobbies (3), gavvers (3), N/R (3), bobby (2), filth (2), the law (2), the bill (1), cop (1), copper (1), dirty bastards (1), the 5-0 (1), keep law and order (1), mr plod (1), noddies (1), plod (1), rozzers (1)
to steal	(to) nick (34), pinch (13), (to) steal (11), rob (7), thief (4), filch (2), five finger discount (1), half-inch (1), lift (1), N/R (1), pilfer (1), purloin (1), to scrump apples (1), skank (1), stealing (1), stole (1), theft (1), yoink (1)
police station	police station (28), cop shop (14), (the) nick (11), N/R (4), (the) station (3), (the) cells (2), pig station (2), copper shop (1), house of law (1), jail (1), law agency (1), noddy shop (1), old bill station (1), pig pen (1)
Money	
not have any money left	skint (42), broke (24), boracic (2), hard up (1), impoverished (1), pauper (1), penniless (1), poor fish (1), shafted (1), short of cash (1)
rich	loaded (27), well-off (11), wealthy (10), rich (9), minted (3), N/R (3), rolling in it (3), well-to-do (2), comfortable (1), coming out of his ears (1), flash (1), flushed (1), loads of money (1), made of money (1), not short of a bob or two (1), opulent (1), plush (1), stinking rich (1), swimming in it (1), wedged (1), well-heeled (1)
money in general	cash (28), money (18), dosh (15), N/R (4), dough (3), ackers (2), currency (1), dollar (1), dollars (1), filthy lucre (1), funds (1), housekeeping (1), lolly (1), loose cash (1), loose change (1), moolah (1), pennies (1), pound notes (1), pounds (1), ranch (1), savings (1), wodge (1)
Eating, Drinking and Smoking	
meals of the day	dinner (48), breakfast (47), lunch (42), tea (23), supper (14), food (3), meals (3), brekkie (2), N/R (2), brunch (1), food time (1), go out to dinner (1), menus (1), roast dinner (1), snack (1), specials (1), sustenance (1)

Notion word or phrase	Variants
food taken to work	lunch (30), sandwiches (14), packed lunch (6), snack (5), dinner (3), lunch box (3), N/R (3), sarnies (3), snacks (2), grub (1), pack lunch (1), picnic (1)
sweets	sweets (36), N/R (9), pudding (4), sweeties (4), confectionery (2), munchies (2), chocolates (1), confection (1), dessert (1), foods (1), puddings (1), spice (1)
types of bread	white (20), brown (18), wholemeal (13), bread (11), loaf (11), N/R (10), rolls (10), roll (5), sliced (5), baguette (4), loaves (4), buns (3), french stick (3), granary (3), baguettes (2), bagels (1), baps (1), best of both (1), bloomer (1), bun (1), ciabatta (1), cottage (1), cottage loaf (1), crusty (1), farmhouse (1), flour (1), french (1), hovis (1), jaws (1), kingsmill (1), loaf of bread (1), malted grain (1), naan bread (1), panini (1), sliced bread (1), slices (1), thick sliced toast (1), tiger bread (1), tin (1), unleavened (1), unsliced (1), wholegrain (1)
non-alcoholic drinks	soft drinks (13), squash (11), N/R (9), lemonade (7), fizzy (6), juice (6), soft drink (6), coke (5), pop (5), water (5), fruit juice (3), juices (3), coffee (2), cordial (2), drink (2), drinks (2), soft (2), tea (2), alcohol free (1), apple juice (1), barley (1), barley water (1), cans (1), cola (1), colas (1), cranberry juice (1), dr pepper (1), fizzies (1), fizzy drinks (1), ginger beer (1), milk (1), non-alcoholic drink (1), non-alcoholic drinks (1), orange barley (1), orange juice (1), sprite (1), tomato juice (1)
cigarettes	fags (42), cigarettes (8), ciggies (6), smokes (4), cancer sticks (2), fag (2), beebs (1), cancer stick (1), cigarette (1), cigs (1), drag (1), grats (1), grens (1), N/R (1), no-no (1), roons (1), snouts (1), weeds (1)
People	
Personality	
intelligent	clever (27), bright (17), brainy (13), boffin (4), intelligent (3), smart (3), N/R (2), smart-arse (2), big head (1), brain (1), brain-box (1), clever clogs (1), crafty (1), egg head (1), genned up (1), head screwed on the right way (1), mastermind (1), very able (1), very bright (1), well-read (1)

Notion word or phrase	Variants
unintelligent	thick (29), stupid (9), dim (7), slow (6), dumb (4), dopey (2), ignorant (2), N/R (2), thicko (2), chimp (1), daft (1), dense (1), educationally sub-normal (1), halfpenny short of a shilling (1), he doesn't go further than Thursday (1), idiot (1), intelligent (1), less able (1), moron (1), not brainy (1), not clever (1), retarded (1), thick as shit (1), thick as two short planks (1), uneducated (1), unintelligent (1)
rude	rude (33), N/R (8), ignorant (5), impolite (2), out of order (2), abrasive (1), abusive (1), bad-mannered (1), bad manners (1), cheeky (1), dirty (1), ill-bred (1), ill-mannered (1), nasty (1), nasty type (1), naughty (1), obnoxious (1), offensive (1), out of line (1), snotty (1), starkers (1), uncouth (1), vulgar (1)
moody	moody (23), grumpy (7), N/R (7), stroppy (4), sulky (3), arsey (2), stressy (2), temperamental (2), bitchy (1), down (1), got out of bed the wrong side (1), having a turn (1), humpy (1), in a rotten mood (1), irrational (1), lairy (1), mardy (1), niggle (1), out of sorts (1), pain in the arse (1), shirty gerty (1), shitty (1), stressed (1), sulk (1), up and down (1), wound up (1)
mean (e.g. with money)	tight (40), stingy (9), mean (6), miserly (4), miser (3), tight-fisted (3), skinflint (2), careful (1), frugal (1), jew (1), jewish (1), long pockets short arms (1), penny-pinching (1), scrooge (1), scroogish (1), tight-arse (1), tight as a duck's arse (1)
Looks, Clothes and Accessories	
soft shoes (e.g. worn by children for P.E.)	plimsolls (47), trainers (20), N/R (2), plimmies (2), flip flops (1), gym shoes (1), pumps (1), slip-ons (1), sneakers (1)
tall	tall (32), lanky (22), lofty (4), N/R (3), beanpole (2), giant (2), massive (2), beanstalk (1), big (1), gangly (1), good-looking (1), lamp-post (1), long (1), lurch (1), tall boy (1)
short	short (34), N/R (6), short-arse (6), small (5), shorty (4), stumpy (4), little (3), midget (2), tiny (2), a bit short (1), dumpy (1), dwarf (1), midgets (1), rat (1), squat (1), titch (1), vertically challenged (1)
attractive	pretty (18), fit (13), attractive (10), good-looking (10), beautiful (6), nice-looking (6), gorgeous (4), N/R (3), smart (3), handsome (2), nice (2), dishy (1), fine (1), looker (1), lovely (1), lush (1), pleasant to look at (1), posh (1), something about him (1), stunning (1), vivacious (1)

Notion word or phrase	Variants
unattractive	ugly (35), unattractive (6), minger (5), plain (5), minging (3), munter (2), rank (2), a dog (1), awful (1), dirt (1), disgusting (1), doggin (1), dull (1), gross (1), hideous (1), N/R (1), not a looker (1), not pretty (1), not very attractive (1), not very pretty (1), ordinary (1), pig (1), rough (1), ugly mug (1)
thin	skinny (36), thin (13), slim (7), lean (4), N/R (3), stick (2), anorexic (1), beanpole (1), emaciated (1), gaunt (1), lanky (1), matchstick (1), rake (1), scrawny (1), skeletal (1), twiglet (1), wafer (1), wiry (1)
fat	fat (25), obese (9), large (6), N/R (4), overweight (4), plump (4), tubby (4), fatty (3), podgy (3), big (2), chubby (2), lardy (2), a bit on the large side (1), chunk (1), chunky (1), disgusting (1), dumpy (1), fat boy (1), fatty arbuttle (1), gross (1), gutsy (1), lard arse (1), massive (1), podge (1), porky (1), rank (1), round (1), stout (1), tubs (1)
clothes in general	clothes (30), N/R (12), gear (7), clobber (6), (the) wardrobe (4), togs (3), apparel (2), kit (2), clothing (1), garment (1), outfits (1), wearing apparel (1)
trousers	trousers (42), pants (6), slacks (6), jeans (5), N/R (5), bottoms (2), bell-bottom (1), britches (1), drainpipes (1), full flare (1), strides (1), trews (1)
spectacles	glasses (52), specs (11), goggles (5), four-eyes (2), N/R (2), eyes (1)
Body	
stomach	tummy (20), stomach (17), gut (13), belly (12), N/R (6), abdomen (1), guts (1), inside (1), pot belly (1), tum (1)
mouth	mouth (36), gob (17), oral (5), cake-hole (4), N/R (4), kisser (2), lips (2), trap (1)
nose	nose (35), hooter (7), N/R (7), conk (7), smell (3), beak (2), snitch (2), snout (2), big (1), nasal (1), nostrils (1), sniff (1), snozz (1)
ears	ears (35), lugholes (10), N/R (7), hearing (3), lugs (3), earholes (2), jugs (2), aural (1), ear (1), earhole (1), hear (1), jug handles (1), lughole (1)
head	head (35), bonce (12), N/R (7), nut (3), cranium (1), napper (1), ned (1), nodder (1), noggin (1), nugget (1), shake (1), skull (1)
bottom	bum (29), arse (26), bottom (14), backside (9), butt (5), rear (3), N/R (2), rear end (2), the brown (1), butt cheeks (1), derriere (1), posterior (1), seat (1)
men's facial hair	beard (40), moustache (12), stubble (11), whiskers (7), tash (6), sideboards (5), bum fluff (4), N/R (3), goatee (2), sideburns (2), fluff (1), full set (1), fuzz (1), growth (1), hirsute (1), mickey (1), need to shave (1)

Relations and Relationships	
Notion word or phrase	Variants
partner (sexual) male/female	(the) wife (17), girlfriend (16), boyfriend (15), husband (12), bird (6), missus (6), N/R (6), partner (6), other half (5), my husband (4), their name/nickname term of endearment (3), (the) bitch (2), bloke (2), better half (1), buddy (1), dear (1), fellow (1), girl (1), good friends (1), lover (1), male (1), man (1), mate (1), my bird (1), my other half (1), my partner (1), my wife (1), partners (1), spouse (1), woman (1)
mother/father	mum (51), dad (49), old man (3), parent (3), ma (2), N/R (2), pa (2), parents (2), pops (2), daddy (1), father (1), mother (1), mummy (1), old dear (1), old lady (1), sweetheart (1), the old boy (1)
brother/sister	sister (30), brother (27), bro (10), sis (10), N/R (7), sibling (5), bruv (3), nipper (3), kid (1), kin (1), my sister (1), siblings (1)
child (boy/girl)	kid (18), girl (13), boy (11), N/R (8), nipper (7), son (7), daughter (6), child (5), lad (3), infant (2), kiddy (2), kids (2), my boy (2), offspring (2), ankle-biter (1), girlie (1), junior (1), kiddies (1), little boy (1), little girl (1), littling (1), master (1), miss (1), my sweetheart (1), toddler (1), youngster (1), youngsters (1)
grandmother/grandfather	grand-dad (40), grandma (18), gran (12), nanny (12), nan (11), granny (10), grandpa (9), grandmother (5), gramps (4), grandparents (4), grandfather (3), N/R (3), nana (2), aged people (1), gramp (1), grandmum (1), grandparent (1), grandpop (1)
friend	mate (40), friend (15), pal (11), chum (3), bosom friend (1), bud (1), buddy (1), comrade (1), confidante (1), mush (1), my friend (1)
person in charge at work	(the) boss (46), manager (8), foreman (5), supervisor (3), charge-hand (2), chief (2), gaffer (2), governor (2), N/R (2), their name (2), boss man (1), director (1), head (1), overseer (1), scumbag (1)
baby	baby (31), N/R (7), kid (6), infant (5), sprog (5), babe (2), little one (2), youngster (2), babe in arms (1), child (1), kiddy (1), little lady (1), little man (1), offspring (1), toddler (1), tot (1)
man/woman	woman (29), man (21), bloke (18), fellow (9), lady (7), N/R (7), bird (6), girl (6), chap (5), guy (4), female (3), male (3), geezer (2), adults (1), bit of stuff (1), chick (1), couple (1), dog (1), dolly-bird (1), he (1), her (1), him (1), humans (1), missus (1), mister (1), party (1), peach (1), person (1), she (1), sweetie (1)

Appendix 3 – Standardised lexical variants and alternative spellings

Being, Saying and Doing		
Saying		
Notion word or phrase	Standardised variant	Variant(s) used by informant(s)
to talk a lot	chatter box	chatterbox
to talk a lot	he/she doesn't stop	he/she don't stop
to talk a lot	talks non-stop	talks non stop
to talk a lot	verbal diarrhoea	verbal diaheria
tell to be quiet	sh	sh sh, shh, sshh, sssh, ssshhh
hello	all right	alright
hello	all right mate	alrite mate
hello	hiya	hi ya, hiyer
hello	good day	gooday
hello	watcher	wotcha
how are you?	all right	alright
how are you?	how's it going	hows it goin, hows it going
how are you?	how's you	hows u, hows you
how are you?	you all right	you alright
how are you?	enquire	inquire
how are you?	are you all right	are you alright
how are you?	hiya	hi ya
how are you?	you all right mate	u alrite mate
goodbye	see you	see ya, c ya
goodbye	see you later	see ya later, s'later
goodbye	cheers	cheer
goodbye	adios	adious
tell tales on someone	tell-tale	tell tale
tell tales on someone	clip	klip
tell tales on someone	poriky-pies	poriky pies
tell tales on someone	tell-tit	tell tit
Doing		
play truant	skive	scive
play truant	hookey	hookie, hooky
to fight	punch-up	punch up
to fight	put 'em up	put em up
to fight	ruckus	rukus
to hit	whacked	wacked
to sleep	bobos	bo bo's
work hard	sweat your cobs off	sweat your cobbs off
to vomit	puke	pueke
to vomit	retch	reach
to vomit	spew	spue

Notion word or phrase	Standardised variant	Variant(s) used by informant(s)
run away from	leg it	leggit
run away from	scarper	skarper
Being		
pleased	chuffed	choughed
insane	loony	looney
insane	doolally	do-lahli
drunk	mullered	mulered
drunk	shit-faced	shit faced
drunk	paralytic	paraletic
drunk	pie-eyed	pie eyed
drunk	rat-arsed	rat arsed
drunk	skinful	skinfull
to be annoyed by someone	irritated	□ggravate, irrated
to be annoyed by someone	aggravated	□ggravated, agrivated
to be annoyed by someone	teed off	tee-ed off
cheated (e.g. financially)	conned	coned
cheated (e.g. financially)	stitched up	stiched up
cheated (e.g. financially)	embezzlement	embezelemt
cheated (e.g. financially)	swizzed	swissed
dirty	minging	mingin
dirty	cacky	kacky
dirty	shitty	shity
left-handed	left-handed	left handed
left-handed	cack-handed	cack handed, kac' handed, kack handed
left-handed	lefty	leftie
left-handed	southpaw	south paw
left-handed	backhanded	back handed
left-handed	left-hooker	left hooker
tired	knackered	knackard, knacked, knakered, nackered, k'd
cold	chilly	chilli
cold	shrammed	shramed
cold	freezing my bollocks off	freezing my bollox off
hot	sweating	swetting
Everyday Life		
Nature and Weather		
cat	moggie	moggy
dog	mutt	mut
dog	muttlow	mutlow
time between summer and winter	autumn	autam, autum

Notion word or phrase	Standardised variant	Variant(s) used by informant(s)
to rain lightly	drizzle	drissle
to rain heavily	stair-rods	stairods, sterods, stair rods
to rain heavily	bucket down	bucked down
The Home		
toilet	khazi	kasy
small walkway between houses	alley	ally
small walkway between houses	alleyway	alley way, ally way
small walkway between houses	cut way	cutway
small walkway between houses	sideway	side way
small walkway between houses	passageway	passage way
long seat in main room of house	settee	seate, setee
long seat in main room of house	sofa	soaffer
television	telly	tele
Crime and the Law		
to be in prison	in the nick	in the knic
the police	gavvers	gavers
Money		
not have any money left	boracic	borassic, brassic
rich	opulent	opulant
rich	well-heeled	well heeled
rich	well-off	well off
rich	well-to-do	well to do
money in general	dough	doe
money in general	moolah	molla
Eating, Drinking and Smoking		
meals of the day	brekkie	breakie
meals of the day	sustenance	sustinence
types of bread	baguette	baget
types of bread	loaves	loafes
non-alcoholic drinks	cranberry juice	cranbury juice
cigarettes	ciggies	ciggys
People		
Personality		
intelligent	brain-box	brainbox
intelligent	smart-arse	smart ass, smartass

Notion word or phrase	Standardised variant	Variant(s) used by informant(s)
intelligent	well-read	well read
unintelligent	stupid	stupit
rude	bad-mannered	bad mannered
rude	impolite	impollite
rude	ill-bred	ill bred
moody	arsey	arsy
moody	temperamental	temperemental, tempermental
moody	mardy	mardi
moody	shitty	shity
mean (e.g. with money)	stingy	stingey, stingie
mean (e.g. with money)	miser	mizer
mean (e.g. with money)	tight-fisted	tight fisted, tite fisted
mean (e.g. with money)	penny-pinching	penny pinching
mean (e.g. with money)	scrooge	scroodge
mean (e.g. with money)	tight-arse	tight arse
mean (e.g. with money)	tight as a duck's arse	tight as a ducks ass
Looks, Clothes and Accessories		
soft shoes (e.g. worn by children for P.E.)	plimsolls	plimsoles, plimsoles, plimsols, plinsoles, plymsoles
soft shoes (e.g. worn by children for P.E.)	slip-ons	slip ons
tall	beanpole	bean pole
tall	good-looking	good looking
tall	lamp-post	lamp post
short	short-arse	short arse, short ass, shortass
attractive	good-looking	good looking
attractive	nice-looking	nice looking
unattractive	unattractive	unatractive
unattractive	minging	mingin
thin	skinny	skinney, skiny
thin	beanpole	bean pole
thin	matchstick	match stick
thin	wiry	wirey
fat	fatty	faty
fat	lard arse	lardarse
clothes in general	clobber	clober
trousers	bell-bottom	bell bottom
trousers	trews	trues
spectacles	four-eyes	four eyes

Notion word or phrase	Standardised variant	Variant(s) used by informant(s)
Body		
mouth	cake-hole	cake hole, cakehole
nose	conk	konk
nose	nostrils	nostrells
ears	earholes	ear holes
ears	earhole	ear hole
ears	lughole	lug hole
head	bonce	bonze
bottom	arse	ass
bottom	butt cheeks	butt cheaks
men's facial hair	stubble	stuble
men's facial hair	tash	tache
Relations and Relationships		
partner (sexual) male/female	missus	misses, missis
partner (sexual) male/female	my husband	me husband
partner (sexual) male/female	bitch	biach
partner (sexual) male/female	fellow	fella
child (boy/girl)	kiddy	kiddie
child (boy/girl)	ankle-biter	ankle biter
child (boy/girl)	girlie	girly
child (boy/girl)	littling	littlin
grandmother/grandfather	grand-dad	granda, grandad, granddad
grandmother/grandfather	granny	grannie
grandmother/grandfather	grandpa	grampa
grandmother/grandfather	grandpop	granpop
grandmother/grandfather	granny	grannie
grandmother/grandfather	nana	nanna
person in charge at work	charge-hand	chargehand
person in charge at work	governor	govenor, guvner
person in charge at work	boss man	bossman
person in charge at work	scumbag	scum bag
man/woman	fellow	fella, feller
man/woman	dolly-bird	dollybird
man/woman	missus	missis

Appendix 4 – Results of the Southampton area language questionnaire

	Informants who responded <i>I'd use this when I was talking to a friend (%)</i>	Informants who responded <i>I wouldn't use this but some people do use it in Southampton/ Eastleigh Borough (%)</i>	Informants who responded <i>I'd use this when I was writing to a friend (%)</i>
<u>Verb forms</u>			
Present tense verb endings – nonstandard –s			
I likes toffees	5.00	56.67	1.67
You likes toffees	0.00	43.33	3.33
She likes toffees	58.33	16.67	23.33
We likes toffees	0.00	51.67	1.67
They likes toffees	1.67	48.33	3.33
Past tense verb forms - to be			
You was singing	5.00	55.00	1.67
We was singing	3.33	56.67	1.67
They was singing	3.33	56.67	5.00
I were singing too	6.67	50.00	0.00
And so were John	1.67	50.00	0.00
But Mary weren't singing	5.00	48.33	0.00
Past tense <i>done</i>			
He done that wrong	15.00	48.33	1.67
<i>Sat</i> and <i>stood</i> as present participles			
She was sat over there looking at her car	61.67	16.67	11.67
And he was stood in the corner looking at it	61.67	18.33	11.67
<i>There was</i> with a plural subject			
There was some singers here a minute ago	20.00	48.33	3.33
<i>Should of</i>			
You should of	45.00	33.33	5.00
You should of left half an hour ago	48.33	35.00	5.00
<u>Negation</u>			
Multiple negation			
Count on me, I won't do nothing silly	10.00	55.00	3.33

	Informants who responded <i>I'd use this when I was talking to a friend</i> (%)	Informants who responded <i>I wouldn't use this but some people do use it in Southampton/Eastleigh Borough</i> (%)	Informants who responded <i>I'd use this when I was writing to a friend</i> (%)
<i>Never as past negator</i>			
My friend broke that, I never	10.00	60.00	1.67
<i>Ain't</i>			
That ain't a bird	18.33	50.00	1.67
That ain't working	20.00	53.33	1.67
I ain't got a clue	23.33	56.67	1.67
<u>Adverbial construction</u>			
<u>Adverbial forms</u>			
He writes really quick	38.33	30.00	10.00
She laughs really loudly	76.67	3.33	25.00
<u>Nominal constructions</u>			
<u>Prepositions</u>			
I'm going up my friend's house later	31.67	41.67	3.33
I'm going down my friend's house later	28.33	41.67	5.00
I'm going over my friend's house later	41.70	36.67	5.00
I'm going round my friend's house later	63.33	20.00	10.00
He knocks his hat off of his head	20.00	41.67	1.67
<u>Demonstrative <i>them</i></u>			
Look at them big spiders	21.67	56.67	3.33
<u>Comparative adjectives</u>			
This is the beautifullest house I've seen	3.33	36.67	3.33
I've never seen a beautifuller one	0.00	40.00	3.33
This is the most beautifullest house I've seen	3.33	38.33	1.67
I've never seen a more beautifuller one	1.67	36.67	3.33
But this is the worstest one I've seen	1.67	40.00	0.00

	Informants who responded <i>I'd use this when I was talking to a friend (%)</i>	Informants who responded <i>I wouldn't use this but some people do use it in Southampton/ Eastleigh Borough (%)</i>	Informants who responded <i>I'd use this when I was writing to a friend (%)</i>
I've never seen a worser one	0.00	45.00	0.00
Regularisation of reflexive pronouns			
John likes doing that hisself too	3.33	46.67	1.67
Yes, lots of people do it theirselves	15.00	48.33	1.67
Relative pronouns			
The films what I like best are horror films	5.00	56.67	1.67
The films as I like best are horror films	3.33	31.67	0.00
The films at I like best are horror films	0.00	30.00	0.00
He's the boy what ate all my sweets	6.67	56.67	1.67
He's the boy as ate all my sweets	0.00	30.00	0.00
He's the boy at ate all my sweets	1.67	28.33	0.00
Unmarked plurality in nouns of measurement			
That town is nearly twenty mile away	18.33	41.67	6.67
To make a big cake you need two pound of flour	31.67	33.33	11.67
This string is three inch long	6.67	35.00	5.00
I've got four dozen eggs	75.00	15.00	25.00

Bibliography

- Aitchison, J., 1994. *Words in the Mind: An Introduction to the Mental Lexicon*. 2nd edn. Oxford: Blackwell.
- Altendorf, U., and D. Watt, 2004. The Dialects of the South of England: Phonology. In: E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, C. Upton, eds. *A Handbook of Varieties of English: A Multimedia Reference Tool. Volume 1: Phonology*. Berlin: Mouton de Gruyter, pp. 178-203.
- Asprey, E., 2001. *A Study in West Midland Phonological Variables*. Unpublished M.A. dissertation, Leeds University.
- Associated British Ports, 2007. *Port of Southampton* [online]. No place: Associated British Ports. Available at: <URL: <http://www.abports.co.uk/custinfo/ports/soton/commodities/contain.htm>> [accessed 18th June 2007].
- Atkinson, P., and D. Silverman, 1997. Kundera's *Immortality*: The Interview Society and the Invention of the Self. *Qualitative Inquiry*, 3(3), 304-325.
- Barry, M. V., ed., 1981. *Aspects of English Dialects in Ireland Volume 1: Papers Arising from the Tape-recorded Survey of Hiberno-English Speech*. Belfast: Institute of Irish Studies, Queen's University of Belfast.
- Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan, 1999. *Longman Grammar of Spoken and Written English*. Harlow, Essex: Longman.
- Burchfield, R., 1985. *The English Language*. Oxford: Oxford University Press.
- Cameron, D., 1988. Part One Introduction. In: J. Coates and D. Cameron, eds. *Women in Their Speech Communities: New Perspectives on Language and Sex*. London: Longman, pp. 3-12.
- Cameron, D., and J. Coates, 1988. Some Problems in the Sociolinguistic Explanation of Sex Differences. In: J. Coates and D. Cameron, eds. *Women in*

- Their Speech Communities: New Perspectives on Language and Sex*. London: Longman, pp. 13-26.
- Chambers, J. K., 2003. *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. 2nd edn. Oxford: Blackwell.
- Chambers, J. K., and P. Trudgill, 1998. *Dialectology*. 2nd edn. Cambridge: Cambridge University Press.
- Cheshire, J., 1982. *Variation in an English Dialect: A Sociolinguistic Study*. Cambridge: Cambridge University Press.
- Cheshire, J., 1991. Variation in the Use of *Ain't* in an Urban British Dialect. In: P. Trudgill and J. K. Chambers, eds. *Dialects of English: Studies in Grammatical Variation*. London: Longman, pp. 54-73.
- Cheshire, J., 1998. English Negation from an Interactional Perspective. In: P. Trudgill and J. Cheshire, eds. *The Sociolinguistics Reader Volume I: Multilingualism and Variation*. London: Arnold, pp. 127-144.
- Cheshire, J., 1999. Taming the Vernacular: Some Repercussions for the Study of Syntactic Variation and Spoken Grammar. *Cuadernos de Filología Inglesa*, 8, 59-80.
- Cheshire, J., V. Edwards, and P. Whittle, 1989. Urban British Dialect Grammar: The Question of Dialect Levelling. *English World-Wide*, 10(2), 185-225.
- Davis, L. M., 1983. *English Dialectology: An Introduction*. Alabama: University of Alabama Press.
- Deuchar, M., 1988. A Pragmatic Account of Women's Use of Standard Speech. In: J. Coates and D. Cameron, eds. *Women in Their Speech Communities: New Perspectives on Language and Sex*. London: Longman, pp. 27-32.

Directorate of Strategy and Development. Southampton. 1993. *City Plan Topic Papers: Population*. Southampton: Southampton City Council.

Docherty, G. J., and P. Foulkes, 1999. Derby and Newcastle: Instrumental Phonetics and Variationist Studies. In: P. Foulkes and G. J. Docherty, eds. *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 47-71.

Docherty, G. J., P. Foulkes, J. Milroy, L. Milroy, and D. Walshaw, 1997. Descriptive Adequacy in Phonology: A Variationist Perspective. *Journal of Linguistics*, 33, 275-310.

Dyer, J., 2002. 'We all speak the same round here': Dialect Levelling in a Scottish-English Community. *Journal of Sociolinguistics*, 6/1, 99-116.

Eastleigh Borough Council, 2006. *Eastleigh Economic Profile and Annual Business Survey* [online]. Eastleigh: Eastleigh Borough Council. Available at: <URL: <http://www.eastleigh.gov.uk/PDF/EconomicProfile2006.pdf>> [accessed 18th June 2007].

Eastleigh Borough Council, 2007. *A Great Place to Work and a Great Place to Live* [online]. Eastleigh: Eastleigh Borough Council. Available at: <URL: <http://www.eastleigh.gov.uk/PDF/personnelbooklet281105.pdf>> [accessed 18th June 2007].

Eckert, P., 2000. *Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten High*. Oxford: Blackwell.

Eckert, P., 2003. Language and Gender in Adolescence. In: J. Holmes and M. Meyerhoff, eds. *The Handbook of Language and Gender*. Oxford: Blackwell. pp. 381-400.

Eckert, P., and S. McConnell-Ginet, 1999. New Generalizations and Explanations in Language and Gender Research. *Language in Society*, 28, 185-201.

- Edwards, V. K., P. Trudgill, and B. Weltens, 1984. *The Grammar of English Dialect: A Survey of Research*. London: Economic and Social Research Council.
- Edwards, V., and B. Weltens, 1985. Research on Non-standard Dialects of British English: Progress and Prospects. In: W. Viereck, ed. *Focus on: England and Wales*. Amsterdam: John Benjamins, pp. 97-139.
- Foulkes, P., and G. J. Docherty, 1999. *Urban Voices – Overview*. In: P. Foulkes and G. J. Docherty, eds. *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 1-24.
- Gal, S., 1978. Peasant Men Can't Get Wives: Language Change and Sex Roles in a Bilingual Community. *Language in Society*, 7, 1-16.
- Gal, S., 1979. Language Shift: Social Determinants of Linguistic Change in Bilingual Austria. London: Academic Press.
- Hughes, A., P. Trudgill, and D. Watt, 2005. 4th edn. *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles*. London: Hodder Arnold.
- Information Commissioner, 1998. *Data Protection Act 1998 (chapter 29)* [online]. London: Stationery Office. Available at: <URL: <http://www.hmso.gov.uk/acts/acts1998/19980029.htm#aofs>> [accessed 20th January 2004].
- Johnson, E., 1996. *Lexical Change and Variation in the Southeastern United States, 1930-1990*. Alabama: University of Alabama Press.
- Kerswill, P., C. Llamas, and C. Upton, 1999. The First SuRE Moves: Early Steps Towards a Large Dialect Project. In: C. Upton and K. Wales, eds. *Leeds Studies in English New Series XXX: Dialect Variation in English: Proceedings*

- of the Harold Orton Centenary Conference 1998*. Leeds: University of Leeds. pp. 257-269.
- Labov, W., 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W., 1972a. *Language in the Inner City: Studies in the Black English Vernacular*. Oxford: Blackwell.
- Labov, W., 1972b. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W., 1994. *Principles of Linguistic Change Volume 1: Internal Factors*. Oxford: Blackwell.
- Labov, W., 2001. *Principles of Linguistic Change Volume 2: Social Factors*. Oxford: Blackwell.
- Labov, W., P. Cohen, C. Robins, and J. Lewis, 1968. *A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City Volume 1: Phonological and Grammatical Analysis*. US Office of Education Cooperative Research Project 3288.
- Ladefoged, P., 1993. *A Course in Phonetics*. 3rd edn. London: Harcourt Brace Jovanovich.
- Leeds Archive of Vernacular Culture, 2005. *Incidental Material – Hampshire*. Leeds: Leeds Archive of Vernacular Culture. Available at: <URL: <http://www.leeds.ac.uk/english/activities/lavc/PDFs/SED39Ha.pdf>> [accessed 13th July 2007].
- Le Page, R. B., and A. Tabouret-Keller. 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.

Llamas, C., 1998. Language Variation and Innovation in Middlesbrough: A Pilot Study. *Leeds Working Papers in Linguistics and Phonetics*, 6, 97-114.

Llamas, C., 1999. A New Methodology: Data Elicitation for Social and Regional Language Variation Studies. *Leeds Working Papers in Linguistics and Phonetics*, 7, 95-118.

Llamas, C., 2000. Middlesbrough English: Convergent and Divergent Trends in a Part of Britain with No Identity. *Leeds Working Papers in Linguistics and Phonetics*, 8, 123-48.

Llamas, M., 2001. *Language Variation and Innovation in Teesside English*. Unpublished Ph.D. thesis, University of Leeds.

Local Government Commission for England, 1994. *Final Recommendations on the Future Local Government of Hampshire: A Report to the Secretary of State for the Environment*. London: HMSO.

Macafee, C., 1983. *Varieties of English Around the World: Glasgow*. Amsterdam: John Benjamins.

Macaulay, R. K. S., 1976. Social Class and Language in Glasgow. *Language in Society*, 5, 173-188.

Macaulay, R. K. S., with the assistance of G. D. Trevelyan, 1977. *Language, Social Class, and Education: A Glasgow Study*. Edinburgh: Edinburgh University Press.

Mather, J. Y., H. H. Speitel, and G. W. Leslie, 1975-86. *The Linguistic Atlas of Scotland*. 3 vols. London: Croom Helm.

Mendoza-Denton, N., 2002. Language and Identity. In: J. K. Chambers, P. Trudgill, and N. Schilling-Estes, eds. *The Handbook of Language Variation and Change*. Oxford: Blackwell, pp. 475-99.

- Milroy, J., and L. Milroy, 1978. Belfast: Change and Variation in an Urban Vernacular. In: P. Trudgill, ed. *Sociolinguistic Patterns in British English*. London: Edward Arnold, pp. 19-36.
- Milroy, L., 1987a. *Language and Social Networks*. 2nd edn. Oxford: Blackwell.
- Milroy, L., 1987b. *Observing and Analysing Natural Language: A Critical Account of Sociolinguistic Method*. Oxford: Blackwell.
- Milroy, L., and M. Gordon, 2003. *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.
- Milroy, L., J. Milroy, G. Docherty, P. Foulkes, and D. Walshaw, 1999. Phonological Variation and Change in Contemporary English: Evidence from Newcastle upon Tyne and Derby. *Cuadernos de Filología Inglesa*, 8, 35-46.
- Newbrook, M., 1999. West Wirral: Norms, Self-Reports and Usage. In: P. Foulkes and G. J. Docherty, eds. *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 90-106.
- Office for National Statistics, 2004. *Census 2001 Key Statistics for Urban Areas in England and Wales*. London: Stationery Office.
- Office for National Statistics, 2005. *The National Statistics Socio-economic Classification User Manual*. Basingstoke: Palgrave Macmillan.
- Office for National Statistics, 2007. *Neighbourhood Statistics* [online]. No place: Office for National Statistics. Available at: <URL: <http://www.neighbourhood.statistics.gov.uk>> [accessed 18th June 2007].
- Office of the Deputy Prime Minister, 2004a. *The English Indices of Deprivation 2004* [online]. London: Stationery Office. Available at: <URL: http://www.communities.gov.uk/pub/446/Indicesofdeprivation2004revised_id1128446.pdf> [accessed 19th June 2007].

- Office of the Deputy Prime Minister, 2004b. *Indices of Deprivation 2004* [online]. London: Stationery Office. Available at: <URL: <http://www.communities.gov.uk/SOA/SOALevelID2004.xls>> [accessed 19th June 2007].
- Orton, H., 1962. *Survey of English Dialects (A): Introduction*. Leeds: E. J. Arnold.
- Orton, H., et al., eds, 1962-71. *Survey of English Dialects (B): The Basic Material*. 4 vols, each of 3 parts. Leeds: E. J. Arnold.
- Orton, H., and M. F. Wakelin, eds, 1967-68. *Survey of English Dialects (B): The Basic Material*. Volume 4 *The Southern Counties*. Leeds: E. J. Arnold.
- Oxford English Dictionary*. 2nd edn. 1989 (ed. J. A. Simpson and E. S. C. Weiner). Additions 1993-7 (ed. J. Simpson and E. Weiner; Michael Proffitt), and 3rd edn. (in progress) Mar. 2000- (ed. J. Simpson). OED Online. Oxford University Press. Available at: <URL: <http://www.oed.com>> [accessed 26th September 2005]
- Parry, D., 1977. *The Survey of Anglo-Welsh Dialects: Volume 1 The South-East*. Swansea: University College of Swansea.
- Parry, D., 1979. *The Survey of Anglo-Welsh Dialects: Volume 2 The South-West*. Swansea: University College of Swansea.
- Partnership for Urban South Hampshire, 2007. *Partnership for Urban South Hampshire* [online]. No place: Partnership for Urban South Hampshire. Available at: <URL: <http://www.push.gov.uk/default.aspx>> [accessed 21st June 2007]
- Patterson, A. T., 1970. *Southampton: A Biography*. London: Macmillan.

- Payne, A. C., 1980. Factors Controlling the Acquisition of the Philadelphia Dialect by Out-of-State Children. In: W. Labov, ed. *Locating Language in Time and Space*. London: Academic Press, pp. 143-178.
- Rickford, J. R., T. A. Wasow, N. Mendoza-Denton, and J. Espinoza, 1995. Syntactic Variation and Change in Progress: Loss of the Verbal Coda in Topic-Restricting *As Far As* Constructions. *Language*, 71, 102-131.
- Robinson, C. C., 1876. *A Glossary of Words Pertaining to the Dialect of Mid-Yorkshire*. London: Trübner & Co. for the English Dialect Society.
- Ross, A. S. C., [1954] 2007. Linguistic Class-Indicators in Present-day English. *Neuphilologische Mitteilungen*, 108, 113-149. [Originally published in *Neuphilologische Mitteilungen*, 55, 20-56.]
- Silverman, D., 2001. *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*. 2nd edn. London: Sage.
- Southampton City Council, 2006. *Southampton Local Transport Plan 2006-2011* [online]. Southampton: Southampton City Council. Available at: <URL: <http://www.southampton.gov.uk/transport/transportplanning/localtransportplan/default.asp>> [accessed 18th June 2007].
- South East England Regional Assembly, 2004. *Frequently Asked Questions* [online]. Guildford: South East England Regional Assembly. Available at: <URL: <http://www.southeast-ra.gov.uk/assembly/faq.html>> [accessed 22nd July 2005].
- Stuart-Smith, J., 1999. Glasgow: Accent and Voice Quality. In: P. Foulkes and G. J. Docherty, eds. *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 203-22.

- Tagliamonte, S., and R. Ito, 2002. Think *really different*: Continuity and Specialization in the English Dual Form Adverbs. *Journal of Sociolinguistics*, 6(2), 236-266.
- Tollfree, L., 1999. South East London English: Discrete Versus Continuous Modelling of Consonantal Reduction. In: P. Foulkes and G. J. Docherty, eds. *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 163-184.
- Trudgill, P., 1972. Sex, Covert Prestige and Linguistic Change in the Urban British English of Norwich. *Language in Society*, 1, 179-95.
- Trudgill, P., 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Trudgill, P., 1983. *On Dialect: Social and Geographical Perspectives*. Oxford: Blackwell.
- Trudgill, P., 1986. *Dialects in Contact*. Oxford: Blackwell.
- Trudgill, P., 1999. Norwich: Endogenous and Exogenous Linguistic Change. In: P. Foulkes and G. J. Docherty, eds. *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 124-140.
- Underwood, G. N., 1988. Accent and Identity. In: A. R. Thomas, ed. *Methods in Dialectology: Proceedings of the Sixth International Conference Held at the University College of North Wales, 3rd-7th August 1987*. Clevedon: Multilingual Matters, pp. 406-427.
- Upton, C., and C. Llamas, 1999. Two Large-Scale and Long-Term Language Variation Surveys: A Retrospective and a Plan. *Cuadernos de Filología Inglesa*, 8, 291-304.

- Upton, C., D. Parry, and J. D. A. Widdowson, 1994. *Survey of English Dialects: The Dictionary and Grammar*. London: Routledge.
- Upton, C., and J. D. A. Widdowson, 1996. *An Atlas of English Dialects*. Oxford: Oxford University Press.
- Upton, G., and I. Cook, 1996. *Understanding Statistics*. Oxford: Oxford University Press.
- Viereck, W., 1980. The Dialectal Structure of British English: Lowman's Evidence. *English World-Wide*, 1(1), 25-44.
- Wakelin, M. F., 1977. *English Dialects: An Introduction*. Revised edn. London: The Athlone Press.
- Wakelin, M. F., 1986. *The Southwest of England*. Amsterdam: John Benjamins.
- Wallace, K., 2000. *Speaker Self-Evaluation and Identity in the Southampton Area*. Unpublished M.A. dissertation, Leeds University.
- Watt, D., and L. Milroy, 1999. Patterns of Variation and Change in Three Newcastle Vowels: is this Dialect Levelling? In: P. Foulkes and G. J. Docherty, eds. *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 25-46.
- Weinreich, U., W. Labov, and M. I. Herzog, 1968. Empirical Foundations for a Theory of Language Change. In: W. P. Lehmann and Y. Malkiel, eds. *Directions for Historical Linguistics: A Symposium*. Austin: University of Texas Press, pp. 95-188.
- Wells, J. C., 1982. *Accents of English*. 3 vols. Cambridge: Cambridge University Press.
- Williams, A., and P. Kerswill, 1999. Dialect Levelling: Change and Continuity in Milton Keynes, Reading and Hull. In: P. Foulkes and G. J. Docherty, eds.

Urban Voices: Accent Studies in the British Isles. London: Arnold, pp. 141-162.

Wright, J., 1905. *The English Dialect Grammar.* Oxford: Clarendon Press.