

## **The Next Generation: Aspects of Grammatical Variation in the Speech of some London Preadolescents**

Levey, Stephen

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author

For additional information about this publication click this link.

<http://qmro.qmul.ac.uk/jspui/handle/123456789/1493>

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact [scholarlycommunications@qmul.ac.uk](mailto:scholarlycommunications@qmul.ac.uk)

ZTH 4325 LEV

QM Library



23 1300788 X

*The Next Generation: Aspects of Grammatical Variation  
in the Speech of some London Preadolescents*

Stephen Levey

Queen Mary

London University

submitted for the degree of Doctor of Philosophy

## ABSTRACT

This thesis is a sociolinguistic investigation into aspects of non-phonological variation in a group of preadolescents recorded in outer east London. Focusing on the analysis of selected grammatical variables, it aims to explore the nature and development of linguistic variation in an age group which has not figured prominently in the foundational sociolinguistic literature.

The study is embedded within a variationist framework, and examines how the distribution of vernacular variables selected from different levels of the grammar can provide important insights into the maturing sociolinguistic competence of preadolescent speakers. The distribution of specific grammatical variables is correlated with the broad social dimensions of age and gender in order to examine the social and linguistic constraints which operate on aspects of variation in this age group. Furthermore, the findings which emerge from this study are contextualized in relation to patterns of variation used by older speakers, and are more broadly situated with regard to related patterns of variation in other dialects of English.

Another primary aim of the study is to contribute to empirical characterisations of grammatical variation in southeastern England, an area in which there has been little systematic quantitative investigation of non-phonological variation. Given that London has been identified in the sociolinguistic literature as the site of considerable dialect levelling and a major locus of linguistic innovation, the study explores preadolescents' active participation in some of the burgeoning linguistic changes that are affecting not only southeastern dialects, but also other contemporary varieties of English.

## Acknowledgements

This thesis has had a lengthy gestation period, not least because its completion has had to compete with sometimes very demanding employment commitments over the past several years.

I am indebted to numerous people and institutions for advice and assistance during the completion of the research which underpins this study. Firstly, I am grateful to the school and the staff where I undertook the linguistic fieldwork. I express my deep gratitude to the children who so willingly gave of their time during the recording sessions, and who were so keen to share their narratives of personal experience as well as a range of personal opinions. The vast majority of the children who were recorded are now well beyond preadolescence. I can only speculate about their retention of some of the sociolinguistic patterns documented in this study.

A number of people in academia willingly supplied advice about some of the technical features of the analyses in this study, and/or sent useful papers which extended my thinking about certain aspects of variation. I am especially grateful to David Adger, Isabelle Buchstaller, Richard Cameron, Karen Corrigan, Alex D'Arcy, Sue Fox, Alexandra Georgakopoulou, Annie Johnson, Miriam Meyerhoff, Jennifer Smith, and James Walker. Very special thanks to David Britain for constant support and advice, as well as for access to important unpublished research. Special thanks also to Sali Tagliamonte for helpful advice about Goldvarb, and for generously supplying me with a steady flow of publications which have had an important impact on the research presented here. Finally, but by no means least, heartfelt thanks to my supervisor, Jenny Cheshire, for her unfailing good humour, her tact, patience, excellent advice and feedback. The influence of Jenny's own extensive research on what follows is pervasive.

I declare that the work presented in this thesis is my own work.

Name: Stephen Levey

Signature: *Stephen Levey*

Date: *10/7/07*

# CONTENTS

LIST OF TABLES	10
LIST OF FIGURES	12
TRANSCRIPTION CONVENTIONS	13
<b>1. INTRODUCTION</b>	<b>14</b>
1.1 Introduction	14
1.2 Preadolescents, language variation and change	15
1.3 Research questions	21
<b>2. METHODOLOGY AND INTRODUCTION TO VARIATION IN THE CORPUS</b>	<b>24</b>
2.1 Theoretical framework	24
2.2 Data collection	25
2.2.1 Geographical location	25
2.2.2 The school as a site for sociolinguistic research	28
2.2.3 Ethical considerations	31
2.2.4 The speakers	33
2.2.5 A note on the external variables of age and gender	34
2.2.6 Accessing the vernacular	36
2.2.7 Elicitation of grammatical variables	38
2.3 Transcription protocols, computerisation and data analysis	40
2.3.1 Transcription and computerisation of data	40
2.3.2 The extension of quantitative methods to higher levels of grammatical variation	43
2.3.3 Data analysis	45
2.4 Variation in the preadolescent corpus	51
2.4.1 Non-standard features	51
2.4.2 Selection of variables in the study	53

<b>3. VARIATION IN THE PAST TENSE OF <i>BE</i></b>	<b>59</b>
3.1 Introduction	59
3.2 Diachronic perspectives on <i>was/were</i> variation	63
3.3 Synchronic perspectives on <i>was/were</i> variation	72
3.3.1 Ecological factors: Urbanity, mobility and isolation	72
3.3.2 Social factors: Age, gender and education	73
3.3.3 Existentials: A case for special consideration ?	75
3.3.4 Vernacular pattern 1	79
3.3.5 Vernacular pattern 2	80
3.3.6 Vernacular pattern 3	84
3.3.7 <i>Was/were</i> variation in the southeast	85
3.4 Methodology	94
3.4.1 Data from the Bergen Corpus of London Teenage Language	97
3.5 Distributional analysis	100
3.5.1 Multivariate analysis	115
3.6 Summary	118
<b>4. VARIATION IN THE RELATIVIZATION SYSTEM</b>	<b>123</b>
4.1 Introduction	123
4.2 Overview of the history of relativization in English	124
4.3 Variation in the relative marker system in contemporary British vernaculars	129
4.3.1 The status of the <i>wh</i> -relativization strategy in contemporary British dialects	129
4.3.2 Subject relative deletion in British dialects	133
4.4 Methodology	135
4.4.1 Deciding what to count	135
4.4.2 Selection of factor groups	137
4.5 Distributional analysis	141
4.5.1 Multivariate analysis	151
4.6 Summary	155

<b>5. TENSE VARIATION IN PREADOLESCENT NARRATIVES</b>	<b>159</b>
5.1 Introduction	159
5.2 A note on tense-switching as an historical phenomenon	161
5.3 Tense alternation in preadolescent narratives	166
5.3.1 Variation in the use of the present perfect	169
5.4 Tense variation from a cross-variety perspective	175
5.5 Methodology	181
5.6 Distributional analysis	184
5.6.1 Multivariate analysis	193
5.7 Summary	196
<b>6. VARIATION IN THE USE OF DISCOURSE MARKER <i>LIKE</i></b>	<b>200</b>
6.1 Introduction	200
6.2 Cross-linguistic perspectives on the the grammaticalisation of <i>like</i> -lexemes in discourse marking functions	204
6.3 The evolution of <i>like</i> as a discourse marker	206
6.3.1 The development of <i>like</i> viewed from the perspective of grammaticalisation	206
6.3.2 <i>Like</i> in older varieties of English and regional dialects	208
6.4 Previous approaches to the study of <i>like</i> in contemporary varieties of English	210
6.5 The sociolinguistic distribution of <i>like</i> in contemporary varieties of English	212
6.5.1 Age	212
6.5.2 Gender	214
6.6 Methodology	215
6.6.1 Deciding what to count	215
6.6.2 Coding for syntactic position	216
6.6.3 Coding for pragmatic function	218
6.7 Distributional analysis	220
6.7.1 Variation in pragmatic function	224
6.7.2 Variation in syntactic position	225
6.8 Summary	234



<b>7. VARIATION IN THE QUOTATIVE SYSTEM</b>	<b>236</b>
7.1 Introduction	236
7.2 Developmental issues relating to the acquisition of quotative strategies and the use of reported speech	237
7.2.1 Age effects and direct speech	238
7.2.2 Gender and direct speech	239
7.2.3 Direct speech as a strategy of embedded evaluation	240
7.3 Previous analyses of variation in quotative markers	241
7.3.1 <i>Say, go</i> and <i>be like</i> contrasted	241
7.3.2 Previous research on the sociolinguistic distribution of <i>go</i>	243
7.3.3 Previous research on the sociolinguistic distribution of <i>be like</i>	245
7.4 Methodology	251
7.5 Variation in the lexical-syntactic devices for marking quotation in the preadolescent corpus	254
7.6 Distributional analysis	257
7.6.1 Distribution of common quotative variants by age and gender	259
7.6.2 Content of the quote	261
7.6.3 Grammatical person	266
7.6.4 Tense variation	268
7.6.5 Multivariate analysis	270
7.7 Summary	273
<b>8. VARIATION IN THE USE OF GENERAL EXTENDERS</b>	<b>275</b>
8.1 Introduction	275
8.2 General extenders in other varieties of English and in other languages	276
8.2.1 Cross-variety perspectives on general extenders	276
8.2.2 General extenders in other languages	278
8.3 Previous research on the sociolinguistic distribution of general extenders	280
8.4 The functions of general extenders	284
8.5 Methodology	290
8.6 Distributional analysis	294
8.6.1 Distribution in narrative	308
8.7 Summary	313

<b>9. CONCLUDING REMARKS</b>	<b>315</b>
9.1 The sociolinguistic competence of preadolescents	315
9.1.1 Summary of analyses	316
9.2 Participation in patterns of variation and change	321
9.3 Educational implications of the research	327
9.4 Directions for future research	330
<b>APPENDIX</b>	<b>332</b>
<b>REFERENCES</b>	<b>336</b>

## LIST OF TABLES

Table 2.1 Corpus information	34
Table 3.1 Past tense <i>BE</i> forms in dialects of Middle English	64
Table 3.2 <i>Was/were</i> generalization for six age groups in the London dialect area	85
Table 3.3 Distribution of <i>was/were</i> in the preadolescent data	100
Table 3.4 Distribution of non-standard <i>was</i> in standard <i>were</i> contexts	100
Table 3.5 Distribution of non-standard <i>were</i> in standard <i>was</i> contexts	101
Table 3.6 Distribution of non-standard <i>was</i> in contexts of standard <i>were</i> by polarity	101
Table 3.7 Distribution of non-standard <i>were</i> in contexts of standard <i>was</i> by polarity	102
Table 3.8 Comparison of male and female use of negative tags in COLT	105
Table 3.9 Two independent multivariate analyses of non-standard <i>was</i> in contexts of standard <i>were</i> in the preadolescent and COLT corpora (affirmative contexts only)	116
Table 4.1 Overall distribution of relative markers	143
Table 4.2 Distribution of relative markers by grammatical category	144
Table 4.3 Distribution of relative markers by the animacy of the antecedent NP	145
Table 4.4 Distribution of relative markers according to the definiteness of the antecedent NP	147
Table 4.5 Distribution of relative markers according to the structure of the matrix clause	148
Table 4.6 Distribution of relative markers according to the length of the relative clause	149
Table 4.7 Distribution of relative markers by gender	150
Table 4.8 Distribution of relative markers by age	150
Table 4.9 Three independent multivariate analyses of the contribution of factors to the probability of <i>what</i> , <i>that</i> and <i>zero</i> in the preadolescent corpus	152
Table 5.1 Composition of the narrative sub-corpus by age and gender of speakers	182
Table 5.2 Overall distribution of tense variants used in complicating action clauses	185
Table 5.3 Location of tense forms in complicating action clauses	186
Table 5.4 Distribution of tense forms in complicating action clauses by grammatical person	189

Table 5.5 Distribution of adverbials of temporal specification used in complicating action clauses according to tense form of the verb	190
Table 5.6 Distributional comparison of tense variation in quotative and non-quotative contexts within narrative complicating action sequences	191
Table 5.7 Distribution of tense forms in complicating action clauses by gender	192
Table 5.8 Distribution of tense forms in complicating action clauses by age	193
Table 5.9 Three independent multivariate analyses of the contribution of factors to the probability of the use of preterite, conversational historical present (CHP), and present perfect forms in complicating action clauses	194
Table 6.1 Frequency of discourse <i>like</i> by age group	221
Table 6.2 Functional distribution of discourse <i>like</i> by age and gender	224
Table 6.3 Syntactic position of <i>like</i>	226
Table 7.1 Overall distribution of quotative variants in the preadolescent corpus	258
Table 7.2 Two separate multivariate analyses of factors contributing to the selection of <i>say</i> and <i>go</i> in the preadolescent corpus	271
Table 8.1 Type/token analysis of general extenders according to age and gender	306
Table 8.2 Distribution of general extenders across four different text types	308
Table 9.1 Summary of findings emerging from multivariate analyses	316

## LIST OF FIGURES

Figure 3.1 Comparison of rates of levelling to non-standard <i>was</i> in contexts of standard <i>were</i> in inner and outer London	86
Figure 3.2 Comparison of non-standard <i>wasn't</i> and <i>weren't</i> in inner and outer London	87
Figure 3.3 Comparison of rates of levelling to non-standard <i>weren't</i> in the vernacular of adolescents and elderly speakers recorded in inner and outer London	89
Figure 3.4 Comparative distribution of non-standard <i>was</i> by grammatical person in the vernacular of inner and outer London adolescents and elderly speakers	92
Figure 3.5 Comparison of rates of levelling to non-standard <i>was</i> in contexts of standard <i>were</i> with rates of levelling to non-standard <i>weren't</i> in contexts of standard <i>wasn't</i> in COLT	103
Figure 3.6 Comparison of rates of levelling to non-standard <i>was</i> by grammatical person in the preadolescent and COLT adolescent corpora	106
Figure 3.7 Percentage of non-standard <i>was</i> in contexts of standard <i>were</i> by age	112
Figure 3.8 Percentage of non-standard <i>was</i> in contexts of standard <i>were</i> by gender	113
Figure 5.1 Distribution of tense variants according to tense of verb in preceding complicating action clause	188
Figure 6.1 Frequency of <i>like</i> by age and gender	221
Figure 6.2 Comparison by age and gender of occurrences of <i>like</i> preceding (and within) noun phrases	230
Figure 6.3 Comparison by age and gender of occurrences of <i>like</i> in clause-initial position	231
Figure 7.1 Frequency distribution of <i>be like</i> in five different corpora of British English	248
Figure 7.2 Distribution of <i>say</i> , <i>go</i> , and <i>be like</i> by age and gender	260
Figure 7.3 Frequency distribution of direct speech (per 1000 words) by age and gender	262
Figure 7.4 Proportional comparison of <i>say</i> , <i>go</i> and <i>be like</i> according to the content of the quote	264
Figure 7.5 Functional distribution of <i>go</i> by age and gender	266
Figure 7.6 Distribution of quotatives by grammatical person	267
Figure 7.7 Proportional comparison of <i>say</i> , <i>go</i> and <i>be like</i> according to the five most frequent tense forms in quotative frames	268
Figure 8.1 Frequency of general extender variants in the preadolescent corpus	295

Figure 8.2 Length of general extenders compared according to the number of word forms	296
Figure 8.3 Frequency of list-like sequences in the corpus	297
Figure 8.4 Frequency of general extenders according to the age and gender of speakers	300
Figure 8.5 Comparison of the frequency of adjunctive and disjunctive general extenders by age and gender	302
Figure 8.6 Comparison of the frequency of <i>and everything</i> and <i>or something</i> by gender	303
Figure 8.7 Distribution of general extenders within narrative macrostructure	309

## TRANSCRIPTION CONVENTIONS

The following transcription conventions are used:

' ,	for extracts of direct speech
...	a long pause (not timed)
?	question mark shows the end of a stretch of talk interpreted as a question
!	exclamation mark shows the end of a stretch of talk spoken emphatically
< >	angled brackets give additional information
<unclear>	indicates an inaudible extract of speech
<i>italics</i>	draw attention to linguistic features of interest
CAPITALISATION	(e.g. AAAH) is used to indicate louder parts of talk
:::	indicates vowel lengthening
[ ]	codes in square brackets relate to speaker details in the preadolescent corpus (initial digits identify the individual speaker; M or F indicates whether the speaker is male or female; 7/8 indicates a speaker in the 7-8 year age range; and 10/11 indicates a speaker in the 10-11 year age range)

# CHAPTER 1

## INTRODUCTION

### 1.1 INTRODUCTION

This thesis is a sociolinguistic investigation into aspects of grammatical variation in the speech of a cohort of primary schoolchildren recorded in the outer east London borough of Redbridge. The focus is on the quantitative analysis of selected syntactic, morphosyntactic and discourse variables which are not only encountered in the vernacular of the preadolescents in this investigation, but are also used by older speakers of numerous varieties of English, thus affording some scope for comparative study.

The research described in this study is embedded within a variationist framework. It grew out of questions relating to the acquisition of variation in childhood; the sensitivity of preadolescents to the social meaning of variation; as well as a concern to explore the possible participation of preadolescents in trajectories of linguistic change which are currently underway in contemporary British English.

Another primary aim of the study is to contribute to empirical characterisations of grammatical variation in London and the southeast of England, an area in which there has been little systematic quantitative study of non-phonological variation (Edwards 1993:219; Kerswill and Cheshire 2004-2007).<sup>1</sup>

The following chapters aim to show that there is ample scope for contributing new insights into grammatical variation in preadolescence, and for exploring the

---

<sup>1</sup> Britain (2001: 125) refers to London English as a 'surprisingly under-researched variety'. The dearth of empirical studies in this area is all the more striking because, as Britain points out, the 'diffusing power' of London and the south-east is central to much contemporary research, particularly with regard to phonological variation, and related theoretical issues such as dialect levelling.

evolution of patterns and the social mechanisms underlying variation in the years that lead up to adolescence (Eckert 2000: 8).<sup>2</sup>

## 1.2 PREADOLESCENTS, LANGUAGE VARIATION AND CHANGE

Children have not figured prominently in the foundational sociolinguistic literature. Local (1983: 449) comments that ‘remarkably little is known about the development and functioning of linguistic variability in the speech of children.’ In a similar vein, Romaine (1984a: 32) claims that little is known about the role and place of children in the overall structure of the adult speech community. Even in cases where variationist studies include data from teenage and preteen populations, there is a tendency for preadolescents in particular to be marginalised (D’Arcy 2005b: 328).

The reasons underpinning the relative neglect of children in sociolinguistic research are of both a practical and theoretical nature. Firstly, the collection of large amounts of data from very young children can be difficult and time-consuming. Furthermore, the telegraphic nature of early child speech can engender a number of methodological and analytical problems (see however MacWhinney 1991). Theoretically, the major focus on child language has, until relatively recently, been on the categorical features of language rather than on variation (see Roberts 2002: 336). Sociolinguistic studies of change based on the apparent-time construct have also tended to attach a number of caveats to the inclusion of children and teenagers on the grounds that vernacular stability is assumed to be achieved only during the early adult years (Bayley 2002: 324).

Perhaps the most salient methodological problem with analysing children’s use of variation hinges on the ability to distinguish clearly patterns of variation which are

---

<sup>2</sup> For some earlier well-known work on sociolinguistic variation in preadolescence, see Labov (1972); Romaine (1975, 1984a); and Reid (1978).



socially motivated from those which are developmental in nature and recur in successive generations (Roberts 2002: 335; Chambers 2003: 206).

Although Labov (1970) originally advanced the view that sociolinguistic patterns found in adult speech communities are acquired during adolescence, recent research suggests that ‘...even at an early stage of development children are reproducing many of the salient features of the adult community’ (Foulkes *et al.* 1999: 19). Chambers (2003: 174) contends that the acquisition of grammatical competence and the development of sociolinguistic competence proceed in tandem, which seems an intuitively reasonable assumption given that the input for acquisition is variable (see Labov 1989), although further substantial empirical research is required to corroborate this claim.

There is certainly accumulating evidence that children’s sensitivity to social and stylistic factors in language emerges early. Romaine (1978) found evidence of gender, age, and style variation in the speech of six-, eight-, and ten-year-old Edinburgh children in their use of word final /r/, indicating that socially motivated variation emerges well before the onset of adolescence (see also Andersen 1990). More recent studies focusing on preschool children (e.g. Roberts 1997, 2002) have argued that even during the early years of language acquisition, children are not merely ‘acquirers’ of the vernacular, but active participants in socially motivated language change (Roberts 2002: 333). Roberts and Labov (1995) report that in Philadelphia, preschool children were found to participate in patterns of language variation and change as evidenced by their adoption of short *a* in environments undergoing lexically determined distributional changes. Preschool children were additionally found to acquire aspects of the lexically conditioned change associated with the Philadelphia short *a* pattern at a higher rate than their parents. Such studies indicate that the speech

of young children can indicate 'sites of change that may be accelerated or otherwise as children grow older' (Roberts 2002: 339).

Although children begin their acquisition of variation relatively early (Roberts 2002: 340), it is still unclear, however, whether the social and linguistic constraints on variables are acquired simultaneously. Further research is also required to ascertain whether the acquisition of variables from different levels of the grammar (e.g. syntactic, morphosyntactic and phonological) follow similar or divergent trajectories. Cameron (1998:76), for example, comments that 'in contrast to phonological variation, syntactic or discourse variation may obey different constellations of constraints.' This issue has not been extensively addressed in the literature, largely because most of the research that has been undertaken on the existence of socially meaningful variation in the language of children has focused on the acquisition of phonological variation, with the extension of the variable to other levels of analysis constituting a relatively recent endeavour (Macaulay 2002: 298).

According to Kerswill (1996: 199), a child's acquisition of dialectal features depends on the 'linguistic level, the complexity of the conditioning, and the child's age.' In Cameron's (2005a: 40) view, young children are engaged in the acquisition of multiple types of social and linguistic constraints on variation, with the result that some constraints may be acquired before others. Evidence that this is indeed the case can be found in several studies. Kovac and Adamson's (1981) study of the contraction and deletion of finite *be* in African American and European American three-, five- and seven-year-olds showed that even though the African American children had acquired grammatical and phonological constraints for the contraction of *be*, the constraints on deletion characteristic of adult speakers had not been acquired by the age of seven. Similarly, Labov's (1989) study of (ing) apicalization in a cohort of

Philadelphian children and their parents revealed that the acquisition of stylistic and linguistic constraints on this variable was dependent on the age of the child, with certain constraints being acquired before others: whereas a seven-year-old boy had acquired the linguistic and stylistic constraints associated with (ing), a six-year-old had mastered only the stylistic variation, and the four-year-old had not acquired either the linguistic or stylistic constraints on (ing).

Smith (2005) compared a number of linguistic variables from different levels of the grammar (including phonology and morphosyntax) in a sizeable corpus of interactional data from children aged between two and four as well as their primary caregivers. Smith (2005: 6) reports that there were interesting frequency modulations in caregivers' use of variables depending on situational context (e.g. routine, discipline, teaching) in which their interactions with their children were embedded. Smith (2005) found that the children followed their caregivers' patterns in the frequency distribution of forms in different situational contexts, corroborating the observation made by Foulkes *et al.* (2005: 199) that 'the choices parents make serve to introduce children to socially structured alternatives.' However, the parallelisms between children's and caregivers' usage were not evident in every case: Smith (2005: 11-12) found that as far as children's mastery of the NP/PRO constraint in third person plural contexts was concerned, the children had acquired the complex set of linguistic constraints on this variable, reflecting caregiver usage and wider community norms, but had not acquired the stylistic constraints associated with this

variable.<sup>3</sup> According to Smith (2005), the children's caregivers were less conscious of particular morphosyntactic variables such as the use of *-s* in third person plural contexts than they were of other variables from different levels of the grammar, such as the lexically conditioned alternation between the diphthong [ʌʊ] and the monophthong [u:] as in r[ʌʊ]nd and r[u:]nd. Whereas caregivers appeared to be consciously aware of the distinction between the local variant [u:] and the more widespread and prestigious diphthongal variant, and were therefore able to modify their usage of this variable according to situational context, Smith (2005: 15) argues that this level of conscious awareness did not extend to verbal *-s* with the result that children did not markedly differentiate their own use of this variable according to situational context.

Smith's (2005) findings are important because although they show that children can acquire complex linguistic constraints in infancy, the social constraints associated with particular variables may not be acquired at the same time. The development of sociolinguistic competence in the use of variables from different levels of the grammar may therefore follow quite diverse acquisitional trajectories. Cheshire *et al.*'s (2005a: 167) claim that there is 'no reason to find more than a loose isomorphism in patterns of variation and change in phonology, syntax and discourse' may equally apply to the nature of children's acquisition of variables from different grammatical sub-components.

---

<sup>3</sup> The NP/PRO constraint refers to the variable use of '*-s*' with third person plural NPs at higher rates than in third person plural pronominal contexts. This constraint additionally predicts that '*-s*' will be favoured when the subject and verb are non-adjacent. In some varieties of Scots (e.g. Buckie in the north east of Scotland, where Smith's 2005 study was based) the '*-s*' inflection occurs with third person plural NPs but not with the pronoun *they* resulting in sentences such as 'the men *goes* shopping' versus 'they *go* shopping'. This constraint is discussed in more detail in Chapter 3 in relation to higher rates of *was* usage with third person plural NPs than in third person plural pronominal contexts.

One of the reasons why sociolinguistic and grammatical competence in the use of linguistic variables may be acquired in different ways and at different times is because social categories such as age and gender, which are conventionally correlated with internal patterns of variation, are not immutable, categorical, or monolithic entities, but dynamically situated constructs (see Schiffrin 1996: 199) that have a differential impact on variation at different points in a speaker's lifespan.<sup>4</sup> Furthermore, a variable which is involved in change may behave quite differently from a stable variable in terms of its potential to be recruited for sociosymbolic purposes (Cameron 2005a: 48). Particularly relevant in this connection are discourse-pragmatic innovations such as discourse marker *like* (e.g. *he was like acting stupid*) and the quotative *be like* form (e.g. *I'm like, 'What's happening ?'*) which are spreading rapidly in contemporary English-speaking communities (Tagliamonte and D'Arcy 2004; D'Arcy 2005a; Tagliamonte 2005; see further Chapters 6 and 7). With regard to quotative *be like*, Tagliamonte and D'Arcy (2004: 505) report that although the youngest speakers in their corpus (10-12 year olds) had acquired the internal constraints on *be like* by the time they were in primary school, the effect of gender-differentiation did not assert itself strongly until speakers were in their mid-teens.<sup>5</sup> This finding is consistent with previous research which suggests that adolescence is the 'focus of development of the social use of the vernacular' (Eckert 1997a: 163; Guy 1990).

Although it appears that adolescents are in the vanguard with respect to the use of vernacular variables, and constitute an important focal point for linguistic innovation and change (Chambers 2003: 194), it is unlikely that the fervid symbolic

---

<sup>4</sup> I use the term 'gender' to refer primarily to the social and cultural facets of the sex difference between females and males.

<sup>5</sup> Cameron (2005a: 26) has recently hypothesised that in the preteen age group, gender-affiliated quantitative differences in variation should be at their most salient. Cameron (2005a: 54, f.n. 9) suggests that future research should target this age group to test this hypothesis.

activity which is said to characterise this age cohort emerges *ex nihilo* (Roberts 2002: 334). This suggests that the period immediately preceding the teenage years is fertile territory for examining the roots of socially motivated variation that foreshadows accelerated social uses of the vernacular in adolescence.<sup>6</sup>

However, rather than simply conceiving of preadolescents as having yet to acquire adolescent patterns of variation, the position taken in the following chapters is one that endorses the point of view adopted by Eckert (2000:10), which stresses the importance of recognising the child's sociolinguistic system at any age as a fully mature linguistic system appropriate for that maturational stage. This perspective requires examining age-specific patterns of variation which may not necessarily be congruent with those of teenagers or adults.

### 1.3 RESEARCH QUESTIONS

In examining aspects of variation in preadolescence, my focus is on the usage of a relatively compressed age cohort consisting of children aged seven to eleven. This age range overlaps with important life stages outlined by Labov (2001: 101), namely alignment to the preadolescent peer group (ages eight to nine) and membership in the preadolescent peer group (ages ten to twelve). The children investigated in this study fall within the age range where there is a characteristic move away from parent-oriented to peer-oriented networks, with individuals accommodating their language in systematic ways to that of their peers over a period of time (Kerswill 1996: 196).

---

<sup>6</sup> Although there is ample evidence corroborating Eckert's (1997a: 163) assertion that adolescents 'lead the entire age spectrum in sound change and in the general use of vernacular variables', this may not be invariably the case. Llamas (2007: 75) discusses contrastive patterns of variation in young adult (19-21) and adolescent male (16-17) speakers in Middlesbrough in the north-east of England who make use of innovatory fronted forms of (th) and (dh) (i.e. *fink* for *think*; *bruvver* for *brother*). Contrary to an expected adolescent lead in the use of these variables, young adult males were found to spearhead the use of the non-standard variants.

Although this age range may seem particularly narrow, it is worth bearing in mind that the preteen years constitute an intense period in language (and literacy) acquisition (Kerswill 1996), and that even relatively small differences in chronological age between preadolescent speakers may have potentially greater cognitive, linguistic and possible social ramifications than a comparable age interval between adult speakers (see Cameron 2005a: 54, f.n. 9).<sup>7</sup>

In targeting the vernacular of this particular age group, I aim to address the following areas of enquiry:

- i) To expand existing knowledge about the maturing sociolinguistic competence of preadolescent speakers by carrying out detailed quantitative analyses of non-phonological variables selected from different levels of the grammar.
- ii) To contribute to a more detailed understanding of how the broad social dimensions of age and gender may constrain grammatical patterns of variation in preadolescence.
- iii) To contextualize variation in preadolescence in relation to variation within older age groups in order to compare and contrast patterns of variation across different age cohorts.
- iv) To situate variation in preadolescence vis-à-vis some of the burgeoning linguistic changes that are affecting contemporary British English as well as other varieties of English.<sup>8</sup>

---

<sup>7</sup> Studies of child language variation based on narrowly calibrated age ranges are not unusual (Roberts 2002: 337). The methodological advantages of working with a compressed age range are discussed by Llamas (2007: 75), who notes that 'fairly broad age cohorts will no doubt reveal differences, but can, at times, mask more finely stratified behaviour.'

<sup>8</sup> See also Meyerhoff (2006: 195), who claims that 'we don't really understand how children keep changes in progress moving [...] and this is an area that merits further research.'

- v) To contribute to urban dialect studies in the southeast of England, with a view to exploring similarities and differences between dialects of British English, as well as examining the ‘historical antecedents and evolutionary mechanisms which give rise to contemporary forms.’  
(Milroy and Gordon 2003: 176).<sup>9</sup>

Chapter 2 discusses the theoretical framework and methodology which underpin the study, and provides an overview of the type of variation found in the speech of the preadolescents I recorded. Chapter 3 to 8 provide detailed quantitative analyses of six key variables, and Chapter 9 discusses the salient findings of the investigation, in addition to offering directions for future research.

---

<sup>9</sup> The extent to which the historical antecedents of synchronic variables can be tracked through time varies according to the nature of the variable studied, as will become evident in the ensuing chapters. Variables such as *was/were* and relative markers have been subjected to variationist analyses both synchronically and diachronically (see e.g. Tagliamonte 1998a; Nevalainen 2006; and Romaine 1982). Conversely, discourse variables such as tense alternation and discourse *like* have been subjected to considerably less historical scrutiny within a variationist framework. Thus, the historical analysis of the variables targeted in this study varies from chapter to chapter in terms of breadth and depth of coverage.



# METHODOLOGY AND INTRODUCTION TO VARIATION IN THE CORPUS

### 2.1 THEORETICAL FRAMEWORK

The analyses of the linguistic variables which constitute the mainstay of this study are embedded within the framework of variationist linguistics, which is indebted to much of the foundational research conducted by Labov (see e.g. Labov 1966, 1972, 1982). A fundamental principle underlying this framework is that variation in language is not random, but is characterized by ‘structured heterogeneity’ (Weinreich *et al.* 1968: 99-100). In practice, this means that a multiplicity of linguistic and social factors may systematically constrain the choices that speakers make between variable linguistic forms (Bayley 2002: 118). One of the working assumptions of variationist theory is that the influence of a constellation of internal linguistic and external social factors is amenable to quantitative modelling. Quantitative modelling can be used to explore systematic variation in phonological, morphosyntactic, lexical or discursive features with the aim of expressing in quantitative terms the strength of a co-occurrence between a linguistic variable and any aspect of the surrounding linguistic environment as well as the social phenomena with which a given variable is associated (Young and Bayley 1996: 254; Bayley 2002: 118). As Sankoff (1988: 151) succinctly notes, contextual constraints operating on the choice between variants may be due to features of ‘the phonological environment, the syntactic context, the discursive function of the utterance, topic

style, situation and personal and/or sociodemographic characteristics of the speaker and other participants.'

A further tenet of the variationist paradigm that has a bearing on the approach adopted in the present study is that synchronic variation in a speech community may reflect ongoing diachronic evolution, with the quantitative details of variation preserving 'linguistic history over several centuries' (Labov 1980: xvii). Thus, in several of the chapters which follow I offer a brief summary of the evolution of selected grammatical variables in the history of English in order to contextualize in greater detail patterns of variation in contemporary preadolescent speech, and to situate these within a broader sociolinguistic perspective.

## 2.2 DATA COLLECTION

### 2.2.1 *Geographical location*

The data on which the linguistic analyses are based consist of recordings of children aged between seven- and eleven-years-old, who were recorded between 2000 and 2004 in a large, mixed multicultural state primary school attended by approximately five hundred students in the outer east London borough of Redbridge.

Situated in the vicinity of a large metropolis of considerable cultural, linguistic and political importance, the geographical location of the study is an important one. London is one of the largest cities in the developed world and is the most heavily populated in the European Union. It is also the site of considerable population change and movement. London tends to have a different age structure from the rest of the United Kingdom, and its population tends to be younger on average (Virdee and Williams 2003). It is also characterised by great ethnic diversity. The 2001 Census showed that 29% of the city's population was drawn from a minority ethnic group,

with London containing a higher proportion of people from most minority ethnic groups than any other region of England and Wales (Virdee and Williams 2003: 14-15). These facts point to the tremendous potential of London as a breeding ground for language variation and change.

London has long been recognised as a major locus of linguistic change. Britain (2002a: 616) comments that in the course of the past five hundred years, London and the southeast have been an influential area of linguistic innovation, with many new forms appearing to originate in this region. According to Nevalainen and Raumolin-Brunberg (2003: 162), waves of migration have had a significant impact on the language of London. Indeed, in the sixteenth and seventeenth centuries, no more than fifteen percent of Londoners had been born in the city (Nevalainen and Raumolin-Brunberg 2003: 164). The historical influx of migrants from other areas of the country promoted dialect mixing, and population growth loosened social networks, which became weaker and more uniplex, thereby facilitating linguistic change (see Nevalainen and Raumolin-Brunberg 2003: 165; Milroy and Gordon 2003: 128-9).

Strong claims have been made about the role of London in linguistic change in the latter half of the twentieth century. Wells (1982, Vol II: 301) asserts that London's 'working-class accent is today the most influential source of phonological innovation in England and perhaps in the whole English-speaking world.' More recently, Rosewarne (1994: 3) has argued that a variety of modified regional speech, which he terms 'Estuary English,' combining features of non-regional and local southeastern pronunciation, is evolving into an influential accent which shows signs of geographical diffusion from the suburban areas of Greater London and the counties of

Essex and Kent where it is apparently most in evidence.<sup>1</sup>

Whatever the status of either London's working-class accent or 'Estuary English,' it is recognised that areas of political and cultural importance tend to operate as focal points for the diffusion of linguistic change (Bynon 1977: 214; McMahon 1994: 229). According to Hernández-Campoy (1999:18) London 'constitutes the principal spatial diffusing nucleus from which innovation normally radiates outwards.' Hernández-Campoy's (1999) claim is consistent with the mechanism of geographical diffusion whereby features spread out in a wave-like fashion from a populous and economically dominant centre, affecting nearby towns and cities before more rural areas (Kerswill 2002: 187-8).

Cheshire *et al.* (1999) report that certain phonological variables characteristic of London English such as T-glottaling and TH-fronting are now found in the northern urban centres of Hull and Middlesbrough.<sup>2</sup> Cheshire *et al.* (1999) attribute the presence of these features in northern locales to the gradual diffusion northwards of consonantal phonological innovations which are presumed to have originated in London and the southeast. It is not clear, however, to what extent syntactic or discourse innovations are emanating from London and its adjacent areas, although Krug (2000: 191) claims that emerging modals (e.g. *wanna*, *gotta*, *hafta*) show a broadly similar south-north pattern of diffusion characteristic of phonological innovations.

---

<sup>1</sup> Recent research (e.g. Przedlacka 2001) has countered many of the impressionistic observations about the nature of Estuary English. Estuary English does not appear to be a unitary variety, and there exist noticeable phonetic differences between speakers who are reported to use this variety. Britain (2005: 998) points out that the southeastern regional mix is still focusing, and that distinct local dialects form part of this mix. Furthermore, some of the features that are claimed to characterise this variety, such as the glottalization of /t/, and /l/ vocalisation, appear elsewhere in Britain, and are not exclusive to the United Kingdom (see Foulkes and Docherty 1999).

<sup>2</sup> Of course, what constitutes London English is problematic in itself, and the focus of ongoing research. This issue is taken up in recent research by Kerswill and Cheshire (2004-2007), which addresses *inter alia* the impact of London's ethnic diversity on variation. This research also investigates linguistic variation in London in terms of differences between inner and outer London boroughs. I return to this particular dimension of variation in Chapter 3.

One of the reasons why it may be premature to draw conclusions about the role of London in the dissemination of non-phonological innovations is that there are markedly fewer quantitative studies of syntactic and discourse variation in this geographical area in comparison with studies of phonological variation.

### 2.2.2 *The school as a site for sociolinguistic research*

The school where the fieldwork was based catered for children aged three to eleven. The proportion of children from backgrounds other than white British was much higher than average, with many of these pupils speaking English as an additional language. At the time of the recordings (2000-2004), approximately 36% of the school population spoke English as a second language. The majority of the children with English as an additional language had cultural backgrounds from the north of the Indian sub-continent. However, the sampling criterion employed in this study was strictly based on native speakers of English who had resided in the locality since infancy, hence children who were acquiring English as a second language, or had spent their infancy in other localities in Britain did not take part in the recording sessions.<sup>3</sup> Thus, children who participated in the study were from monolingual backgrounds, and were representative of the school's white ethnic majority.<sup>4</sup> Almost without exception, the parents of the children who were recorded had similar ethnic affiliations, and were native English speakers.

---

<sup>3</sup> The requirement that children should have been resident in the locality since infancy was motivated by sociolinguistic research which suggests that the ability to acquire a second dialect diminishes with age (see Payne 1980; Chambers 2003: 179).

<sup>4</sup> The influence of ethnicity on preadolescent patterns of variation is an important topic to pursue, although I do not discuss it in this study. It is clear from comments made to me by some of the 10-11 year old girls that I recorded that they were aware of stylistic variation associated with particular ethnic groups, and appropriated such stylistic resources for their own purposes. However, examples of the phenomenon of *crossing* (Rampton 1995) were not apparent in any of the recordings I made of the children. It is also important to note that the variables foregrounded in the present study were certainly not exclusive to the white ethnic majority in the school, although a broader and more detailed analysis encompassing data from children with different ethnic backgrounds may well have highlighted important differences as well as areas of convergence.

Most children in the school came from homes that were not as socially or economically advantaged as the national average.<sup>5</sup> From a socioeconomic perspective, the majority of children who took part in this study came from families that can be broadly defined as working-class, based on available indices such as locality and the occupational status of the children's parents (see appendix).<sup>6</sup>

For three of the four years during which the data were collected, I was employed as a teacher in the school, initially on a part-time basis, and later working in a full-time post. My role as an adult male authority figure working within a normative institutional setting presented methodological problems with regard to data collection which warrant further comment here.

Mendoza-Denton (2002: 479) notes that the researcher's identity as well as ideological positioning in relation to interviewees has profound implications for sociolinguistic studies of variation. Within the school setting, the clearly defined social roles and expectations associated with teachers and students can crucially determine and constrain patterns of social behaviour. Eckert (2000: 70) notes that 'conventional sociolinguistic wisdom tells us that school and other normative institutions are problematic sites for the study of the vernacular.' However, if schools are fraught with methodological pitfalls for the researcher, they also compensate by allowing relatively easy access to age-stratified cohorts of children within a restricted geographical space, in addition to being invaluable locations for gathering information about children's vernacular usage.

In spite of the obvious deterrents to undertaking sociolinguistic research with children in an institutional setting where asymmetries of power and age are salient,

---

<sup>5</sup> The source of this information is the Office for Standards in Education.

<sup>6</sup> As far as I was able to ascertain from the children I recorded and the records I had available, the vast majority of the parents of the children had been resident in the locality for a considerable length of time, and the majority had grown up themselves in the southeast. Some of the parents had attended the school selected as the fieldwork site when they had been children.

one of the advantages that I possessed was that I had worked in the school a year before beginning fieldwork, which meant that I was already familiar with many of the children who took part in the study, and knew more about their interests, personalities and family backgrounds than an outsider would. As Milroy and Gordon (2003: 71) note, the local knowledge afforded by 'insider' status extends researchers' explanatory possibilities, and my employment in the school enabled me to make some useful preliminary ethnographic observations.<sup>7</sup> These initial observations were methodologically beneficial because they helped me to identify children who would be potentially keen to participate in the study, and to formulate appropriate questions which would capitalise on the children's interests (see Romaine 1984a: 25). As Baugh (1980: 42) stresses, researchers '[...] will probably have little success in collecting colloquial speech through direct enquiries that are not framed in terms of the consultants' cultural perspectives.' Careful participant observation ensured that my knowledge of individual children's interests could be fully exploited during the recording sessions to help overcome the constraints of formality, and collect samples of vernacular speech.

Additional steps were also taken to minimise the impact of power and age asymmetries on the quality of the recording sessions with the children. The children who took part in the study were assured that whatever was recorded would not be discussed with anyone else in the school. The recordings themselves were always

---

<sup>7</sup> It is important to point out, however, that the present study makes very limited use of ethnographic methods. Eckert (2000: 69) claims that the pursuit of local categories is central to an ethnographic methodology. The pursuit of local meanings necessarily requires extended participant-observation and considerable time investments, both of which are not always practical for everyone. A study of the children's social networks, or communities of practice, would no doubt have been illuminating for investigating the emergence of localised styles in which preadolescents may engage, but practical constraints made an ethnographically oriented study very difficult for me to pursue. Unlike Eckert (2000), I did not fully enjoy the researcher's privilege of mobility. During the time in which I worked in the school where the recordings were made, my institutional responsibilities often meant that I had to be in certain places at specific times, even during certain recreational periods. Thus, institutional constraints relating to my contractual responsibilities in the school were a major impediment to collecting the kinds of sociometric data that are often used in ethnographic studies to establish particular social groupings of speakers.

made in a quiet private area in the school, either in a small group room, or, less often, in an empty classroom during lunchtime.<sup>8</sup> In the areas where the recordings took place, I often sat with the children on the carpet and let the children operate the recording equipment to make the recording sessions seem as informal and casual as possible. The fact that most of the children who were recorded were keen for the sessions to be repeated on a regular basis is an indication that they felt neither obliged to take part in the study nor intimidated during the recording sessions themselves.<sup>9</sup> In fact, the general level of enthusiasm exhibited by most of the children, and their willingness to expound their views and share narratives of personal experience, bear testimony to Eckert's (2000: 71) observation that because power and status asymmetries with adults are acutely felt by children, particularly in institutional settings, many are keen to explore new kinds of relationships with adults when the opportunity arises.

### 2.2.3 *Ethical considerations*

A fundamental prerequisite underpinning ethical research with human beings is the principle of informed consent (Milroy and Gordon 2003: 79). All the children who took part in the study participated on a purely voluntary basis, and were

---

<sup>8</sup> Recording children in different situational contexts (e.g. in the classroom) would have facilitated examination of patterns of style-shifting. However, I did not pursue this goal for a variety of reasons. Firstly, a primary aim of the study was to tap children's vernacular speech rather than to investigate style-shifting *per se*. Secondly, and perhaps of greater importance given my need to maintain a positive rapport with the children, the more formal setting of the classroom constitutes a site of 'continual power struggles between adults and children' (Eckert 2000: 73). As a part-time member of staff in the school with potential authority over children, recording the children in the classroom would have possibly exacerbated existing power asymmetries of age and power which might have adversely affected not only my relationship with the children outside the classroom, but also the quality of recorded data I would have been able to collect.

<sup>9</sup> On more than one occasion, children involved in the study recruited the headteacher and other members of staff in the school to track me down on the premises so that they could begin recording sessions as soon as possible during recreational periods.



guaranteed anonymity and confidentiality.<sup>10</sup> Letters were also sent home to parents about the project.

Children were asked if they were willing to be recorded for a study that was part of a project based at London University, and which was directed at finding out more about schoolchildren's lives, their recreational interests and hobbies as well as their language, including their views on how they and other people spoke.

In the normative setting of the school, any attempt to investigate children's language by an adult may be potentially construed by children as an evaluative exercise, which could seriously undermine efforts to access children's informal speech. Thus, following the recommendations made by Milroy and Gordon (2003: 80), the aims of the study were framed in very general terms, and language was described as just one of several different foci of the project. Furthermore, I made a concerted effort to reassure the children who took part that the project had nothing whatsoever to do with school. Recordings were only made during recreational periods and never during lesson time.

Cameron *et al.* (1993: 81) point out that the unequal partnership between the researcher and the subjects of research has prompted several different conceptualizations of the type of ethical framework in which researchers should operate. Previous sociolinguistic studies have invoked the principle of debt incurred according to which researchers are obliged to use their knowledge and expertise for the benefit of a community which has provided information to researchers (see Labov 1982: 173). Rickford (1997: 182) similarly advocates the notion of 'service in return' for linguistic information gathered from a community. In Chapter 9, I briefly discuss some of the educational implications arising from the findings of this study, including

---

<sup>10</sup> In line with current practice, the names of any children which appear in this study are pseudonyms.

the need for professionals to have a greater awareness of the vernacular norms of children's spoken language.

In terms of more practical acts of gratuity, one way in which I thanked the school community for allowing me to carry out recordings on site was to return to the school when I was later employed on a part-time basis elsewhere, and volunteer my services in order to offer extra tuition to small groups of children who were preparing for standardised government tests.<sup>11</sup>

#### 2.2.4 *The speakers*

A major aim of the study was to obtain extended samples of vernacular speech from a range of children who volunteered to be recorded as part of the study. The recording sessions took place intermittently over the course of four years (2000-2004). The children were recorded using a Sony TCM-939 cassette recorder as well as either an unobtrusive clip-on microphone, or a larger, more sensitive stand-alone microphone placed on a flat surface close to where the children were sitting.<sup>12</sup>

In total, 48 children (27 girls and 21 boys) aged between seven and eleven took part in the study.<sup>13</sup> Table 2.1 below gives a breakdown of the corpus by age and gender. The total number of words recorded for each speaker group is also displayed.

---

<sup>11</sup> Although this was not necessarily a linguistic favour in return for data gathered from the school community (see Wolfram 1993), it was nevertheless an act which was intended to directly benefit children in the school.

<sup>12</sup> The recording equipment was initially treated with curiosity by the children, but once they were fully engaged in talk, the equipment was quickly forgotten.

<sup>13</sup> This age range is not intended to be definitive of preadolescence. Researchers vary in the way that they define the preadolescent age cohort chronologically (e.g. Kerswill 1996: 191 includes children aged six to twelve in this age group).

TABLE 2.1

## CORPUS INFORMATION

Speaker group	Number of speakers	Number of words
7-8 year old girls	11	16,903
7-8 year old boys	9	17,629
10-11 year old girls	16	19,892
10-11 year old boys	12	15,710
TOTAL	48	70,125

Before the start of the recording sessions, I collected background information on each child I recorded, including their name, date of birth, as well as the locality in which they had spent their childhood, and the occupation of their parents. Further details relating to the speakers and their parents can be found in the appendix.

### 2.2.5 *A note on the external variables of age and gender*

Speakers in the present study were divided according to the salient social dimensions of age and gender. Since both these categories have been problematised in the sociolinguistic literature (see e.g. Cheshire 2002a; Milroy and Gordon 2003; Cameron 2005a), they require further discussion here. Milroy and Gordon (2003: 39) comment that the assignment of chronological age to an individual has no explanatory value in itself. Age can, however, be recruited as an insightful analytical construct when it is contextualized in terms of its social significance and used to illuminate meaningful differences in life experiences. Eckert (1997a: 155) advocates a life-stage approach to age rather than one based simply on chronology because the significance

of age lies in the fact that 'the individual's place in society, the community and the family changes through time.'

In the present study, the age range of seven to eleven years represents the period in which it is claimed that the influence of the peer group is in the ascendant, with children accommodating their speech to that of their peers and older children (Kerswill 1996: 196). The older children (10-11 years old) who took part in this study were the most senior in their institution, and were viewed as 'role models' by the younger children. Indeed, the developmental imperative for the 10-11 year olds to be seen as 'grown-up' was institutionally sanctioned, with the most senior children in the school being continually entreated by adults to act as exemplars of mature and responsible behaviour to be emulated by the younger children in the school.

In brief, the passage from late childhood to adolescence constitutes a key life stage, at least in many western industrialised societies. In the British educational system, this passage also coincides with the transition from primary to secondary education at the end of Year 6. Previous research has identified this transition as the period in which a 'marketplace of identities' emerges in late preadolescence in connection with the salient social dimensions of sexuality, gender, race and class (Eckert 1997b).

Just as there has been a call for age to be contextualized in relation to its social significance at particular points in the lifespan, there has been a growing concern to situate gender within a social constructionist framework (Holmes 1997: 195-6), rather than viewing gender differences from the reductionist perspective of a fixed binary opposition that has an unchanging impact on linguistic or social behaviour. Following Butler (1990), many researchers stress the performative nature of gender as something that emerges from 'what people *do* rather than what they intrinsically are' (Cameron

2005: 488-9). Gender can therefore be foregrounded to different extents in different situations, and even at different moments within a single interaction (Cheshire 2002a: 435). Even though I categorise speakers in this study according to age and gender, I do not intend this procedure to imply that either age or gender are categorical attributes of speakers' social identities. Following Macaulay (2005: 39) and Milroy and Milroy (1997: 53), I use gender and age as broad methodological and exploratory extralinguistic variables that are drawn on to investigate the combinatory effects of external and internal linguistic factors on a given linguistic variable (Tagliamonte 2002a: 731).

#### 2.2.6 *Accessing the vernacular*

Before outlining the methodological steps that were taken to access samples of children's vernacular speech, it is expedient at this juncture to consider briefly what the vernacular is. Labov (1984a: 29) has used this term to designate the variety which is learned in the preadolescent years. It has also been described as 'the language used by ordinary people in their everyday affairs' and 'the style in which the minimum attention is given to the monitoring of speech' (Labov 1972: 69, 208). However, while one of the primary objectives of many sociolinguistic studies has been to gather examples of speech which are maximally removed from formal, high-status varieties, the notion of the vernacular as the most unmonitored, unself-conscious form of speech is not unequivocally straightforward. Exaggerated, stylised vernacular performances may be highly self-conscious activities, as Schilling-Estes (1998) has pointed out. Thus, while I use the term 'vernacular' in the sense in which it is employed by Eckert (2000: 17) as the 'language of locally based communities,' I endorse Milroy and Gordon's (2003: 50) view that the vernacular is fundamentally an

abstraction, and that any speech, even of an informal nature, varies substantially in accordance with situational context.

In order to tap the vernacular, the children who participated in the study were normally recorded in single-sex pairs, or, less frequently, in small single-sex friendship groups in order to encourage peer interaction and convergence towards the speech norms of the peer group. The recording sessions were of variable length, but generally lasted between fifteen and twenty-five minutes depending on the topic of discussion and the children's willingness to talk. In all the conversations I recorded, I endeavoured to play the part of a receptive and sympathetic listener rather than an equal participant in order to maximise the number of personal narratives, reminiscences, and opinions collected from the children (see Macaulay 2002: 288). I managed to sustain a very colloquial format during the recording sessions, and avoid the type of interactional style which is characteristic of conventional classrooms where rights to speak, introduce, and sustain topics are typically asymmetrical between children and adults (Romaine 1984a: 170).

Interview techniques successfully adopted in previous sociolinguistic research were employed to minimise the effects of the Observer's Paradox (see Labov 1970: 32). Speakers were encouraged to share narratives of personal experience centred on topics of high emotional involvement such as arguments, fights, and serious accidents, with a view to triggering shifts towards vernacular norms during narration (see Labov 1984a: 32). Speakers were also asked about games they played, as well as about their attitudes towards language, and their opinions of the differences between the behaviour of boys and girls. The children were recorded by the same interviewer (myself) in similar locations so that most of the situational variables underlying the sociolinguistic interviews were kept constant.

Many children were very willing to attend more than one recording session, and I encouraged this. Sometimes a child would be recorded with the friend that she/he had brought along to their initial recording session, or, on other occasions, a child would request to be recorded with a different friend from their peer network. Not only did this help to increase the number of speakers who participated in this study, it was also methodologically beneficial because it encouraged rapport-building with children who were involved in repeated recording sessions. Furthermore, subsequent interviews with the same children enabled me to draw on stories that they had recounted in previous interviews to elicit further samples of vernacular speech (see also Tagliamonte 2006a: 45).

#### 2.2.7 *Elicitation of grammatical variables*

One of the major difficulties associated with the application of the sociolinguistic variable to the study of non-phonological variables is the relative infrequency of syntactic variables in spontaneous discourse. Rickford *et al.* (1995: 106) observe that '[...] unlike phonological variables which show up with high frequencies in [sociolinguistic] interviews, syntactic variables often involve special semantic and pragmatic circumstances which occur rarely or unpredictably in interview settings.' Many interesting syntactic variables such as relative clauses are infrequent in spontaneous spoken discourse (Milroy and Gordon 2003: 63), although, as I demonstrate in Chapter 4, even a somewhat restricted number of relative clauses can yield insights into significant grammatical patterning (see also Cornips and Corrigan 2005a: 102).

In the case of other aspects of variation targeted in this study, there were relatively few difficulties encountered in eliciting appropriate variables. Given that

the interview protocol was largely directed at eliciting narratives of personal experience (see Labov and Waletzky 1967; Labov 1972), which, by their very nature, focus on events that have occurred in the past, it was relatively straightforward to collect examples of past tense *BE* forms, which constitute the focus of Chapter 3. Similarly, since many of the elicited narratives were ‘performed’ rather than simply recounted, it was not too difficult to collect examples of tense-switching, which is a typical stylistic feature of many dramatised narrative recounts (see Chapter 5; Schiffrin 1981; Wolfson 1982). Nevertheless, it is important to concede that even successful sociolinguistic interview techniques are subject to pragmatic constraints that impose restrictions on the type of constructions that can be easily elicited. To take an example cited by Milroy and Gordon (2003: 63), interrogative constructions are likely to be infrequent in the speech of interviewees, which in the case of *was/were* variation examined in Chapter 3, complicates the analysis of morphosyntactic variation in this particular syntactic environment owing to a restricted number of relevant tokens in the corpus.

In other cases, problems with the elicitation of constructions which are intended to form the basis of quantitative analysis can be circumvented with a little ingenuity (Milroy and Gordon 2003: 173). For example, I was able to elicit examples of quotative variants (e.g. *I said*, ‘Where are you going’ and *he’s like*, ‘It’s none of your business’) by encouraging the children I recorded to report what was said during heated exchanges they had engaged in with parents, friends and siblings. Even in a relatively small corpus, this technique yielded over 500 tokens of quotative forms, which is sufficient for exploring consistent effects in patterns of variation (see Chapter 7; Young and Bailey 1996: 259).



## 2.3 TRANSCRIPTION PROTOCOLS, COMPUTERISATION AND DATA ANALYSIS

### 2.3.1 *Transcription and computerisation of the data*

In line with previous research on corpus-based analysis of variation (e.g. Poplack 1989), transcriptions of the recordings of the children were used to construct a computerised database which could be easily manipulated so that the linguistic features under investigation would be maximally accessible. In terms of the degree of fidelity of the computerised transcriptions to the original recordings, there is always a difficult balancing act to be negotiated between the size of the corpus and the level of detail that can be reasonably incorporated in an automated database when analysts have to operate with practical constraints such as time and human resources (Poplack 1989: 430). Tagliamonte (2006a: 54) proposes that transcription protocols should be drawn up in accordance with two major aims: firstly, computerised transcriptions should be sufficiently detailed in order to conduct efficient linguistic analyses; and secondly, transcriptions should be easily readable.

In order to meet these aims, a transcription protocol was devised which enabled targeted linguistic features to be tagged and identified efficiently, including features that were not phonetically overt such as zero relative markers (e.g. *that's the woman* [ $\emptyset$  *I saw*]). The transcription protocol also accounted for dysfluencies such as hesitations and false starts which are characteristic of all natural speech data, and ensured that other important contextual information such as facial expressions, sound effects and gestures were identified and highlighted in the transcription in a systematic fashion as these phenomena were of relevance to some of the variables targeted for analysis (see e.g. Chapter 7 which deals with variation in quotative forms). Following Poplack (1989), idiosyncratic transcriptions were adopted in

certain cases to isolate high frequency forms easily. Thus, the discourse particle *like* was consistently transcribed as *lyke* in order to distinguish it from other homographs in the database with different grammatical functions (i.e. *like* as a main verb, a conjunction, or a suffix; see Tagliamonte 1998b).

The identification of non-standard grammatical features in the corpus was based on the findings of previous research (see especially Cheshire *et al.* 1993; and Cheshire *et al.* 2005a) which has highlighted a set of geographically widespread morphosyntactic variables such as multiple negation, ‘default singulars’ (e.g. *there was pens on the table*), non-standard relative markers (e.g. *the girl what I saw*) that are found in contemporary urban British vernaculars (see further 2.4.1 below). Although these (and other) non-standard features are amply instantiated in the preadolescent data, it is important to bear in mind that coding procedures, as Sigley (1997a: 101) remarks, invariably involve some idealisation of the data. The identification of non-standard morphosyntactic features employed in this study rests on the conventional variationist procedure of classifying non-standard variation in relation to prescriptively defined standard alternatives. Although this approach frequently yields a sufficient number of tokens for quantitative analysis using a sizeable corpus, researchers working within less orthodox frameworks have argued that higher level variables such as those drawn from syntax may require a more extensive approach which requires the analysis of a broader set of forms drawn from different components of the grammar (see Cheshire 2005a).

As far as transcribing the data was concerned, each line of transcription was coded so that information about the individual speaker, her/his age and gender was easily recoverable. The presence of non-standard morphosyntactic features as well as the use of a number of discourse markers were also coded for. The identification of

selected non-standard features in the computerised corpus using a systematic coding string is illustrated (with the accompanying text) in (1) below.

(1) Example of transcription protocol

(9F6MAPXXXYXXXX) these houses had lyke ... these rooftops  
(9F6CAPXCXXXSXX) and my dad kept on saying to me, 'There's ghosts  
living up there'

In the first coding sequence, the number 9 identifies the individual speaker, F indicates that the speaker is female, and 6 identifies the speaker as belonging to the 10-11 age cohort. The fourth character, M, indicates that extract consists of an independent clause; the following A indicates that the verb is active; and the P indicates the tense of the verb, in this case the simple past. The eight symbol code that follows the P is designed to highlight the presence or absence of key morphosyntactic, syntactic or discourse variables. X is the default 'fill' character which is used when there is no example of a targeted feature in the transcription line. In the first coding string, the only noteworthy feature is indicated by Y, which indicates the presence of a discourse marker, in this case *like*, which precedes an untimed pause represented by three dots. In the next coding string, the first C indicates that the second line of transcription begins with the co-ordinating conjunction *and*; the second C in the coding sequence indicates the absence of agreement between the verb in the extract cited and the postverbal associate NP; and the S indicates the presence of a verb of quotation, in this case *say*.

The use of a systematic coding procedure as well as a formalised transcription protocol guaranteed that recurrences of the same linguistic feature were treated

consistently (Poplack 1989: 433). This later facilitated automated extraction of the linguistic variables that constitute the focus of the present study.<sup>14</sup>

Once the corpus had been transcribed and coded, it was stored in text files that were analysed using TACT (Text Analysis Computing Tools, version 2.1). This is a versatile program which is capable of using text files, word-frequency lists, concordances, and type-token statistics. It offers adequate flexibility for descriptive linguistic analysis (Kennedy 1998).

### 2.3.2 *The extension of quantitative methods to higher levels of grammatical variation*

The definition of the envelope of variation is an important methodological prerequisite in any accountable quantitative study of linguistic variation. However, the circumscription of the variable context runs into a number of methodological problems when the concept of the linguistic variable, which was originally applied to the study of phonological variation, is extended to the study of syntactic, morphosyntactic and discourse variation, as is the case in this study.

Whereas phonological variables are based on arbitrary form-meaning relations, this is not the case for syntactic and discourse variables, where, as Cheshire (2005a: 479) points out, variable forms can convey important differences in propositional and attitudinal meanings. The contrastive semiotic functions of syntactic variables as opposed to phonological variables have been viewed as constituting a major impediment to the extension of the linguistic variable beyond phonology (Sankoff 1973; Cheshire 1987; Romaine 1984b: 410). Whereas variants such as *talking* and *talkin'* may be referentially equivalent, albeit with different social and

---

<sup>14</sup> Note that that the extraction of key variables was an initial phase in the organisation and manipulation of the data, and that substantive recoding was subsequently necessary in order to prepare the data for more detailed quantitative analyses of specific features.

stylistic connotations attached to them, the referential equivalence of postulated syntactic variants such as *they broke into the liquor closet* and *the liquor closet was broken into* (see Weiner and Labov 1983) is more difficult to establish, particularly as meaning in this case is highly dependent on the context of the utterance, and is subject to pragmatic conditioning (Cornips and Corrigan 2005a: 104; Romaine 1984b). Cheshire (2005a: 482) notes that a compromise has been tacitly reached in the variationist paradigm in relation to the referential indeterminacy of syntactic variants, with many analysts opting to relax the condition of semantic equivalence in favour of equivalence of discourse function. Dines (1980: 15), for example, suggests that variables can be identified on the basis of their having a 'common function in discourse.' Dines (1980: 15-16) adds that 'this does not preclude the possibility of a semantic tie existing between variants, but it removes the necessity of one.' This is the position which is adopted in this study. Thus, in Chapter 7, for example, which focuses on variation in the quotative system, I follow Buchstaller (2006a: 5) in positing a (*say*) variable which encompasses 'all strategies used to introduce reported speech, sounds, gesture and thought by self or other.' The advantage of this approach is that it enables the analyst to compare and contrast individual variants with other variants which have a common function in a quotative context.<sup>15</sup>

---

<sup>15</sup> Other problems in extending variation analysis to non-phonological variables relate to the adherence to Labov's principle of accountability. Labov (1982: 30) originally argued that the variable context should be defined in a systematic and accountable manner, claiming further that every occurrence of a given variable should be noted and 'where it has been possible to define the variable as a closed set of variants, all non-occurrences of the variant [should be noted] in the relevant circumstances.' However, defining the envelope of variation is by no means always straightforward in the case of non-phonological variables (see further 6.6.1)

### 2.3.3 *Data analysis*

An important tool which is widely used in variationist research for investigating variable linguistic phenomena is multivariate analysis. Various computerised versions of the variable rule program exist, but for the purposes of this study, the main analytic tool which was used was Goldvarb 2001.<sup>16</sup> Computerised versions of the variable rule program are particularly suited to the analysis of natural language data because they are generally able to make allowances for badly distributed cells (e.g. full cells versus near-empty ones; see Tagliamonte 2006a: 133).<sup>17</sup> It is for this reason that Tagliamonte (2006a: 130, 132-3) counsels against the use of other statistical procedures such as analysis of variance, or ANOVA, to analyse sociolinguistic data on the grounds that simple additive models are not suitable for situations where ‘the application frequencies are very different in different environments, or when there are a large number of different environments’ (Cedergren and Sankoff 1974: 337).

According to Bayley (2002: 129), the strength of variable rule analysis lies in its ability ‘to give precise and replicable measures of the strength of a wide range of contextual influences on the choice among variable linguistic forms. These conditioning factors may be either social or linguistic.’ An additional advantage of using programs such as Goldvarb is that they offer fast and efficient ways of organising and reconfiguring the data so that distributional patterns within the data can be carefully scrutinised before statistical analysis is undertaken.

In order for a distributional analysis to be carried out, the variants of a linguistic variable are coded for a number of different factor groups. Each factor group represents an aspect of the contextual environment (external or internal

---

<sup>16</sup> Goldvarb X, a later version of the variable rule program, was used when it became available in 2005.

linguistic) which is hypothesised to affect the choice of a particular variant. In this study, factor groups are largely constructed on the basis of constraints that have either been identified in previous analyses as operating on particular variables, or which could be inferred, on the basis of diachronic evidence, as having an effect on the variable in question. The rationale informing the choice of factor groups is described in greater detail in each of the chapters which focuses on key linguistic variables in the preadolescent corpus.

Each factor group, or set of ‘classificatory dimensions’ assumed to influence variation (Sigley 2003: 227), comprises a number of different factors which together constitute the set of possible values for a particular contextual environment. The coding schema that was employed for each factor group in each of the multivariate analyses was invariably elaborate in the initial stages of quantitative analysis in order to ensure maximum flexibility in the subsequent re-coding phases of the data ( see Tagliamonte 2006a: 200). As Young and Bayley (1996: 257) remark, this methodological flexibility enables analysts to approach the data with liberal rather than conservative hypotheses as it is easier to remove certain factor groups than it is to introduce new factor groups once the data have been coded.

Another prerequisite for statistical analysis, following the coding of contextual factors which are hypothesised to condition the choice of linguistic variants, is the identification of potential interaction effects in the data. This is particularly important in programs such as Goldvarb where statistical modelling is based on independent factor groups (Sigley 1997a: 252), and where there is no overt check for interaction effects incorporated into the variable rule program. Tagliamonte (2006a: 151) notes that interactions are often apparent when factor weights generated by the variable rule program are compared with corresponding proportions in the marginal results. Salient

disparities emerging from such comparisons can provide initial indications of potential interactions in the data. Furthermore, notable fluctuations in the factor weights assigned by the variable rule program to each factor as it proceeds from one level to the next in the step-up/step-down regression procedure offer additional indications of the presence of substantive interaction effects. The standard approach to addressing the problem of potential interaction effects in quantitative data is to systematically cross-tabulate every factor group with every other factor group (Young and Bailey 1996: 287; Tagliamonte 2006a: 182). This practice ensures that problems associated with poorly distributed cells, empty cells as well as possible dependence between factor groups are identified and resolved before conducting multivariate analysis of the data. Once potential interaction effects have been pinpointed, new models of the data can be constructed which eliminate the confounding influence of strong interaction effects.<sup>18</sup> However, successful reconfiguration of the data to resolve interaction necessarily depends on the accuracy with which such effects are identified in the first place.

In terms of the statistical analysis of the data, automated significance testing of factor groups is achieved by means of the step-up/step-down analysis. During the step-up analysis each factor group is added one at a time, beginning first with one factor group, then adding the next in an incremental process so that eventually a full model is constructed with all factor groups included. This procedure is then followed by the step-down analysis during which the program discards one factor group at a time until only one factor group remains. During each computational step of the analysis, the log likelihood is calculated as well as the factor weights for the factors in

---

<sup>18</sup> Programs such as Goldvarb are not without their limitations and have difficulty handling data which are essentially skewed (Tagliamonte and Temple 2005: 301, f.n. 25). Such difficulties can usually be overcome by removing factor groups; collapsing factors within a factor group; or collapsing factor groups together (Tagliamonte 2006a: 189). These decisions cannot be undertaken on an *ad hoc* basis, and need to be conducted in accordance with linguistically defensible principles.



a factor group. The log likelihood is a measure of the goodness of fit between the model and the data (see Bayley 2002: 128), with log likelihoods closer to zero reflecting better models than those that are further away from zero. The factor weights, expressed as decimals, provide a 'numerical measure of the strength or influence of each factor, relative to other factors in the same group, on the linguistic variable under investigation' (Bayley 2002: 126). Conventionally, factor weights are interpreted as follows: the closer the decimal number is to 1, the more favourable the effect is; conversely, the closer the number is to zero, the more disfavoured the effect is. A factor weight of 0.5 is claimed to neither favour nor disfavour the application value. Although this interpretation of factor weights is pervasive in the variationist literature, Tagliamonte (2006a: 145) observes that it is the relative position of factor weights with respect to each other which is of particular relevance with regard to the interpretation of the results.

Statistically significant factor groups are indicated in the output at the end of the step-up/step-down analysis. The factor groups in the best stepping-up and stepping-down runs of the analysis, which should normally be the same, are those which are significant at the 0.05 level. In this study, significant factor groups are highlighted in bold text in those tables which report multivariate analyses of the data. Non-significant factor groups, which are obtained from the first iteration of the step-down analysis (Tagliamonte 2006a: 252), are displayed in unbolded text.

Another important statistic that is reported in tables displaying the results of multivariate analyses is the input (also known as the input probability, the mean effect or the corrected mean). This figure is reported at the top of each table and is rounded to two decimal places. Paolillo (2002: 79) defines the input as 'the average frequency

of occurrence of the application value of the dependent variable.' This figure is helpful in indicating the overall probability of the dependent variable in the data.

In terms of the interpretation of statistical evidence generated by the variable rule program, the analyses discussed in this study make use of two additional indicators, the range value and the constraint hierarchy, which can both be used to refine the interpretation of the results. The range, which is calculated by subtracting the lowest factor weight from the highest factor weight in each significant factor group, provides an indication of the relative strength or magnitude of each significant factor group. The constraint hierarchy refers to the ordering of the factors from more to less within a factor group (see Tagliamonte 2002a: 752). Both the range and the constraint hierarchy can be used when comparing statistically significant evidence across data sets which may differ both in terms of size and in terms of the frequency of forms being investigated. In Chapter 3, for example, the range and the constraint hierarchy are particularly useful for exploring underlying grammatical similarities and differences between preadolescents and adolescents in their use of morphosyntactic variation in past *BE* forms.

Although statistical analysis is an integral part of this study, it is important to stress that programs such as Goldvarb are powerful heuristic tools for exploring quantitative data, and 'do not in and of themselves provide explanations about linguistic structure or the meaning of the social distribution of linguistic variants' (Bayley 2002: 130). Statistical evidence and numerical formalisms merely help to construct a foundation on which to build explanations of linguistic phenomena, and therefore need to be supplemented by careful qualitative analysis of the data.

In Chapters 6 and 8, which deal with discourse marker *like* and general extenders respectively, Goldvarb is not used to conduct multivariate analyses of the

data. Variationist analyses of discourse variables are notoriously complex (Tagliamonte 2006a: 83). In the case of discourse *like* examined in Chapter 6, it is relatively straightforward to identify actual occurrences of this marker in discourse, but far more complex to delimit where it could occur, but did not (see Tagliamonte 2006a: 83; and D'Arcy 2005a). As the focus in this chapter is on overt occurrences of *like* in discourse only, and excludes consideration of where it could have potentially occurred but did not, I restrict the investigation to a distributional analysis of this discourse feature. Given that Goldvarb is designed for binomial analysis (Tagliamonte 2006a: 160), it requires that both an application value (i.e. a variant defined as the outcome of a variable rule) and a non-application value (i.e. a context in which a variable rule does not apply) be specified. Thus, the program is not appropriate for the statistical analysis of overt occurrences of *like* only. Nevertheless, the program was still used to carry out distributional analysis of discourse *like*, as well as to cross-tabulate social and linguistic categorizations of the data.

Similar problems are encountered in the analysis of general extenders. In this case, the sheer number of variants, many of them nonce forms, makes statistical analysis via the use of the variable rule program problematic. In her study of general extenders in Canadian French, Dubois (1992: 194) resolves this problem by grouping variants into homogeneous classes based on the structural components of variant forms before conducting a multivariate analysis of the data. While such a solution is feasible for large data sets containing several hundred tokens, categorizing the small number of general extender tokens (N=148) in the preadolescent corpus into several groups would have resulted in a small number of tokens in each group. As Milroy and Gordon (2003: 168) note, statistical significance is generally less easily achieved in small samples which are additionally subject to greater random fluctuations. For these

reasons, Goldvarb was only used to facilitate distributional analysis of general extender variation in the corpus, and statistical analysis of various configurations of the data were subsequently conducted using the chi-square test.<sup>19</sup>

## 2.4 VARIATION IN THE PREADOLESCENT CORPUS

### 2.4.1 *Non-standard features*

The preadolescent corpus contains a wealth of non-standard features. Many of these non-standard forms are widespread in other urban varieties of English (see e.g. Cheshire 1982). According to Coupland (1988:35), the currency of certain grammatical variables in diverse geographical areas suggests that they are characteristic of British social dialects rather than exemplifying regionally circumscribed forms (see also Cheshire *et al.* 1993; Cheshire *et al.* 2005a). These common features (illustrated below in examples taken from the preadolescent corpus) include non-standard negative forms (a-d); *never* as a past tense negator (e); *them* as a demonstrative (f-g); unmarked adverbs (i.e. without *-ly*) (h); non-standard comparatives and superlatives (i-j); non-standard past tense forms (k-o); unmarked plurality (p); non-standard prepositional usage (q); and non-standard reflexives (r).<sup>20</sup>

(a) you *couldn't* see *none* of the audience [2M10/11]<sup>21</sup>

(b) he started kicking me in my legs so I *couldn't* take it *no more* [7M10/11]

---

<sup>19</sup> The chi-square test was also used to assess statistical levels of significance in various distributional configurations of the data relating to discourse marker *like* in Chapter 6.

<sup>20</sup> This list is intended to be illustrative rather than exhaustive. Further variables could also be added such as the 'new' modals (Krug 2000). In some cases, however, there is limited information about the regional or social distribution of non-standard features (see Britain *in press*).

<sup>21</sup> On the basis of an analysis of data in the *British National Corpus*, Anderwald (2002: 109-114) claims that multiple negation is subject to a regional cline in terms of frequency: it is most frequent in the south of England, less so in the Midlands, and least frequent in the north of England. In the preadolescent data, however, it is relatively rare. The frequency index for multiple negation (with the percentage of the non-standard form calculated in relation to the total number of standard and non-standard variants) is 15.8%. There is no significant gender distribution. For males, the frequency index is 16.13%, and for females, 15.56%.

- (c) she doesn't wanna do it she *don't* wanna do it, so just leave her alone  
[16F10/11]
- (d) what if they *ain't* a pedigree ? [26F7/8]
- (e) I knew they were going to do it but I *never* knew what time and I *never*  
knew they had all creepy crawly hands and everything [15F10/11]
- (f) I'm like one of *them* ski jumpers that jump off the thing [12M7/8]
- (g) then it was just *them* two left [5M10/11]
- (h) he just got out the car really *quick* [15M7/10]
- (i) I think they're *more naughtier* [23F7/8]
- (j) cos Lisa's *most populaarest* out of us all [15F10/11]<sup>22</sup>
- (k) they *done* the Ouija board again [28F7/8]
- (l) they *brung* me out after [29F7/8]<sup>23</sup>
- (m) I *seen* the first one [15M7/8]
- (n) and then he *run* away [18M7/8]
- (o) she *come* over and saw me [23F7/8]<sup>24</sup>

<sup>22</sup> Quirk *et al.* (1985: 462) rule out inflectional forms for trisyllabic adjectives, although Kytö and Romaine (1997: 342) report such forms as *frustratedest* and *beautifullest* in the *British National Corpus*.

<sup>23</sup> Some of these non-standard past forms conform to a postulated past tense schema characterised by the phonetic configuration /u/ followed by a nasal and/or a velar consonant (see Bybee and Moder 1983; Cheshire 1994). Hogg (1988: 38) has claimed that the tendency exhibited by strong verbs to evolve in the direction of a prototypical phonological shape containing /u/ suggests that this vowel can be seen as an 'ideophonic marker of past forms'.

<sup>24</sup> Poplack *et al.* (2002: 103) point out that non-standard *come* has featured in the English grammatical tradition since at least the late sixteenth century. Tagliamonte (2001: 43) notes that non-standard *come* exhibits high levels of non-standardness in relation to other strong verbs. This is also the case in the preadolescent data where it is the most non-standard strong verb, with a frequency index of 43.8%. Thus, contrary to reports that this is a recessive feature of non-standard British dialects (Tidholm 1979:147), it is clearly a robust variant in the corpus. In accordance with the widely attested gender-differentiated pattern (Chambers 2003: 116), preadolescent males use the non-standard variant (47%) more than females (40%). A preliminary multivariate analysis of the data for *come/came* variation revealed that grammatical person was significant, with first person singular subjects strongly favouring the non-standard variant, and third person plural subjects disfavouring it. The constraint ranking for grammatical person raises some interesting issues regarding the preservation of attested historical patterns of variation. In a study of strong verbs in English from Chaucer to Caxton, Long (1944:129, cited in Tagliamonte 2001: 51) notes that the form *come* occurs three times as frequently in the preterite singular than in other contexts. Tagliamonte (2001: 46) observes that in written English, *came* did not supplant *come* in the midland and southern dialects of English until the fifteenth century.

- (p) it costs one thousand *pound* [16M7/8]
- (q) we saw that a man had come *off of* his motorbike [16F10/11]
- (r) like he's just like describing *hissself* [4M10/11]

#### 2.4.2 *Selection of variables in the study*

In a small corpus of some 70,000 words, frequency was obviously of paramount concern in the selection of variables that were appropriate for quantitative analysis. Infrequent features such as non-standard negation were not considered eligible candidates for detailed quantitative study.

Aside from the intrinsic interest of studying language variation in the vernacular of preteen schoolchildren, an age cohort which has not been extensively subjected to detailed empirical linguistic analysis from a variationist perspective, the research focus on the participation of preadolescents in patterns of variation and change meant that it was important to include a selection of grammatical variables that have been implicated in trajectories of change in contemporary British English. For these reasons, the study concentrates on a detailed investigation of six key variables. These variables are described below:

##### (i) *Was/were variation*

- (a) so all my friends *was* saying, 'You shouldn't have... be... made up with her been friends with her.' [16F10/11]
- (b) we *was* like... pretending... pretending we *was* asleep on the sofa [14M7/8]
- (c) I *weren't* that injured [13M10/11]

Variation between *was* and *were* has been variously explicated in terms of analogical levelling or conjugation regularization (see Tagliamonte 1998a;

Tagliamonte and Smith 2000) as well as with reference to vernacular universals (Chambers 2004).<sup>25</sup> Chapter 3 situates variation between *was* and *were* in the preadolescent corpus in relation to vernacular patterns attested in other varieties of English as well as with regard to diachronic patterns and constraints that can be extrapolated from the available historical evidence. Quantitative analysis of the preadolescent data explores the internal and extralinguistic constraints which operate on *was/were* variation, and compares the results with data from adolescent speakers recorded in London.

(ii) *Relativization*

- (a) it was about these men [*who* had been eaten by werewolves] [27F7/8]
- (b) he was doing an interview with the people [*that* owned the bakery] [13M7/8]
- (c) it's a series [Ø I'm watching now] [7F10/11]

Variation between *wh*-relativizers, the relative marker *that*, and the zero relative marker is sufficiently frequent in the preadolescent corpus to permit quantitative analysis. In spite of the fact that this is one of the most 'studied areas of recent grammatical change' (Tagliamonte 2006b: 492), little is known about the internal contextual constraints which operate on this domain of variation in southeastern varieties of English, nor how regional norms in the southeast may differ from those elsewhere in Britain. Furthermore, as Diessel and Tomasello (2005: 882) point out, the majority of studies which are concerned with relativization in child language focus on experimental studies of comprehension rather than on acquisition, with the result

---

<sup>25</sup> According to Chambers (2003: 265), 'certain variables appear to be primitives of vernacular dialects in the sense that they recur ubiquitously all over the world.' Possible candidates for the status of vernacular primitives include recurrent patterns of non-standard subject-verb concord; conjugation regularization; and multiple negation. Chambers (2003) maintains that traditional diffusionist explanations for vernacular universals are implausible; instead, it is claimed that such features appear to be the result of 'natural outgrowths' of the language faculty (Chambers 2004: 128) which can offer important insights into the cognitive motivations underlying processes of language variation and change.

that there is less information about how variation in relativization strategies patterns in natural language data obtained from children.

Chapter 4 examines constraints on the variation between choice of relative marker in the preadolescent corpus, and explores whether the paucity of *wh*-markers, which is claimed to be a feature of spoken English (see Romaine 1982), is also characteristic of the preadolescents' vernacular usage.

iii) *Tense variation*

(a) and then a car was coming...behind

so *I tried* to get back on the path

but *it's hit* the ba...my back wheel

and *I've gone* off on to the road

and *scraped* all my arm [13M10/11]

Chapter 5 looks at tense alternation between the conversational historical present (the CHP), the simple past and the present perfect in narrative discourse. Most of the previous research on variation in tense-aspect morphology in narrative has focused on the discourse functions of the alternation between the simple past and the conversational historical present. However, it is not clear what discourse functions may be implicated by switches between the simple past and the present perfect, as illustrated in the example above. Lindstedt (2000: 371) claims that when a perfect construction can be used as a narrative tense, it has ceased to be a perfect. This raises the question of whether the present perfect in preadolescent narratives is making incipient inroads into the territory conventionally associated with the simple past. Does the use of the present perfect in foregrounded narrative clauses reflect a possible change that is underway in certain varieties of English, or can it be explained in



developmental or discourse-pragmatic terms? To address these issues, Chapter 5 builds on the foundational research conducted on the alternation between the simple past and the CHP in narrative (e.g. Schiffrin 1981; Wolfson 1982) in order to explore the structural and pragmatic motivations underlying tense switching between the simple past and the present perfect in preadolescent discourse. Comparisons are drawn with other varieties of English in which similar phenomena can be observed in adult populations, as well as with other variationist studies of temporal reference which have highlighted a discrepancy between normative accounts of distinctions between tense forms and their patterning in natural language data where such distinctions may be neutralized (Poplack and Turpin 1999: 161; Bybee and Hopper 2001:4).

iv) *Discourse LIKE*

- (a) after we went *like* to a disco [19F7/8]
- (b) there's *like* fireworks going up [14M7/8]
- (c) he started drinking *like* sugar water [3M10/11]

There are currently a number of discourse-pragmatic features which either appear to be gaining ground or are involved in ongoing change in several varieties of English (see Tagliamonte 2005; D'Arcy 2005a; Rickford *et al.* 2005). Discourse *like* is one such feature which has been specifically correlated with the language of the youngest generation, particularly adolescents, who are claimed to be spearheading the use of such discourse forms. Less is known, however, about how discourse *like* is used by preadolescent speakers in Britain, with some researchers (e.g. Miller and Weinert 1995) claiming that this discourse particle is not characteristic of preadolescent vernacular usage. Chapter 6 reassesses these claims by presenting a

distributional analysis of discourse *like* in preadolescent speech. Contrary to popular stereotypes which associate this discourse marker with late twentieth century youth, Chapter 6 situates discourse *like* in a diachronic context and looks at the existence of ‘embryonic variants’ (Trudgill 2002) which are possible precursors to modern discourse uses of *like*. Attention is also drawn to a number of cross-linguistic analogues to *like* in a range of languages, suggesting that such markers evolve for functional reasons. Furthermore, drawing on cross-varietal research, Chapter 6 attempts to show that discourse *like* is grammatically constrained, and not simply an unsystematic insertion into discourse.

v) *Variation between quotative markers*

(a) my dad *said*, ‘We’ll take you to the doctor’s ’ [25F7/8]

(b) he *goes*, ‘Can you do it please?’ [16M7/8]

(c) boys *say*, ‘Come here man,’ and everything like that

and girls are *like*, ‘I don’t... we’re not men.’ [15F10/11]

Recent research (Tagliamonte and Hudson 1999: 168) suggests that variation between quotative markers is a good place to examine the ‘burgeoning global mega trends of language change.’ Particular interest in the quotative system as a site for language innovation has been fuelled by the rapid spread in worldwide varieties of English of the new *be like* quotative which Labov (*personal communication* in Cukor-Avila 2002: 21-22) cites as ‘one of the most striking and dramatic linguistic changes of the past three decades.’ However, as Tagliamonte and D’Arcy (2004: 494) note, the age at which this variant is appropriated by speakers is not yet well understood. Chapter 7 analyses over 500 quotative contexts in the preadolescent data and examines the extent to which *be like* has made incursions into the children’s

vernacular. The focus in this chapter is on the quotative system in its entirety, which not only enables proportional comparisons to be made between the vernacular quotatives *be like* and *go* and the more neutral variant, *say*, but also facilitates contrastive analysis of the different pragmatic functions of quotative markers in preadolescent discourse.

vi) *General extenders*

(a) they've ate all their insides *and everything* [22F7/8]

(b) boys try to be all hard *and stuff* [20F10/11]

(c) I don't know if they were chasing us *or anything* [13M10/11]

General extenders such as *and stuff*, *and everything*, etc., are common in informal conversation in which 'lack of precision is one of the most important features of vocabulary' (Crystal and Davy 1975:111-12). Previous research (Overstreet 1999) suggests that these structures express a range of referential and interpersonal meanings, and depend on shared social knowledge and common cultural scenarios in guiding utterance interpretation. Chapter 8 focuses on a distributional analysis of these forms in preadolescent speech, and explores the intersection between age and gender in patterns of general extender variation. The investigation of variation between different general extender forms is also extended to a quantitative exploration of the distribution of particular variants in specific narrative sub-components.

## CHAPTER 3

### VARIATION IN THE PAST TENSE OF *BE*

#### 3.1 INTRODUCTION

Variation in the past tense of *BE* has been widely researched by a number of linguists working on an array of worldwide varieties of English (see e.g. Cheshire 1982a; Tagliamonte 1998a; Smith and Tagliamonte 1998; Britain 2002a; Pietsch 2005a; Pietsch 2005b; Cheshire and Fox 2006 on British English; Feagin 1979; Schilling-Estes and Wolfram 1994; Wolfram and Sellers 1999 on American English; Meechan and Foley 1994 on Canadian English; Eisikovits 1991 on Australian English; Hay and Schreier 2004 on New Zealand English). Examples of *was/were* variation in a number of vernacular varieties of English are given below:

- (1) cos I *were* going to smash him up, and he had to get Rick (Reading English, UK, Cheshire 1982a: 44)
- (2) they *was* best mates (York English, UK; Tagliamonte 1998a: 155)
- (3) we *were* all thegither... I think we *was* all thegither (Buckie English, Scotland; Smith and Tagliamonte 1998: 148)
- (4) it *weren't* me and it *weren't* Linda (Ocracoke English, North Carolina, US; Schilling-Estes and Wolfram 1994: 280)
- (5) there *wasn't* the distractions that there are today (Ottawa English, Canada; Meechan and Foley 1994: 64)
- (6) you felt like you *was* really in it (Sydney English, Australia; Eisikovits 1991: 250)

(7) *we was* always good friends (New Zealand English; Hay and Schreier 2004: 223)

The pervasiveness of variation in past *BE* in worldwide varieties of English has generated a number of claims relating to the apparent universality of this phenomenon in dialects of English. These claims have combined insights from many different areas of linguistics including the variationist paradigm, formal syntactic theory as well as from dialect typology (see e.g. Anderwald 2002; Tagliamonte 2002a; Adger and Smith 2005; Rupp 2005). For some researchers, morphosyntactic variation in past *BE* reflects analogically based levelling (Wolfram and Schilling-Estes 2003: 131), while for others, past tense *BE* variation is considered to exemplify vernacular primitives (see Chapter 2; and Chambers 2003: 266).<sup>1</sup>

Examples of what Chambers (2003: 266) terms as ‘default singulars’ (i.e. subject-verb non-concord) occur frequently enough in the preadolescent data to permit quantitative analysis. The following examples from the preadolescent data illustrate *was/were* variation in both positive and negative contexts:

- (8) *was* they crucified? [16M7/8]
- (9) Shaggy and Scooby Doo *was* running away [18M7/8]
- (10) he saw golf clubs that *was* in the house [4M10/11]
- (11) my brother *was* saying it *weren 't* his fault [16F10/11]
- (12) we *was* playing a game [19M7/8]
- (13) they *was* chasing us [13M10/11]
- (14) these bad guys *was* trying to... catch it [15M7/8]
- (15) my socks *was* already pulled down yeah [18M7/8]

---

<sup>1</sup> There is evidence that the phenomena discussed here have cross-linguistic counterparts. Nevalainen (2006: 365-6) discusses one such case in vernacular Finnish in which the verb *olla*, used as a copula, is levelled to the singular third-person forms *on* ‘is’ (present) and *oli* ‘was’ (past) with the third-person plural pronoun *ne/he (they)* and plural NP subjects. These ‘default singulars’ are encountered in Finnish in both existential and non-existential contexts.

- (16) he *were* all like this <PERFORMS GESTURE> [15F10/11]
- (17) I *were* pretending I *were* superman [12M10/11]
- (18) I *was* following my dad but I *weren't* cos I couldn't see... see anybody [19F7/8]
- (19) it ended up we *wasn't* talking to each other for about two weeks [3M10/11]
- (20) I had to go to hospital to see if there *was* anything wrong with it but there *weren't* anything wrong with it [19F7/8]
- (21) there *was* these sharks [24F7/8]
- (22) there *was* these three dogs [26F7/8]

In terms of the broad conditioning effects constraining variation in past *BE*, a number of salient tendencies have been identified based on comparative sociolinguistic research focusing on non-standard *was* in a range of vernaculars used in geographically disparate locales. Firstly, existential constructions (see 21-22 above) appear to strongly favour the use of non-standard *was* (Britain 2002a: 19-20). The other major constraint which is reported to have conditioned variation in past *BE* diachronically is the NP/PRO constraint (Bailey *et al.* 1989: 294) or the Northern Subject Rule (see examples 10, 14 and 15 above; for further details of the operation of this constraint in older varieties of English, see Murray 1873: 211-2; Ihalainen 1994: 221).<sup>2</sup>

Chambers (2004: 133-4) claims that subject types which promote the use of non-standard *was* can be graded hierarchically, with expletive *there* subjects favouring non-standard *was* the most, and the third person plural pronoun *they* being

<sup>2</sup> This constraint was introduced and discussed in Chapter 1, footnote 3. It is also important to point out that the Northern Subject Rule has scope beyond past *BE* (see further Britain 2002a: 20; Pietsch 2005a).

the most inhibiting subject type. While expletive *there* and the pronoun *they* are frequently cited as the subject types which embody the most predictable conditioning effects constraining variation in past *BE* across a range of dialects, Chambers (2004: 133) concedes that there is cross-dialectal variation in the grammatical subjects which occupy intermediate positions in this hierarchy. According to Tagliamonte (1998a: 158), the findings of contemporary research on variation in past *BE* suggest that the hierarchy of constraints embodied by subject types is highly consistent across varieties, with the second person singular *you* favouring *was* after existentials, followed by NP plurals, the first person plural *we*, and finally *they*.

Viewed in conjunction, these salient findings generate a number of research questions that are pertinent to the nature and extent of variation in past *BE* in the preadolescent corpus:

- First, which environments favour the use of non-standard patterns of subject-verb concord in the preadolescent data, and how do these compare with the results reported for other dialects?
- What evidence is there that the Northern Subject Rule (originally a feature of northern varieties, although attested in other southern British dialects; see Godfrey and Tagliamonte 1999) has diffused to southeastern dialects, and is present in the vernacular of preteen speakers? <sup>3</sup>
- How can patterns of variation in the preadolescent corpus be contextualized in relation to the findings reported by Anderwald (2001), and Cheshire and Fox (2006) for *was/were* variation in London English?

---

<sup>3</sup> Some researchers (e.g. Montgomery 2001: 213) have recently queried the extent to which northern Britain can be treated as a homogenous area with regard to *was/were* variation. Citing historical evidence from West Yorkshire, Montgomery (2001: 212) claims that the NP/PRO constraint was disappearing for the past tense copula by the late eighteenth century.

- Finally, in what way, if at all, are gender and age implicated in preteen patterns of variation in past *BE*? Do the findings for the preadolescents examined here support Anderwald's (2001: 14) tentative hypothesis that younger speakers (under 35) in London are restructuring patterns of *was/were* variation, resulting in a remorphologized paradigmatic distinction (Schilling-Estes and Wolfram 1994), with *was* used in positive contexts and *weren't* in negative ones?

In order to address these questions, I first examine variation in past *BE* from a diachronic perspective in order to explore the historical precursors of contemporary patterns of non-standard usage, and to investigate the possible persistence of historical patterns of variation in contemporary preadolescent speech.

### 3.2 DIACHRONIC PERSPECTIVES ON *WAS/WERE* VARIATION

Variation in past *BE* is a longstanding phenomenon in the history of English. Lass (1993: 170) notes that from an historical perspective, the verb *BE* is not a 'single verb' but an assortment of semantically related paradigm fragments in other Germanic languages as well as in Old English. Similarly, Pyles and Algeo (1993: 127) refer to *BE* as a 'badly mixed up verb'. Even in modern standard English, *BE* retains a person and number distinction in the past tense that does not exist for any other standard English verb (Tagliamonte and Smith 1999: 8).

A cursory examination of past tense *BE* variants in earlier varieties of English furnishes ample evidence of widespread geographical variation in the distribution of variant forms. The following table, adapted from Mossé (1952: 84), gives an overview of the forms used in four dialects of Middle English.



TABLE 3.1

PAST TENSE *BE* FORMS IN DIALECTS OF MIDDLE ENGLISH

Standard English	North	West Midland	East Midland	South
I	<i>was, wes</i>	<i>was</i>	<i>was</i>	<i>was</i>
you	<i>was, wes</i>	<i>wore</i>	<i>were, wast</i>	<i>weore</i>
he, she, it	<i>was, wes</i>	<i>was</i>	<i>was</i>	<i>was</i>
we, you, they	<i>wer, war, wes</i>	<i>woren</i>	<i>were(n)</i>	<i>weore, waere</i>

The variant forms attested in the south exhibit parallels with the modern standard English system. In the north, on the other hand, there is a more widespread use of *was*, particularly in the second person singular. The northern provenance of the form *you was* is further corroborated by evidence from the Oxford English dictionary. Nevalainen (2006: 356) points out that the earliest citations of this variant in the OED are all attributable to northern writers:

(23) *you was* our drowrie and our dayes darling (OED; c.1450, Henryson  
[Scottish poet] *Mor. Fab.* 19)

(24) in dud frese ye *was* schryned with better frese lynyd (OED; a1529,  
Skelton [Yorkshire-born poet] *Poems agst. Garnesche* 46)

[examples cited in Nevalainen 2006: 356]

Nevalainen (2006: 360) adds that *you was* gained momentum in dialects of English in the eighteenth century, which suggests that there has been some reweighting in the

grammatical constraints operating on past *BE* over time (Nevalainen 2006: 356).<sup>4</sup>

Smith and Tagliamonte (1998: 108) point out that although the available historical evidence shows that the use of *was* in the second person singular is amply attested in northern dialects, the extent to which *you was* was employed in southern dialects is less easy to substantiate. A similar difficulty obtains in the case of efforts to substantiate the time-depth of *we was* in southern dialects, although Nevalainen (2006:360) notes that in the *Corpus of Early English Correspondence* (which includes material from 1410 to 1681) *we* is the only plural pronoun which occurs with any frequency with *was*.<sup>5</sup>

As far as northern dialects are concerned, some researchers have argued that there appears to have been an expansion in the use of *we was* over time. Smith and Tagliamonte (1998:117) argue that although *we was* was not common in the Middle English period, the fact that it now appears frequently in Buckie English (spoken in the northeast of Scotland) suggests that it may be an extension of an older pattern.

With regard to other subject types, Nevalainen (2006: 360) observes that there are few examples of the variant *they was* in the *Corpus of Early English Correspondence*. This finding accords with contemporary variationist research which, as noted above, has consistently found the third person plural pronoun *they* to be the subject type which is least likely to co-occur with non-standard *was*. Nevalainen

---

<sup>4</sup> Petyt (1985: 237, f.n.147) claims that the use of *you was* when addressing one person was common from the sixteenth century until the eighteenth century. Noah Webster (1789/1967: 233-74) comments that '*you* is commonly used with the plural of the verb *be*, *you were*; in conversation, it is generally followed by the singular, *you was*.' Strang (1970: 140) suggests that in the seventeenth and eighteenth centuries, the use of the contrastive forms *you is* (singular) and *you are* (plural), and, by extension, *you was* (singular) and *you were* (plural), was motivated by changes in the pronominal system which culminated in the loss of *thou*. However, as Strang (1970: 140) notes, the use of forms such as *you is* /*you are* to restore the number distinction formerly expressed by *thou* and *ye* was restricted to informal and non-literary usage, and eventually attracted censure from prescriptivists.

<sup>5</sup> The *Corpus of Early English Correspondence* contains written material from four areas in particular: the Court, London, East Anglia, and the North (Nevalainen 2006: 357). However, Nevalainen (2006) does not provide an analysis of *we was* according to these regions, so it is hard to ascertain the extent to which *we was* is present in older varieties of London English and how the frequency of this variant varied from region to region.

(2006: 361) does, however, cite an example from a London merchant writing in the sixteenth century who uses non-standard *was* with a plural noun phrase:

- (25) I dyd mystake the prices of wyne, when I showed you at my last being  
at Dene that the prices *was* but iij li the tonne (1546 JJOHNSON 604)

At this juncture, it becomes important to consider example (25) cited above within the broader perspective of the extent to which the Northern Subject Rule operated in historical varieties of London English. Recall that this variable rule is associated with the contrastive use of *was* with full NPs as opposed to the inhibiting context of the third person plural pronoun (the NP/PRO constraint), as well as being associated with contexts where the subject and the verb are non-adjacent. The operation of the Northern Subject Rule in older varieties of English is illustrated in the following examples cited in Visser (1963: 72):

- (26) þaire wordes *was* for noght (c. 1352 Minot War Ballads)
- (27) his wordes *was* not byleved (1523-5 Ld. Berners, Froiss. I, 81)
- (28) many great conflyctes *was* betwene them ( 1529, John Rastell,  
Pastyme of Pl. 172)
- (29) þei souȝt alle so serliche þurh cites & smale townes in wodes & alle  
weies þat *was* þer a-boute (c. 1350 Will. of Palerne 2149)
- (30) they toke shyppyng and sayled to Dover and *was* there by noone  
(1523 -5, Ld. Berners, Froiss. V. 357 )

Although the Northern Subject Rule appears to be historically associated with northern English dialects and the dialects of Scotland (McCafferty 2004: 52; Mustanoja 1960: 481-2), there is some evidence that over time, it came to operate over a much wider geographical area. Tagliamonte (2002a: 741) claims that the association of *was* with plural NPs appears to have spread to the south in the sixteenth

and seventeenth centuries.<sup>6</sup> Similarly, Britain (2002a: 36) reports that the Northern Subject Rule is operative in the vernacular of older speakers born in the northwest of the Fens in eastern England, and Godfrey and Tagliamonte's (1999) study of verbal *-s* in Devon in the southwest of England revealed statistically significant effects associated with the operation of the Northern Subject Rule.

Evidence for the operation of the Northern Subject Rule in historical varieties of London English, particularly with regard to past *BE*, is sparse, although the historical influx of migrants from the north of England to the capital lends some weight to the possible spread of the Northern Subject Rule to London in the late fifteenth and sixteenth centuries. Klemola's (1996: 185-191) research is particularly instructive in this connection as it suggests that the period between 1500 and 1800 witnessed substantial migration to London from the northern counties of England. Wright (2002: 244) contends that the Northern Subject Rule had permeated the London vernacular by the late 1500s, and supports Schendl's (1996) claim that the speech of migrating northern traders was an important vector for the transmission of the Northern Subject Rule to the speech of Londoners.

An early source of information about patterns of verbal agreement in the language of Londoners can be found in the Cely Letters written between 1472 and 1488 by a family of wool merchants resident in the capital. Bailey *et al.* (1989: 290) cite evidence of constraints associated with the Northern Subject Rule in these letters. Example (30) below from Bailey *et al.* (1989: 290) shows the effects of the non-

---

<sup>6</sup> Tagliamonte and Smith (2000: 154) cite Visser (1963:71-2), who reports a large number of examples of *was* with plural nouns in the sixteenth and seventeenth centuries, suggesting that this pattern was pervasive throughout Britain. Tagliamonte and Smith (2000: 163) further claim that in contemporary dialects the 'favoring effect of plural NPs [...] is attested all over Britain.' However, the contrastive use of *was* with third person plural noun phrases versus plural pronouns is not replicated in the findings for all varieties (see Britain 2002a: 36).

adjacency of the verb and the subject on variable patterns of verbal agreement in a coordinated VP structure:

- (30)   owr father and mother wher in good heyll and *sendys* yow ther  
          blessyngs

Although the writers of the Cely letters may not have necessarily been born and raised in London (Bailey *et al.* 1989: 287), the evidence furnished by this data source provides valuable insights into the fluctuating norms characteristic of dialect mixing in London in the past.

There is also textual evidence that the Northern Subject Rule affected patterns of verbal concord in the speech of prominent Londoners in the Early Modern English period. Schendl (1996: 150) cites the following examples from the writings of Elizabeth I, which illustrate some of the variable effects associated with the Northern Subject Rule:<sup>7</sup>

- (31)   for if neither they can doo that they *promise & wantes* greatest good  
          (Elizabeth, *Boethius* 48. 11)
- (32)   for wicked men *giues* this good turn to dignitie, that they *spot* them  
          with their own infection ( Elizabeth, *Boethius* 50. 23)

Nevalainen and Raumolin-Brunberg (2003: 169) comment that although such examples indicate that the Northern Subject Rule might have been operative in southern Early Modern English, including the writings of high status individuals such as Elizabeth I, the nature of this evidence is sporadic, indicating that the Northern Subject Rule never attained the supralocal status enjoyed by other features of northern origin such as the verbal suffix *-s*.

---

<sup>7</sup> As shown in examples 31 and 32 above, the Northern Subject Rule affects other verbs, although Pietsch (2005b: 152) claims that those dialects which exhibit patterns that conform to the Northern Subject Rule tend to reflect the effects of the constraints of this rule more frequently with *BE* than with lexical verbs.

Further evidence that the Northern Subject Rule percolated down the social hierarchy affecting the speech of ordinary Londoners in the sixteenth century is discussed in Wright (2002). Wright's (2002) detailed examination of the language used in London prisoners' depositions in the Bridewell Court Minute Books reveals that a variable NP/PRO constraint operated in third person present indicative forms, with noun phrases triggering verb forms terminating in '-s' or 'th', or zero if the verb was preceded by a pronominal subject, as illustrated in the following example taken from Wright (2002: 252):

- (33) Melcher Pelse sayeth that John Thomas and his wife are bawdes they  
*dwell* in Seathing lane very many marryners & other Englishe men  
lewdly *resortes* thither (Bridewell records, fo 124V, 28 Dec. 1576)

In brief, then, a synthesis of the available textual evidence from historical sources suggests that at least some of the constraints associated with the Northern Subject Rule appear to have diffused southwards and may have been operative in the speech of Londoners, at least from the Early Modern English period, although the extent to which the constraints associated with this rule were productive in past *BE* in older varieties of London English requires further substantiation. However, Nevalainen's (2006) multivariate analyses of the probability of *was* with plural subjects in three dialect regions (the North, East Anglia, and London) between the early fifteenth and late seventeenth centuries sheds some light on the processes of diachronic modification that non-standard *was* has been undergoing over the past several centuries. Whereas Nevalainen (2006: 362) found *was* to be positively favoured in the North and with plural NP subjects in the earliest sub-period examined (i.e. 1440-1519), in the last sub-period, 1640-1681, Nevalainen (2006: 363) reports that region was not selected as a significant factor, and that in contrast with the

earliest sub-period, plural NP subjects were no longer found to favour *was*.<sup>8</sup> Although Nevalainen (2006) does not present separate multivariate analyses of historical data for each geographical region, her findings are nevertheless important because they indicate that the conditioning factors operating on variation in past *BE* have been subject to change in the past several centuries.

A second major finding emerging from Nevalainen's (2006) analysis of historical data relates to the stability of existential subjects as a context which is particularly propitious to the use of non-standard *was*. The fact that there is minimal fluctuation in the factor weights for this subject type over time in the data discussed by Nevalainen (2006) is testimony to the strength of this conditioning effect in the history of English. As Britain and Sudbury (2000: 5) remark, non-agreement in existential constructions is a variable with considerable time-depth (see Meechan and Foley 1994; Tagliamonte 1998a; Denison 1998). According to Pietsch (2005a: 156), the absence of agreement in clauses beginning with *there* is attested as far back as the Old English period (Visser 1963: 62). Historical examples of variability in agreement in existential constructions in both past and present contexts are given below (all examples are from Visser 1963: 73-74):<sup>9</sup>

(34) *þær wæs syx hund manna* (The Blickling Homilies, 203, 27 [971 A.D.]

'there was six hundred men'

(35) *and þere was in þat tyme many gode holy men & holy heremytes* (c 1400

Mandev. 51, 35)

---

<sup>8</sup> In contrast to extensive reports of non-standard *was* in historical records, there is very limited information about non-standard *were* in earlier varieties of English. Forsström (1948: 228) reports the use of *were* with a singular subject, and Nevalainen (2006:360) provides the following example from a letter written by a London merchant: *Mr Hatten youre neighbour, who taried to speke with some of the same clarkes (for owne of them were then at home)...* (1545 JJOHNSON 400).

<sup>9</sup> The tendency for verbs to exhibit lack of agreement when a plural subject is in postverbal position in non-existential constructions has also been diachronically variable for centuries, as illustrated in the following example cited in Visser (1963: 73): *with rose and swete flores was strawed halles and bouris* (c1350 Alisaunder 1025).

(36) *there was* many Dukes, Erles, and Barons (1533 Ld. Berners, Huon 2, 22)

(37) *there's two crowns* for thee, play (a1592 Marlowe, Jew Malta IV, v)

(38) *there is some things* I can't resist ((1849-50) Thackeray, Pendennis II,

316)

In contrast with the variation in the factor weights for other subject types that Nevalainen (2006) examined, the robustness of existentials as a potent conditioning factor for non-standard *was* over a considerable time period corroborates recent claims that variable agreement phenomena with existentials constitute an exceptional case ( see further 3.3.3 below; Pietsch 2005a: 156; Wilson and Henry 1998).

To summarise, the overview of historical variation in past *BE* allows us to extrapolate the following conclusions: first, written records provide ample exemplification of variation in past tense *BE* forms, with variation correlating to some extent with geographical location. Second, the nature of the grammatical subject, and syntactic configuration play pivotal roles in conditioning variation in past *BE*. Third, recent historical research by Nevalainen (2006) suggests that there has been some reweighting of contextual effects conditioning variation in past *BE*, particularly in connection with certain subject types, although the use of non-standard *was* in expletive *there* constructions containing a postverbal plural NP has been a consistently salient trend for several hundred years.

The diachronic picture of variation in past tense *BE* forms constitutes an essential backdrop to the synchronic analysis of the preadolescent data discussed below. By grounding the synchronic data within a broader diachronic perspective, it becomes possible to investigate which synchronic variants may be remnants of older patterns that have persisted in present day varieties (see Labov 1980: xvii; Tagliamonte 2006b: 478). By the same token, departures in synchronic data from



historically attested patterns afford some tentative insights into evolving trends in past tense *BE* variation.

### 3.3 SYNCHRONIC PERSPECTIVES ON *WAS/WERE* VARIATION

#### 3.3.1 *Ecological factors: Urbanity, mobility, and isolation*

Chambers (2004) specifically addresses some of the correlations between variation in past *BE* and urbanity and mobility with reference to worldwide varieties of English. Of crucial importance in the emergence of patterns of variation is the role played by ecological and historical factors. According to Tagliamonte and Smith (1999: 21), the more isolated and separate a variety is from mainstream developments, the greater the likelihood of levelling to *was* (see also Tagliamonte 2002a: 734). Similarly, Chambers (2004: 138) makes a link between urban complexity and the frequency of non-standard *was*, and contends that sizeable urban settlements exhibit grammatical patterns which are closer to standard English in contrast with the greater non-standard usage attested in socially isolated communities. In support of this claim, Chambers (2004) cites findings reported for Tristan da Cunha English, a variety, which, until relatively recently, evolved in a context of isolation from other varieties of English. In this dialect, rates of levelling to non-standard *was* are reported to be particularly high (see Schreier 2002; Chambers 2004: 138, Figure 2). According to Schreier (2002: 92-5), levelling to *was* in past *BE* in Tristan da Cunha English is, to some extent, a structurally inherent process which has been shaped by the ecological particularities of the island of Tristan da Cunha, including its unusual settlement

history and its absence of interaction with standard English.<sup>10</sup>

In brief, a comparison of rates of levelling to non-standard *was* in a variety of locations reveals that ecological factors may conspire with language-internal tendencies in shaping the evolution of patterns of variation in past *BE*.

### 3.3.2 *Social factors: Age, gender and education*

Tagliamonte (1998a: 175) points to the importance of extralinguistic factors in conditioning patterns of *was/were* variation. *Was/were* variation has been found to be more frequent among working-class speakers than middle-class speakers (see also Feagin 1979; Eisikovits 1991). Non-standard patterns of variation in past *BE* have also been found to correlate with speakers' educational backgrounds, with less educated speakers making greater use of non-standard variants. Tagliamonte (2002a: 742) argues that the use of non-standard *was* may be affected by contact with prescriptive norms. For example, in their study of *was/were* variation in Samaná English, in which levelling to non-standard *was* is highly prevalent, Tagliamonte and Smith (1999: 14) report that the speaker who had by far the largest number of occurrences of standard *were* occurrences in her speech was the speaker who had the

---

<sup>10</sup> My use of the term 'levelling' in the discussion above with regard to the use of non-standard *was* in several worldwide varieties of English such as Tristan da Cunha English refers primarily to processes of conjugation regularization that are found in varieties of English which, for geographic or social reasons, are often less permeable to mainstream norms, and tend to preserve older non-standard features that are able to perdure in the absence of contact with exogenous normative influences associated with the standard language (see Milroy and Milroy 1991). In sociolinguistic studies of grammatical change (see e.g. Britain 2002a), the term 'levelling' often has broader ramifications, and is discussed in relation to dialect contact phenomena, specifically in relation to koinéisation processes which additionally involve mixing and simplification. Thus, in Britain's (2002a) research on *was/were* variation in the east and southeast of England, which is discussed at several points in this chapter, levelling is viewed as a diminution in marked or minority forms in situations of dialect contact, 'where the number of variants in the output is dramatically reduced from the number in the input' (Britain 2002a: 16). Britain (2002a) further invokes the notion of reallocation to explain structural changes in past *BE* which result from dialect mixing. These changes appear to involve the refunctionalisation of allomorphic variants of past *BE* which survive levelling processes, resulting in the evolution of new linguistic uses (see also Kerswill 2002).

most education. Exposure to literacy and exogenous normative pressures therefore appears to have an inhibiting effect on the use of non-standard *was*.

Age too is implicated in patterns of variation and change in past *BE*, as we shall see below (see also Tagliamonte 1998a; Britain 2002a; Britain and Sudbury 2000). In several cases, age and gender interact in accelerating change. For instance, Tagliamonte (1998a: 156) notes the role of young women in giving impetus to the use of non-standard *was* and *weren't* in specific syntactic environments.

In other cases, however, there are differences across communities in the role played by gender in conditioning patterns of *was/were* variation. Smith (2000: 70) reports that in Buckie English, there was no statistically significant difference in younger males' and females' use of non-standard *was* in contexts of standard *were*. In the light of previous research suggesting that men used stigmatized variants of stable sociolinguistic variables more than women (see e.g. Labov 1990), Smith (2000) interprets the absence of a statistically significant difference between the sexes as evidence that non-standard *was* is not a stigmatized variable in this group. A similar conclusion is arrived at by Tagliamonte and Smith (1999: 15) for Samaná English, although this time, gender was found to exert a statistically significant effect on the use of *was*, with females favouring its use, whereas men disfavoured it.<sup>11</sup>

Cross-variety comparisons highlight the need to take into consideration social embedding as well as internal constraints which operate on *was/were* variation. The complex interplay between external and internal factors is important in tracking recurrent trends in *was/were* variation across dialects, as well as identifying where

---

<sup>11</sup> Tagliamonte and Smith (1999: 15) infer from this finding that non-standard *was* is not stigmatised in Samaná English because the gender-related effect operates 'in the opposite direction to what would be expected if *was* were a negatively evaluated form.' However, Rickford (2006:131) criticises this line of interpretation on the grounds that 'it certainly is not the case that anytime women appear to favour a variant we can infer that the variable is *not* socially stigmatised in their speech community' (see also Escure 1991 on the interplay between gender roles and linguistic variation in Belizean Creole).

linguistic constraints are susceptible to social manipulation, which may lead to dialect differentiation (see Horvath and Horvath 2003: 168).

### 3.3.3 *Existentials: A case for special consideration?*

Both Tagliamonte (2002a:162) and Pietsch (2005a: 156) agree that non-standard *was* in existential contexts constitutes ‘a special case’ as far as *was/were* variation is concerned. Recall that high rates of non-standard *was* in existentials is a longstanding phenomenon with a heritage that extends back to the Old English period (Visser 1963: 62; Tagliamonte 1998a: 157).

Tagliamonte’s (1998:175) study of *was/were* variation in York English highlights the ‘dramatic distinction between existential constructions and all other grammatical persons’. Analogous findings are reported in Eisikovits’s (1991) study of variation in the speech of 40 Sydney adolescents, with the frequency of occurrence of non-standard *was* in existentials attaining an overall rate of 88.9%. Eisikovits (1991: 244) notes that the prevalence of non-standard *was* in existentials clearly differentiates this environment from other contexts in her data where the rate of non-standard *was* usage is markedly lower. Britain (2002a: 28) also observes that in the speech of certain Fenland dialect users, non-standard *was*-levelling in plural existentials behaves ‘radically differently’ from other contexts. Similarly, in a cross-variety comparison of patterns of *was/were* variation in a number of different varieties located in North America and Britain, Tagliamonte (2002a: 745) reaches the conclusion that high rates of levelling to non-standard *was* in existential contexts is a pan-community effect.

Other studies of contemporary dialects confirm that whereas rates of non-standard *was* with non-existential subjects fluctuate in apparent time, rates of non-

standard *was* in existentials remain relatively stable. For example, Smith (2000: 64) reports that although there was a gradual decrease in the use of non-standard *was* in Buckie English across three generations, rates of *was* usage in existentials remained high irrespective of the age and gender of speakers.

In spite of the apparent stability of non-standard *was* in existentials reported in previous studies, some researchers who have studied patterns of variation in present and past *BE* in existential constructions claim that this syntactic context is a site of change in progress. Britain and Sudbury's (2000:13) comparative study of singular verbs forms in plural existential contexts based on data from Falkland Island English and New Zealand English indicates a gradual progression to higher rates of non-concord in existential constructions when viewed from an apparent-time perspective, with the younger generation leading this change. Tagliamonte (1998a: 181) also raises the possibility of change in progress, and notes that non-standard *was* in existential constructions is favoured by the youngest generation (20-30) in York English, with women leading this change.

The fact that levelling to non-standard *was* appears to have evolved rather differently in existential contexts as opposed to non-existential contexts is highlighted in Hay and Schreier's (2004) recent research based on New Zealand English. Hay and Schreier (2004: 233) note that although both existential and non-existential contexts showed a steady decrease in the use of non-standard *was* in the early history of New Zealand English, possibly in response to contact with exogenous standardizing influences, singular concord in plural existential contexts subsequently increased, whereas non-existential contexts continue to exhibit negligible use of non-standard patterns of concord. Variable concord phenomena in existential and non-existential contexts have not followed parallel trajectories in New Zealand English, and do not

appear to be subject to similar linguistic constraints, which leads Hay and Schreier (2004: 229) to conclude that they are best studied as two separate phenomena. Rupp (2005: 260) also alludes to the possibility that non-standard concord in existentials and non-existentials are essentially separate phenomena, although she acknowledges the possibility that non-standard concord may simply be more 'advanced' in expletive *there* sentences than with other subject types. Pursuing a similar line of argument to Hay and Schreier (2004), Pietsch (2005a: 156) justifies treating variable patterns of concord in existential and non-existential contexts as separate on the grounds that these environments have different sociolinguistic status in dialects of English.

In their study of variable patterns of concord in existentials in Canadian English, Meechan and Foley (1994: 82) claim that non-concord is the norm. Furthermore, in contrast to non-standard concord in non-existential contexts, non-standard agreement in existentials has also been found to encroach to a limited extent on written English (see Martinez Insua and Palacios Martinez 2003: 268). Lack of agreement in existentials also occurs relatively frequently in informal 'educated' speech (Quirk *et al.* 1985: 1405). Tagliamonte (1998a: 184) observes that even relatively well-educated female speakers in York show near categorical use of non-concord in past existential constructions.

Additional insights into the nature of *was/were* variation in existential contexts can also be gained from considering the role of these constructions in discourse management and the sequential organisation of information. Breivik (1990: 153; 1999: 10) notes the use of the contracted form *there's* as a presentative formula preceding both singular and plural noun phrases. Cheshire (1999b: 71) argues that expletive *there* constructions with singular concord are prefabricated phrases that can

be usefully deployed in conversation for introducing new information (compare French *il y a*; German *es gibt*; Spanish *hay*)<sup>12</sup>.

Corroborative evidence that *there is/there was* may be evolving into a generalised presentative can be found in Walker (2007). On the basis of a study of patterns of variable agreement in plural existentials in spoken Canadian English, Walker (2007) reports that *there's* appears to have become a lexicalised construction (see also Schilling-Estes and Wolfram 1994:285). Furthermore, Walker (2007) suggests that *there's* may actually be grammaticalizing as part of a larger morphosyntactic string involving particular modifiers (e.g. *there's a lot (of)*; see Bybee 2003: 603 ).<sup>13</sup>

For Henry (2002: 269), the widely attested pattern of high rates of non-standard concord in existentials is not a question of 'how there is agreement with a post-verbal noun phrase, but how there is optional agreement.' Various solutions have been proposed within formal linguistic theory to the problems raised by concord in existential contexts, particularly in relation to the assignment of case and the control of agreement (Britain and Sudbury 2000: 9; see further Meechan and Foley 1994).

One of the major implications of syntactic theory for studying patterns of *was/were* variation in existential and non-existential contexts is that what are commonly assumed in variationist analyses to be related environments for variants of the same variable are treated within existing syntactic frameworks as syntactically remote (Cornips and Corrigan 2005b: 9). For example, Wilson and Henry (1998: 11)

---

<sup>12</sup> Henry (2005: 118) notes that the effect of the position of the subject NP on agreement patterns is attested in other languages such as Arabic, which has number agreement where the subject NP precedes the verb, but not in those cases where it follows it. Nevalainen (2006: 366) points out that singular concord is the only option with existentials in both vernacular and standard Finnish.

<sup>13</sup> Further evidence that expletive *there* constructions with singular concord may be grammaticalizing as part of a longer morphosyntactic string can be found in Henry (2005: 115-16), who notes that for some speakers, non-agreement is permissible when the plural postverbal subject is preceded by a quantifying expression such as *lots*, but is not permissible when there is no number or quantifier before the postverbal subject.

point out with reference to the Principles-and-Parameters version of generative grammar that singular concord in expletive *there* sentences would be explained in relation to a different parameter setting than that which would be invoked to explain singular concord in other contexts. Wilson and Henry (1998: 12) additionally note that because levelling to *was* in plural existentials is almost categorical or obligatory for some dialect speakers who may exhibit very low rates of levelling with other grammatical subjects, the inclusion of existential contexts within the envelope of variation may potentially obscure or distort the interpretation of community patterns of variation and change in the use of *was /were* in other environments.

Adopting a different line of argumentation based on the discourse-pragmatic functions of the existential *there* construction, Cheshire (1999b: 72) suggests that it might be more appropriate to consider *there's* and *there was* as variants of topic-introducing constructions rather than as part of a broader analysis of patterns of variable verbal concord. I return to this issue in section 3.4 below.

### 3.3.4 Vernacular Pattern 1

Cross-variety comparisons of *was/were* variation have revealed that there are two dominant patterns attested in English dialects. The first, and apparently the most common pattern, is for the use of *was* to generalize across grammatical person in both positive and negative contexts, giving rise to forms such as '*They was, wasn't they?*' (see Britain 2002a: 19). This pattern is attested in a number of vernacular Englishes, including African American Vernacular English (Labov *et al.* 1968); Alabama English (Feagin 1979); Samaná English of the Dominican Republic (Tagliamonte and Smith 1999); as well as Buckie English in Scotland (Smith 2000). In an overview of *was/were* variation in non-standard British English, drawing on data in the *British*



*National Corpus*, Anderwald (2001: 11) claims that *was*-generalization principally occurs in three dialect areas: Scotland and Ireland, and the northwest Midlands.<sup>14</sup> Furthermore, Anderwald (2001: 10) notes that for 11 out of 18 areas tested, *was*-generalization is the predominant option in clauses with positive polarity.

### 3.3.5 *Vernacular Pattern 2*

The second major attested pattern involves variable levelling to *was* in positive polarity clauses, with variable levelling to *weren't* predominating in negative polarity clauses, yielding forms such as '*you was good, weren't you?*' Chambers (2004: 132) posits a relationship between the second vernacular pattern and the dominant pattern of *was*-generalization, claiming that we should expect to find evidence in some vernacular varieties for the second attested pattern developing from the putatively more basilectal pattern of *was*-generalization. To date, there are few studies which provide empirical corroboration for this claim, although Britain (2002a:30) presents evidence of a more levelled paradigm consisting of the forms *was* and *weren't* in the Fenland dialect of eastern England which seems to have emerged from a much more variable system that was based on a possibly once levelled *were* paradigm.

Evidence for the second vernacular pattern can be found in geographically disparate locations including Reading English in the south of England (Cheshire 1982a); York English in the northeast of England (Tagliamonte 1998a); and, as already indicated, the Fens in eastern England (Britain 2002a); as well as in Ocracoke English on the Outer Banks of North Carolina, in the United States (Schilling-Estes

---

<sup>14</sup> Anderwald (2001:11) concedes that the data for Scotland and Ireland are exiguous, although she cites additional evidence based on the *Northern Ireland Transcribed Corpus of Speech*. Anderwald's (2001:11) claim that the dialect of the northwest Midlands is characterized by extensive *was*-levelling in both positive and negative contexts is queried by Rupp (2005: 264), who notes the use of non-standard *weren't* in this area in the vernacular of younger male speakers, although she tentatively suggests that non-standard *weren't* in this area is a relatively recent development.

and Wolfram 1994). Anderwald (2001: 12) claims that this mixed type generalization pattern accounts for 58% of all British dialect regions which could be subjected to statistical analysis.

In one of the earliest discussions of the role of polarity in conditioning patterns of *was/were* variation, Cheshire (1982a: 46) notes that whereas non-standard *were* in positive polarity clauses attained a figure of 4% in the English spoken by Reading teenagers, the figure dramatically increased to approximately 40% in negative polarity clauses. Cheshire (1982a: 46) suggests that the use of non-standard *were* in negative contexts is a 'relic of an earlier dialect form'. However, the extent to which this variant is a historical remnant is difficult to ascertain.<sup>15</sup> According to Nevalainen (2006: 355), *weren't* generalization is not widely found in earlier varieties of English except with plural subjects and singular *you*, as in the following example taken from the fifteenth century (cited in Nevalainen 2006: 355):

(39) I conceyvyd þat at þe wryghtyng off þat letter *ye weere nott* serteyn off the  
delyng betwyn Towneshende and me. (1474 J2PASTON I, 477)<sup>16</sup>

In contemporary dialects, the 'split' between *was* in positive clauses and *weren't* in negative ones contrasts with standard English usage in which *was* and *were* are paradigmatically configured to mark person-number distinctions. Schilling-Estes and Wolfram (1994: 289) argue that in Ocracoke English, the reconfiguration of the allomorphs of past *BE*, *was* and *weren't*, to mark polarity distinctions is best viewed in terms of a remorphologization process which is functionally motivated. Kortmann

---

<sup>15</sup> Not all variants of the *was/were* variable can be extensively treated diachronically owing to the paucity of data contained in previous studies. However, access to historical sources such as the *Linguistic Atlas of Late Middle English* and computerised corpora such as the *Helsinki Corpus* means that the diachrony of particular variants can now be more comprehensively addressed in future research.

<sup>16</sup> Wolfram and Schilling-Estes (2003:132) argue that in contemporary dialects, non-standard variants of *were* in negative contexts are restricted to the contracted form *weren't*; in other words, the non-contracted variant *were not* as in *I were not there* is claimed to be an ungrammatical string, at least in the varieties analysed by these researchers. Wolfram and Schilling-Estes (1994) further claim that the suppletive form *weren't* has developed through analogy with *ain't*.

and Wagner (2005:3) refer to the paradigmatic split between positive *was* and negative *weren't* as a prime example of iconicity where a morphological distinction is recruited to encode a maximal semantic and cognitive difference (i.e. affirmation versus negation).

In those dialects where there appears to be a functional realignment of forms according to polarity, rates of generalization of non-standard forms are frequently reported to be asymmetrical in positive and negative contexts. Anderwald's (2001: 7) wide-ranging survey of *was/were* variation in British dialects highlights the propensity for the generalization of non-standard forms to be more strongly favoured in negative clauses than in positive ones.<sup>17</sup> Anderwald (2001: 7) adds that the favouring effect of negative contexts on generalization strategies receives some support from the apparent cross-linguistic trend for negation to be generally treated as marked from a typological perspective when compared to affirmative contexts (see Greenberg 1966: 50; Croft 1990: 93). One of the consequences of this tendency is for forms in affirmative contexts to be more frequent and irregular, in addition to exhibiting greater allomorphic variation than their negative counterparts (Croft 1990:72). Leaving aside morphological variation in past *BE* for the moment, there is additional evidence in non-standard English that morphological distinctions that exist in positive clauses may be reduced in negative clauses. For example, in certain non-standard varieties, *ain't* and *don't* appear to be grammaticalizing as invariant negative particles which are competing with standard English inflected negative auxiliaries (see Anderwald 2002: 152). Examples of this type of change can also be adduced for other languages: Campbell (1980) discusses the case of the non-standard Finnish spoken by immigrants in Fitchburg in the USA, in which an invariant negative

---

<sup>17</sup> This is not invariably the case. Recent unpublished research by Cheshire and Fox (2006) indicates that levelling to non-standard *was* (66%) in contexts of standard *were* in outer London is marginally greater than levelling to non-standard *weren't* (63%) in contexts of standard *wasn't*.

particle, *ei*, is used for all persons in place of an inflected negative verb in standard Finnish.

Cross-variety comparisons of British dialects that instantiate a functional split between positive *was* and negative *weren't* reveal substantial differences in the frequency of levelling to non-standard *weren't*. In Tagliamonte's (1998a: 165) study of past *BE* in York English, *weren't* levelling was especially high in tag questions (57% versus 15% in other contexts), whereas Britain (2002a: 30) reports that levelling to non-standard *weren't* in the Fens in eastern England is very advanced overall at 88.48%. Britain (2002a: 17) claims that the younger generation in particular are playing a pivotal role in the morphological restructuring of the past tense *BE* paradigm in the Fens, with young people using *was* in affirmative clauses and *weren't* almost exclusively in negative contexts. Britain (2002a: 33-4) argues that scrutiny of intergenerational patterns of *was/were* variation in Fenland varieties indicates that this pattern is diffusing northwards across the Fens, with the youngest generation employing a bipartite *was/were* system (i.e. *I was/I weren't*) aligned in terms of polarity, whereas the oldest speakers investigated by Britain (2002a) used *was* and *were* variably in affirmative contexts and *weren't* in negative contexts. From an apparent-time perspective, the trend which emerges from Britain's (2002a) data is one where a simplified system used by the younger generation is replacing a more variable pattern employed predominantly by the oldest speakers in Britain's (2002a) corpus. Britain (2002a: 26) suggests that these changes may be the result of recent contact between Fenland dialects and varieties from the southeast.

The extent to which similar restructuring processes are in progress elsewhere in Britain warrants further investigation (see also below). Cheshire *et al.*'s (1993: 80) survey of British dialect grammar, which was based on questionnaires distributed to

schools in Britain, found, for example, that non-standard *were* was reported only in negative contexts by all four schools surveyed in the central Birmingham area, whereas other schools in the West Midlands and East Midlands reported the use of non-standard *were* in other constructions. Cheshire *et al.* (1993: 80-1) suggest that whereas variation in the past tense forms of *BE* persist in the Midlands more generally, in the core metropolitan region of Birmingham, allomorphic variation in past *BE* may have been reduced resulting in a more streamlined, refunctionalised split between *was* in positive contexts and *were* in negative contexts.

### 3.3.6 *Vernacular Pattern 3*

A third vernacular pattern that is not widely discussed in the research literature, and is apparently much more liminal than the other two patterns, concerns levelling to *were* in positive and negative polarity clauses. Britain (2002a: 19) remarks with regard to this pattern that there is little information about 'its present socio-geographical distribution or the linguistic constraints operative on such varieties.' Wolfram and Sellers (1999: 94) claim that 'levelling to *were* [...] appears to be a minority levelling option attested in selected regions of England and some coastal regions of the United States.' Shorrocks (1999: 168-9) notes that *were*-levelling can be found in Bolton in northwest England, and Moore's (2004: 386) research on variation in the speech of adolescent girls in a Bolton high school similarly showed an overwhelming tendency towards levelling to *were*, which had been recruited by some speakers for stylistic purposes. Anderwald's (2001: 10) analysis of *was/were* variation in contemporary British dialects revealed that *were*-generalization is the preferred option to *was*-generalization in four areas: the south Midlands, the central north, the north-east, and London. However, as we shall see below, *were*-generalization in

positive and negative contexts is not a characteristic feature of the vernacular of contemporary London youth.

### 3.3.7 *Was/were variation in the southeast: Evidence from inner and outer London and neighbouring East Anglian dialects*

Data displayed in Anderwald (2001: 5, Table 1) suggest that in London generalized *were* (15.2%) in positive clauses is more frequent than generalized *was* (10.1%). However, when the data are reconfigured by age group, a more nuanced picture emerges of *was/were* variation in the capital. Table 3.2 summarises patterns of *was/were* generalization in six different age cohorts in the London dialect region. On the basis of the data displayed below, Anderwald (2001: 14) suggests that in London there may be a shift from a predominantly *were*-generalised pattern to a mixed system, with speakers under 35 years of age advancing this change.

TABLE 3.2

#### *WAS/WERE* GENERALIZATION FOR SIX AGE GROUPS IN THE LONDON DIALECT AREA (AFTER ANDERWALD 2001: 14)

Age group	<i>was</i>	<i>were</i>	<i>wasn't</i>	<i>weren't</i>	Type
0-14	8.6%	1.2%	4.3%	32.1%	mixed
15-24	8.3%	6.5%	12.5%	26.3%	mixed
25-34	31.0%	3.3%	11.7%	28.9%	mixed
35-44	10.7%	44.1%	4.5%	67.1%	WERE
45-59	5.3%	21.3%	0%	24.1%	WERE
60+	6.3%	5.7%	5.0%	11.5%	mixed ? <sup>18</sup>

<sup>18</sup> Anderwald (2001: 14) notes that over 60-year-olds are sparsely represented in the data that were examined, thereby making it difficult to draw any preliminary conclusions about the nature of *was/were* variation in this age group.

According to recent preliminary research conducted by Cheshire and Fox (2006), focusing on the English of adolescents in London (see Kerswill and Cheshire 2004-2007), patterns of variation in past tense *BE* in the speech of the younger generation are more complex than the data displayed in Table 3.2 suggest. An important finding emerging from Cheshire and Fox's (2006) study of *was/were* variation in London, which has a bearing on the preadolescent data examined in section 3.5 below, is the linguistically heterogeneous nature of the London dialect area with regard to variation in past *BE*.

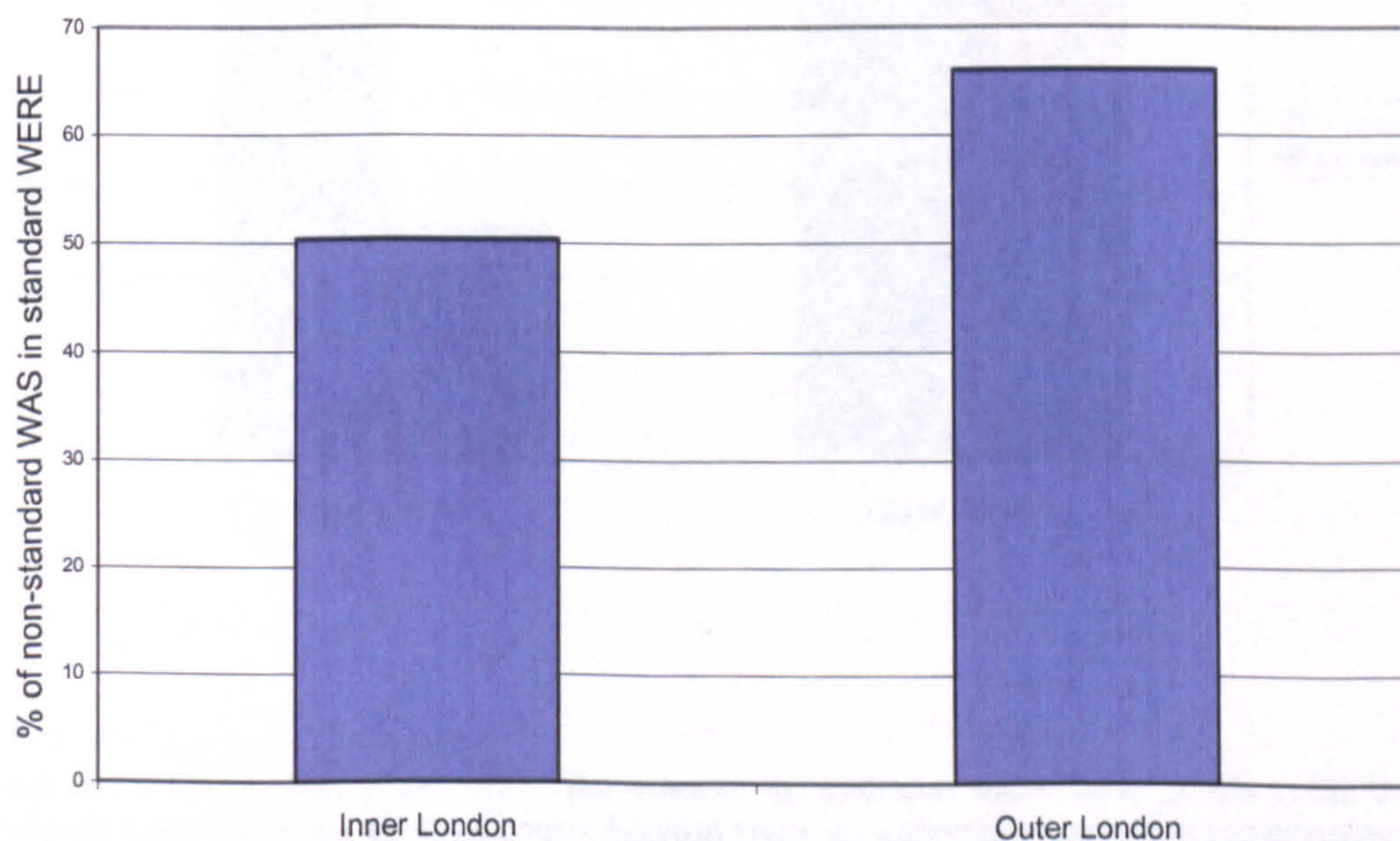
Figure 3.1 below compares rates of levelling to non-standard *was* in inner and outer London.

FIGURE 3.1

COMPARISON OF RATES OF LEVELLING TO NON-STANDARD

*WAS* IN CONTEXTS OF STANDARD *WERE* IN INNER AND OUTER LONDON

(FIGURES ABSTRACTED FROM CHESHIRE AND FOX 2006<sup>19</sup>)

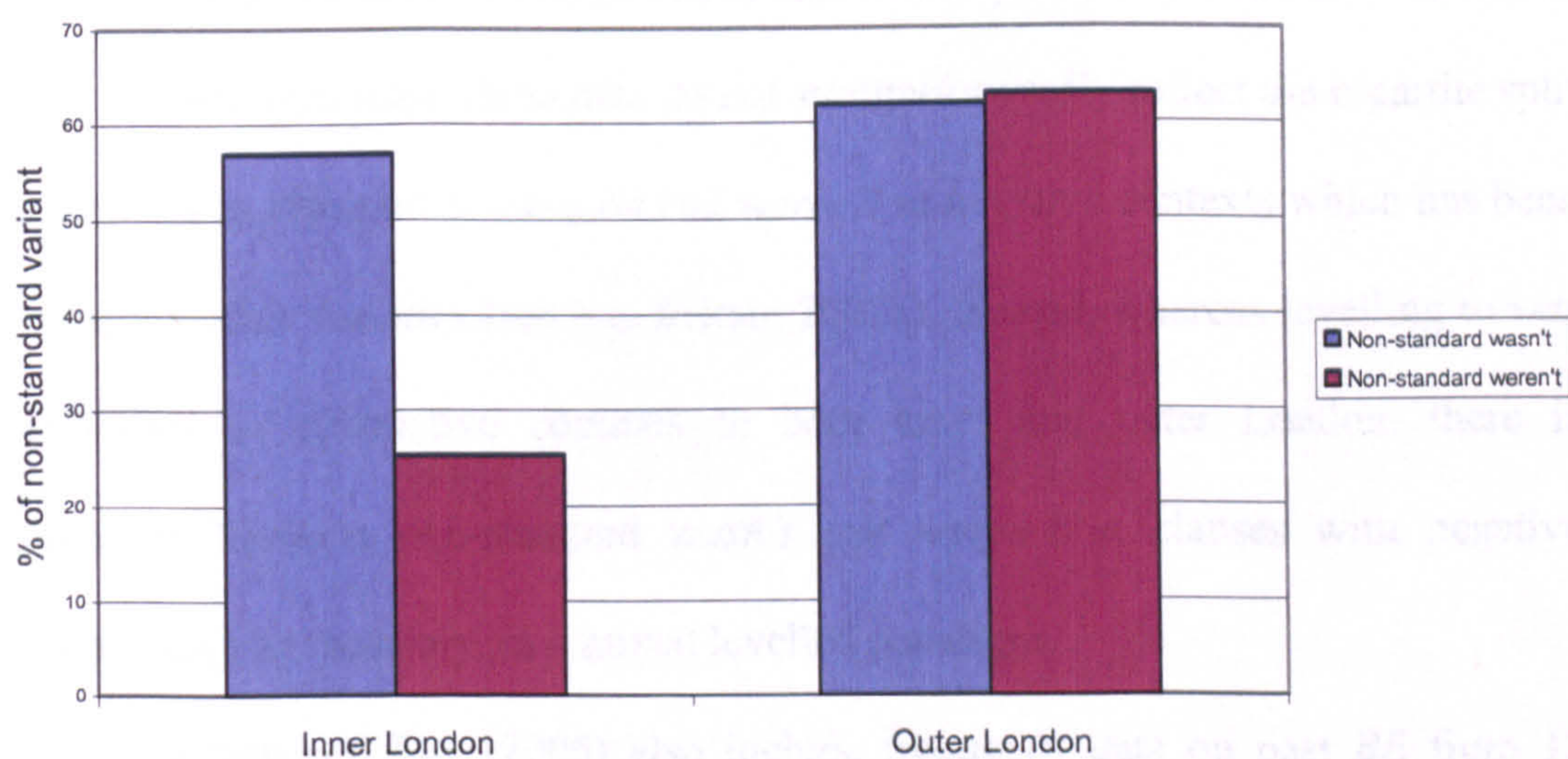


<sup>19</sup> Figures do not take into account levelling to non-standard *was* in existential contexts (see 3.3.3 above).

Although Cheshire and Fox (2006) report that levelling to *was* in positive clauses was the predominant pattern in their data, partially corroborating the findings for younger Londoners discussed in Anderwald (2001:140), inner London speakers were found to exhibit lower rates of levelling to *was* than speakers in outer London. Moreover, Cheshire and Fox (2006) point to correlations between patterns of *was/were* variation and the ethnic background of speakers, with, for example, speakers of Bangladeshi heritage using non-standard *was* at markedly lower rates than speakers from white British backgrounds.<sup>20</sup>

FIGURE 3.2

COMPARISON OF NON-STANDARD *WASN'T* AND *WEREN'T* IN INNER AND OUTER LONDON (FIGURES ABSTRACTED FROM CHESHIRE AND FOX 2006)



<sup>20</sup> Research focusing on variation in other areas of the grammar additionally points to the importance of ethnicity, as well as the inner and outer London areas, as important parameters constraining variation in London English. Kerswill and Cheshire (2004-2007) examine adolescent vernaculars in the inner London borough of Hackney and the outer London borough of Havering. Both boroughs differ in terms of ethnic populations; the proportion of recent migrants; and socioeconomic class. Such contrasts offer fertile territory for tracking linguistic innovations in inner versus outer London. Cheshire *et al.* (2005b) report on innovations in the diphthong system that are not found outside inner London. Torgersen *et al.* (2006) discuss ethnicity in London as a source of innovation in the vocalic variables used by inner London teenagers.



Although overall numbers are small, Cheshire and Fox's (2006) data also reveal differences between inner and outer London in terms of the levelling strategies (and their respective rates of occurrence) that can be distinguished in negative clauses. Figure 3.2 above compares non-standard *wasn't* and *weren't* in inner and outer London.

Whereas rates of levelling to non-standard *wasn't* exceed those of levelling to non-standard *weren't* in inner London, the reverse state of affairs obtains in outer London where rates of levelling to non-standard *weren't* are marginally greater than those for non-standard *wasn't*. If change is underway in patterns of variation in past *BE* in London youth, as hypothesised by Anderwald (2001: 14), then the data above suggest not only that this may be taking place at different rates in inner and outer London, but that this may also involve greater competition between allomorphic variants in negative clauses than has been reported in previous studies of variation in past *BE*. At the very least, these data do not straightforwardly reflect the bipartite split between *was* in affirmative contexts and *weren't* in negative contexts which has been reported for other varieties (see e.g. Britain 2002a). Instead, whereas levelling to *was* is prevalent in affirmative contexts in both inner and outer London, there is competition between non-standard *wasn't* and *weren't* in clauses with negative polarity, resulting, possibly, in a mixed levelled paradigm.

Cheshire and Fox (2006) also include frequency data on past *BE* from 12 elderly speakers aged 70 and above (6 elderly speakers recorded in inner London, and 6 recorded in outer London), which permits some provisional comparisons to be made from an intergenerational perspective. In terms of the social distribution of non-standard *weren't*, the study by Cheshire and Fox (2006) suggests that this variant is more frequent in the speech of adolescents in comparison with the usage of older

speakers. Figure 3.3 below compares rates of levelling to non-standard *weren't* in contexts of standard *wasn't* in the vernacular of the adolescents recorder in inner and outer London and compares these with rates of levelling to non-standard *weren't* in the speech of the elderly speakers.

FIGURE 3.3

COMPARISON OF RATES OF LEVELLING TO NON-STANDARD *WEREN'T* IN THE VERNACULAR OF ADOLESCENTS AND ELDERLY SPEAKERS RECORDED IN INNER AND OUTER LONDON (FIGURES ABSTRACTED FROM CHESHIRE AND FOX 2006)

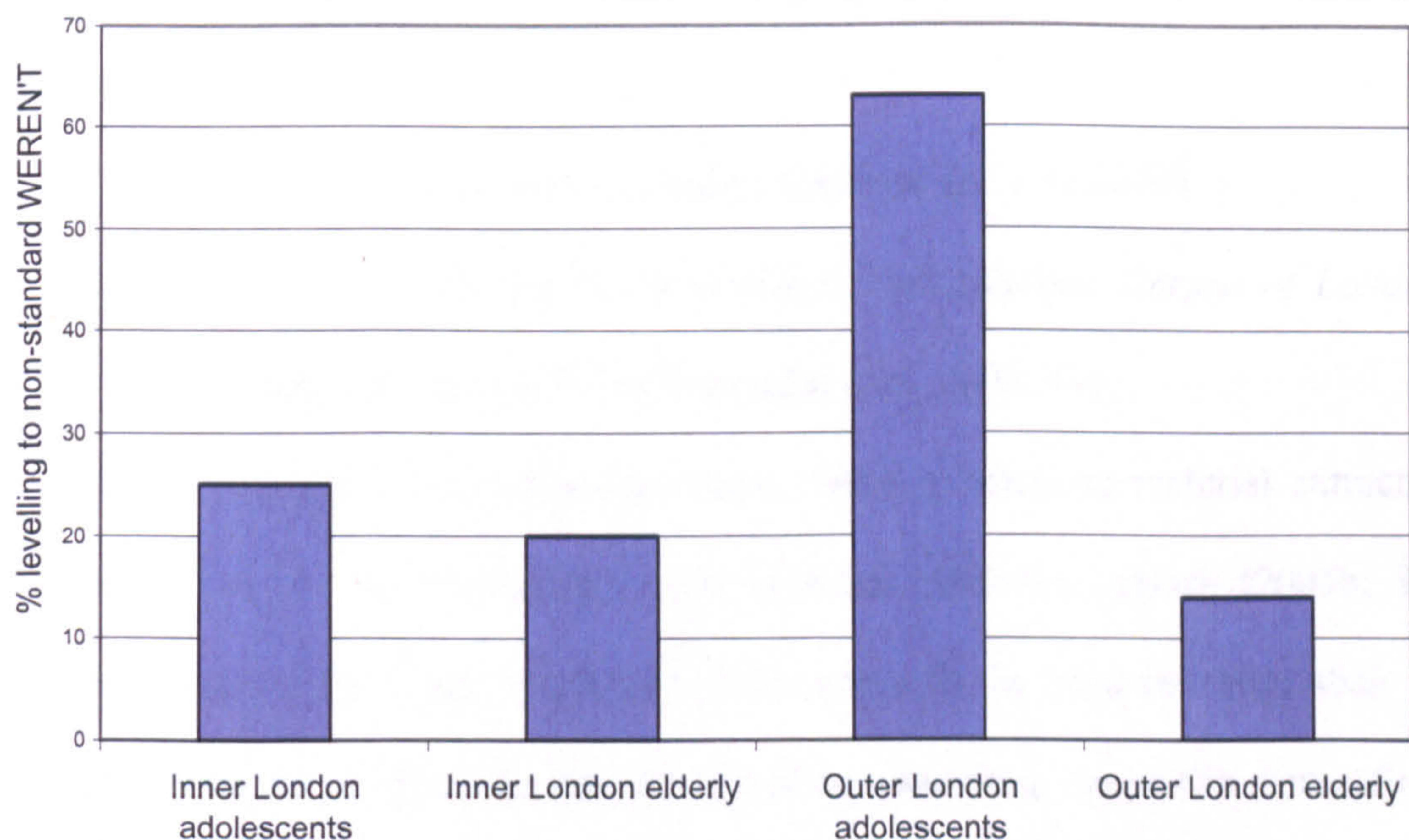


Figure 3.3 shows that adolescents make greater use of non-standard *weren't* than elderly speakers. This trend is evident in both inner and outer London, but is particularly salient in the case of the outer London speakers. A similar trend towards an increase in the use of non-standard *weren't* is documented by Wolfram and Schilling-Estes (2003) in certain North American vernaculars spoken in the Mid-Atlantic region of the United States. However, whereas Wolfram and Schilling-Estes

(2003) claim that this trend is restricted to peripheral varieties spoken in isolated communities in the United States, accelerated rates of levelling to *weren't* evident in Cheshire and Fox's (2006) cross-generational data indicate that a similar change may also be a feature of mainstream urban varieties in Britain.

Another important point raised by Cheshire and Fox (2006) in their observations on non-standard allomorphs of *BE* in negative polarity clauses is that the high frequency of non-standard *weren't* with *it* subjects, particularly in tag constructions in outer London, may be indicative of the grammaticalization of *weren't it* to a discourse marker, comparable to the use of the clause-final marker *innit*, which is also attested in the speech of London teenagers. Examples of these features are given below:

(40) that's not good *weren't it* (cited in Cheshire and Fox 2006)

(41) no but you only like Cody or Hagar *innit* (*Bergen Corpus of London Teenage Language*, cited in Stenström *et al.* 2002: 49)

Using evidence from the mid-twentieth century, based on material extracted from the *Survey of English Dialects*<sup>21</sup> (Orton *et al.* 1963-71), Pietsch (2005b: 80) notes that generalized *were* forms are more common in negative tags than in affirmative clauses, suggesting that the use of non-standard *weren't* in clause-final tags is by no means a recent phenomenon. Using data from the *British National Corpus*, Anderwald (2002:178) reports that nine British dialect areas favour non-standard *weren't* in tag questions more frequently than in non-tag environments. In data gathered in York, Tagliamonte (1998a: 164) observes that 80% of negative tags occurred with non-standard *weren't*, which demonstrates the existence of a similar pattern in contemporary northern dialects. Consistent with the findings reported in

---

<sup>21</sup> The *Survey of English Dialects* gathered data from non-mobile older speakers in rural areas. Many of these speakers were born over a century ago.

Cheshire and Fox (2006), Tagliamonte (1998a: 163) reports that third person *it* in York English exhibits a high frequency of non-standard *weren't*. In fact, according to Tagliamonte (1998a: 163), the effect of negative polarity on the use of non-standard *weren't* is particularly salient with this subject type when it is contrasted with other grammatical subjects.

In contrast with the York data, however, Cheshire and Fox's (2006) study reveals that there are important differences between northern and southern dialects in terms of the differentiated pattern of non-standard *was* usage according to grammatical person. Figure 3.4 below shows the distribution of non-standard *was* by grammatical person in the speech of the adolescents and elderly speakers recorded in inner and outer London (existentials are not included).

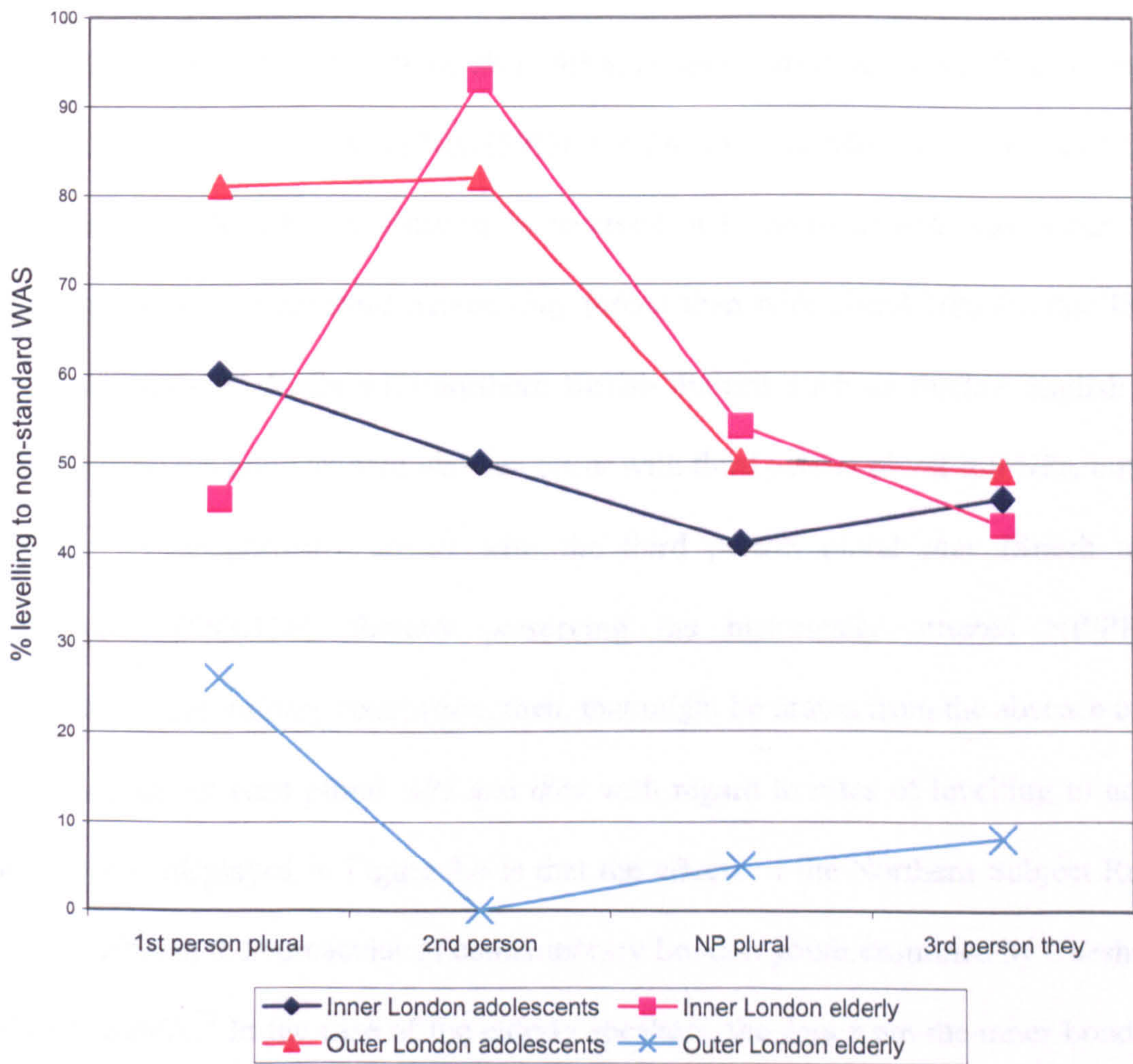
Of particular interest are the relatively high rates of levelling to non-standard *was* with the second person in the speech of the adolescents, and the inner London elderly speakers. The outer London elderly speakers, in marked contrast with the outer London adolescents, show no use of non-standard *was* with the second person.

From an historical perspective, it is interesting that the use of non-standard *was* with the second person is discontinuous with diachronic patterns of variation attested in southern dialects in the Middle English period, during which, according to extant historical sources, there was a divide between northern and southern dialects, with *were* predominating with the second person in the south and *was* in the north (Forsström 1948: 22). The contemporary distribution of *was* with the second person in adolescent speech in inner and outer London would seem to offer some support for Nevalainen's (2006: 367) observation that this grammatical person has been the locus of change as far as the use of non-standard *was* is concerned, although further

synchronic and diachronic research is required to substantiate this point (see further 3.5 below).

FIGURE 3.4

COMPARATIVE DISTRIBUTION OF NON-STANDARD *WAS* BY GRAMMATICAL PERSON IN THE VERNACULAR OF INNER AND OUTER LONDON ADOLESCENTS AND ELDERLY SPEAKERS (FIGURES ABSTRACTED FROM CHESHIRE AND FOX 2006)



Recall too that the existing textual evidence from Middle English dialects (see Table 3.1 above) indicates that non-standard *was* was not a characteristic feature of southern dialects with the first person singular. However, the contemporary London

data examined by Cheshire and Fox (2006) show that non-standard *was* co-occurs comparatively frequently with this grammatical person in the speech of the adolescents, and, to a lesser extent, in the usage of the elderly speakers.

The data displayed in Figure 3.4 also show that the use of non-standard *was* with other grammatical persons differs in important respects from the constraint hierarchies posited for northern dialects based on differential frequencies of non-standard *was* across the verbal paradigm (see Tagliamonte 1998a; Smith and Tagliamonte 1998). Note that for the outer London adolescents, the frequency of non-standard *was* with third person *they* (49%) is only marginally lower than the rate of non-standard *was* with plural NPs (50%). For the inner London adolescents, on the other hand, the NP/PRO distinction is reversed, with non-standard *was* occurring slightly more often with third person *they* (46%) than with plural NPs (41%). This pattern contrasts markedly with northern British dialects such as Buckie English in Scotland in which non-standard *was* can occur with third person plural full NPs, but is found to be categorically absent with the third person plural *they* (Smith and Tagliamonte 1998:116), thereby preserving the historically attested NP/PRO distinction. A preliminary conclusion, then, that might be drawn from the absence of a sharp contrast between plural NPs and *they* with regard to rates of levelling to non-standard *was* displayed in Figure 3.4 is that the effects of the Northern Subject Rule are not visible in the vernacular of contemporary London youth examined by Cheshire and Fox (2006).<sup>22</sup> In the case of the elderly speakers, the data from the inner London speakers maintain the NP/PRO distinction, whereas those from the outer London elderly speakers show a reversal of this constraint, although this finding should be interpreted with some caution as the overall number of tokens (N=20) of non-standard *was* in contexts of standard *were* for the outer London elderly speakers is very small.

---

<sup>22</sup> These are my interpretations of the data in Cheshire and Fox (2006).

Does the absence of a clearly operative NP/PRO constraint in the vernacular of the London adolescents investigated by Cheshire and Fox (2006) reflect an age-graded effect, or is it representative of a more widespread regional pattern? Britain *et al.* (2005) offer some purchase on this question by claiming that in certain dialects in eastern and southeastern England an alternative pattern to the Northern Subject Rule exists. In the Fens, Britain (2002a: 26) reports that in plural contexts, noun phrases disfavour non-standard *was* more than the third person plural pronoun. The replication of this finding in various geographical locations throughout rural and urban East Anglia leads Britain *et al.* (2005) to posit the existence of an East Anglian or Southern Subject Rule which operates variably, and favours a reversal of the NP/PRO constraint associated with the Northern Subject Rule.

In section 3.5 below, I adduce additional data from London preadolescent and adolescent speakers which exhibit parallelisms with the findings reported in Britain (2002a), Britain *et al.* (2005), as well as Cheshire and Fox (2006).

In the following section, I detail the procedures which I adopted for circumscribing the envelope of variation in the preadolescent data and in additional datasets.

### 3.4 METHODOLOGY

Initially, every instance of *was* and *were* in both positive and negative contexts was extracted from the preadolescent corpus. Following previous studies of *was/were* variation in contemporary British dialects, syllable structure and vowel quality were taken into account in identifying individual variants (see Tagliamonte 1998a: 160).

As both Tagliamonte (1998a: 160) and Britain (2002a: 20) point out, there is some variation in the phonetic realisations of *was* and *were* in contemporary dialects

which may be remnants of older variant pronunciations attested, for example, in the *Survey of English Dialects* (Orton *et al.* 1963-71). For the most part, phonetic variation in individual realisations of *was/were* in the preteen data did not pose any serious obstacle to the identification of variant forms as instances of either *was*, *wasn't*, *were* or *weren't*. However, there were a small number of forms used by some of the older (10-11) boys which were difficult to assign to either *wasn't* or *weren't*:

(42) I think it was the Rock, *wannit* [wɒnɪ?] [7M10/11]

(43) they were going like that, *wont* [wɒn?] they [12M10/11]

(44) he beat them yeah that *was (sic)*, *wannit* [wɒnɪ?] [2M10/11]

The variant *wont* is attested in the *Survey of English Dialects* (Orton *et al.* 1963-71) for parts of East Anglia, suggesting that it is not restricted to outer London and that it was possibly more common and widespread in the past. According to Britain (2002a: 36), forms such as *wont* may well be reanalysed by learners as *weren't*. Note too that in examples (42-44) above, the italicised variants all occur in tag questions, which previous studies have identified as being a particularly favorable environment to *weren't* levelling (see e.g. Tagliamonte 1998a: 165). However, this environment alone would not be a sufficient diagnostic for classifying forms such as *wont/wannit* as variants of *weren't/weren't it*.

Other varieties of English have also been found to contain the past negative form *wont*. Wolfram and Sellers (1999) note the existence of negative *wont* in Lumbee Vernacular English spoken in North Carolina, although they add that it is a marginal variant in their data. Similarly, Hazen (1998: 221) maintains that in Warren County, North Carolina, there is evidence of a tripartite division among negative forms for past *BE* (i.e. *wasn't*, *weren't* and *wont*).<sup>23</sup> According to Hazen (1998: 225,

---

<sup>23</sup> Cheshire *et al.* (1999) also cite the form *wont* as a distinctive feature of the negative past *BE* paradigm in Hull.



234) not only does *wont* correlate with factors such as speaker sex, ethnicity, and cultural identity in Warren County, it can function as a sociolinguistic marker for some speakers. Furthermore, Hazen (1998) claims that *wont* is morphophonologically differentiated from other allomorphs in the negative past *BE* paradigm, and is not derived synchronically from either *wasn't* or *weren't*.

The very small number of *wont* forms in the preadolescent data precludes a more thorough analysis of the sociolinguistic status of this variant in the children's vernacular. Given that there is no principled way of categorically assigning this form to either *wasn't* or *weren't*, and that there is some recognition of *wont* in previous research as a distinct variant (Hazen 1998), I decided to exclude this variant from the quantitative analysis of *was/were* variation. In fact, apart from a handful of examples of *wont* in the data, the otherwise limited existence of variant phonetic past *BE* forms in the preadolescent data supports Britain's (2002a: 35) observation that phonetic levelling of realisations of past *BE* is a noteworthy feature of contemporary southeastern dialects.

Following Smith (2000: 52), subjunctive usage was not examined, although conditional sentences containing a second person singular, first person plural, third person plural subject or plural pronoun were included as the form *were* would feature in standard contexts with these subject types. Thus, sentences of the type *you'd think twice if you were going to be arrested* were included, whereas sentences such as *if I were you, I would go out* were not.

Reformulations, as well as incomplete and anomalous utterances, were also excluded from the final analysis of the data. Each remaining token of *was* and *were* was coded for a variety of factors that previous research has shown to effect *was/were* variation. The main factor groups that were incorporated into the coding protocol

were grammatical person and polarity, as well as the age and gender of the speaker. In addition, characteristics of the subject noun (e.g. collective subjects and conjoined subjects), as well as syntactic configuration (i.e. subject-verb adjacency/non-adjacency) were coded for.

Although several previous studies (e.g. Tagliamonte 1998a; Tagliamonte and Smith 2000) have included existential contexts in quantitative analyses of patterns of *was/were* variation, I decided to exclude existentials from the analysis of the data principally on the grounds that existential and non-existential contexts are syntactically remote environments (see also 3.3.3 above). This decision was further endorsed by an initial distributional analysis of *was/were* in the preadolescent data, which revealed that in expletive *there* constructions non-standard *was* was categorical with postverbal plural subjects; in other words, there were no variable concord patterns to analyse in this environment.

#### 3.4.1 *Data from the Bergen Corpus of London Teenage Language (COLT)*

In order to extend the analysis to other age groups to investigate Anderwald's (2001: 14) hypothesis that *was/were* variation may be a locus of change among the younger generation of Londoners, I also examine patterns of *was/were* variation in the vernacular of London teenagers who were recorded in 1993 as part of an investigation into adolescent language. The *Bergen Corpus of London Teenage Language (COLT)*, which now forms part of the *British National Corpus (BNC)*, has already formed the basis of sociolinguistic analyses of discourse-pragmatic features (see e.g. Andersen 2001; Stenström *et al.* 2002), but, to the best of my knowledge, patterns of *was/were* variation have not yet been analysed in this corpus.

The adolescent data that I draw on come from the CD-ROM version of the COLT corpus. The corpus consists of just under half a million words and is based on recordings made in five different London school boroughs: Hackney, Tower Hamlets and Camden (Inner London); Barnet (Outer London); and Hertfordshire, located within the London Metropolitan area.

Given that the COLT corpus was collected using a very different research design from my own, and was based on speakers from a variety of social and ethnic backgrounds, the interpretation of the data emerging from the quantitative analysis of *was/were* variation in the COLT corpus must necessarily be cautious.

Some remarks are in order about how the COLT corpus was constructed, and about the social profiles of contributing speakers. Volunteers aged between 13 and 17 were issued with small portable tape recorders, and were asked to record the conversations that they took part in, preferably with speakers of their own age, over a period of 3 to 5 days. Volunteers were additionally requested to keep a log of whom they spoke to, the age and gender of their interlocutors, as well as to supply details of where the conversations took place (Stenström *et al.* 2002: 5).

There were 9 female and 21 male volunteers who were responsible for making the recordings. The nine female volunteers recorded approximately 44% of the entire corpus compared to the males' 56% (Stenström *et al.* 2002: 7). In addition to gender, volunteers were classified as being either 'white' or 'ethnic minority', a classification which is somewhat crude, as Stenström *et al.* (2002: 23) acknowledge. Furthermore, information on the ethnic background of recorded speakers is not always recoupable, as not all participants in the recorded conversations furnished details regarding their ethnicity. Similar difficulties exist with retrieving information about speakers' social class membership. Social classification of the volunteers who made

the recordings was undertaken by classifying the speakers into three different social groups based on their borough of residence, their parents' occupations, and whether or not their parents were employed (Stenström *et al.* 2002: 21). Only volunteers who were entrusted with recording equipment, and their families, are classified socially, which means that a certain amount of recorded material in the COLT corpus cannot be indexed in terms of speakers' social class affiliations.

There is also some variation in the final corpus in the age of the speakers who were recorded. The greater part of the corpus, constituting 85% of the recorded material, is made up of speakers in early adolescence (10-13); middle adolescence (14-16); and late adolescence (17-19). In order to restrict my analysis to teenage speakers in the COLT corpus, I excluded any data that came from young children or adults.

Given the limitations I have highlighted above with regard to the recoverability of aspects of individual speakers' social profiles in COLT, the only external variable that I factor into the analysis of adolescent patterns of *was/were* variation is gender.

In order to maximize comparability between the preadolescent and adolescent data sets, existential contexts were also excluded from the quantitative analysis of *was/were* variation in the COLT corpus.<sup>24</sup> After the further exclusion of reformulations as well as incomplete and anomalous utterances, each token of *was/were* which was extracted and retained for analysis was coded for grammatical person and polarity.

---

<sup>24</sup> In contrast to the preadolescent data, there are variable concord patterns in the adolescent data in expletive *there* constructions with postverbal plural subjects. Rates of levelling to non-standard *was* in this environment are nevertheless high, reaching 53%. Rates of singular concord with postverbal plural subjects are more advanced in present tense existential constructions (77.7%) than in the past tense.

### 3.5 DISTRIBUTIONAL ANALYSIS

Table 3.3 below shows the overall distribution of *was/were* in the preadolescent data.

TABLE 3.3

#### DISTRIBUTION OF *WAS/WERE* IN THE PREADOLESCENT DATA

	Standard <i>was</i>	Non-standard <i>was</i>	Standard <i>were</i>	Non-standard <i>were</i>	TOTAL
N	1034	119	101	18	1272
%	81	9	8	1.4	

Table 3.3 shows that in terms of the overall distribution of *was/were* in the data, non-standard *was* is more frequent than non-standard *were*. These preliminary findings are consistent with Anderwald's (2001: 14) observation that the vernacular of the youngest generation of Londoners is a mixed system, rather than one which is based on *were*-generalization, which, according to Anderwald (2001: 14), is characteristic of the vernacular of older Londoners (35-59).

Table 3.4 below shows the percentage of non-standard *was* in contexts of standard *were*.

TABLE 3.4

#### DISTRIBUTION OF NON-STANDARD *WAS* IN STANDARD *WERE* CONTEXTS

	<i>was</i>	<i>were</i>	TOTAL
N	119	101	220
%	54	46	

The overall rate of levelling to non-standard *was* in contexts of standard *were* is quite high. Conversely, Table 3.5 below shows that non-standard *were* is comparatively rare in contexts of standard *was*.

TABLE 3.5

DISTRIBUTION OF NON-STANDARD *WERE* IN STANDARD *WAS* CONTEXTS

	<i>were</i>	<i>was</i>	TOTAL
N	18	1034	1052
%	1.7	98.3	

Although the data presented in Tables 3.4 and 3.5 confirm the pervasiveness of levelling to non-standard *was* in the data, which contrasts markedly with the rarity of non-standard *were*, these data camouflage the strength of a polarity effect that constrains patterns of *was/were* variation. A more complex state of affairs obtains when levelling strategies in positive clauses are compared with those in negative clauses, as shown in Tables 3.6 and 3.7 below.

TABLE 3.6

DISTRIBUTION OF NON-STANDARD *WAS* IN CONTEXTS OF STANDARD *WERE* BY POLARITY

	<i>was</i>	<i>were</i>	TOTAL
POSITIVE			
N	118	97	215
%	55	45	
NEGATIVE			
N	1	4	5
%	20	80	

TABLE 3.7

DISTRIBUTION OF NON-STANDARD *WERE* IN CONTEXTS OF STANDARD *WAS* BY POLARITY

	<i>were</i>	<i>was</i>	TOTAL
POSITIVE			
N	4	1010	1014
%	0.4	99.6	
NEGATIVE			
N	14	24	38
%	37	63	

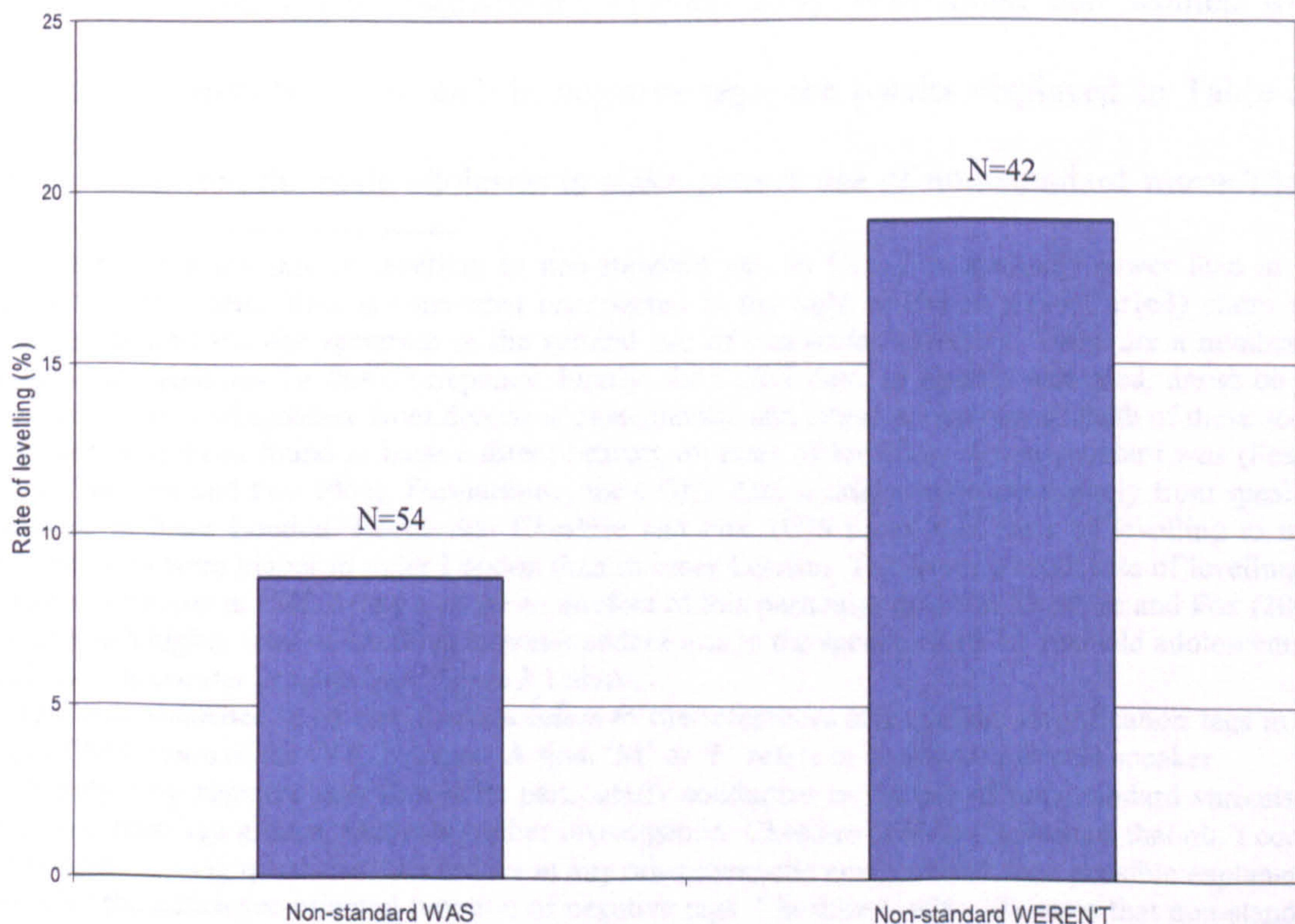
Although the number of tokens for negative clauses is restricted, particularly in the case of the distribution of non-standard *was* in contexts of standard *were*, the data indicate that levelling strategies are sensitive to polarity distinctions, with levelling to non-standard *was* being preferred in positive clauses, whereas negative clauses favour levelling to non-standard *weren't*. Although this finding suggests that the preadolescents share a more generalized dialect pattern with other speakers in the southeast and in East Anglia (see e.g. Britain 2002a), and, possibly, other urban centres in England (see Trudgill 1990: 98), in favouring non-standard *were* in clauses with negative polarity, rates of levelling to *weren't* are less high than those reported elsewhere. Britain (2002a: 30) reports that rates of levelling to non-standard *weren't* in the Fens are very advanced, with over 88% of standard *wasn't* contexts realised as non-standard *weren't*. Note too that rates of levelling to non-standard *weren't* do not exceed those of levelling to non-standard *was* in affirmative contexts, which is at variance with the claim made by Anderwald (2001: 7) that any generalization strategy

is more strongly favoured in negative environments than positive ones. Nevertheless, in spite of the frequency differences in non-standard *weren't* across varieties, the data displayed above in Tables 3.6 and 3.7 are congruent with patterns of *was/were* variation that are characteristic of dialects in which non-standard *was* and *were* are functionally configured to mark polarity distinctions (i.e. Vernacular Pattern 2; see 3.3.4).

Additional evidence of this 'paradigmatic split' can be found in the COLT corpus. Figure 3.5 below compares rates of levelling to non-standard *was* in contexts of standard *were* with those of levelling to non-standard *weren't* in contexts of standard *wasn't* in the adolescent data.

FIGURE 3.5

COMPARISON OF RATES OF LEVELLING TO NON-STANDARD *WAS* IN CONTEXTS OF STANDARD *WERE* WITH RATES OF LEVELLING TO NON-STANDARD *WEREN'T* IN CONTEXTS OF STANDARD *WASN'T* IN COLT





In the adolescent data, polarity exerts a very strong conditioning effect on patterns of *was/were* variation. In affirmative contexts, there are no instances of non-standard *were*, and in negative contexts, there are no examples of non-standard *wasn't*. In the adolescent corpus, then, the configuration of non-standard *was/were* to mark polarity distinctions is even more clear-cut than in the preadolescent data.<sup>25</sup>

Further differences exist between the preadolescent data and the COLT corpus with respect to the use of non-standard *weren't*. In the adolescent data, 24% (N=10) of non-standard *weren't* forms are found in tag questions. Moreover, 80% (N=8) of non-standard *weren't* forms embedded in tag questions co-occur with the pronoun *it*, as shown in the following examples:

(45) well it was this thing *weren't it* it was parked up [B136103, 12-1/M]<sup>26</sup>

(46) you're the only one saying *weren't it* [B134902, 6-12/F]

As indicated above, elevated frequencies of non-standard *weren't* in tag questions are reported in several other studies (e.g. Tagliamonte 1998a; Pietsch 2005b).<sup>27</sup> In contrast to Tagliamonte (1998a: 179), who found that women were extending non-standard *weren't* in negative tags, the results displayed in Table 3.8 below show that the male adolescents make greater use of non-standard *weren't* tags

---

<sup>25</sup> Note too that the rate of levelling to non-standard *was* in COLT is markedly lower than in the preadolescent corpus. This is somewhat unexpected in the light of Eckert's (1997a:163) claim that adolescents lead the age spectrum in the general use of vernacular variables. There are a number of possible explanations for this discrepancy. Firstly, the COLT data, as already indicated, draws on the vernacular usage of speakers from diverse socioeconomic and ethnic backgrounds. Both of these social variables have been found to have a direct bearing on rates of levelling to non-standard *was* (Feagin 1979; Cheshire and Fox 2006). Furthermore, the COLT data is drawn more extensively from speakers recorded in inner London. Recall that Cheshire and Fox 2006 found that rates of levelling to non-standard *was* were higher in outer London than in inner London. The lower overall rate of levelling to non-standard *was* in COLT may also be an artefact of this particular data set. Cheshire and Fox (2006) found much higher rates of levelling to non-standard *was* in the speech of 16-19 year-old adolescents in both inner and outer London (see Figure 3.1 above).

<sup>26</sup> The code sequence in square brackets refers to file references and speaker identification tags in the CD-ROM version of the COLT corpus. A final 'M' or 'F' refers to a male or a female speaker.

<sup>27</sup> Exactly why negative tags should be particularly conducive to the use of non-standard variants, at least in certain age groups, warrants further investigation. Cheshire (1982b: 165) notes that *ain't* occurs more often in a tag questions than it does in any other syntactic environment. One possible explanation lies with the addressee-oriented function of negative tags. Cheshire (1996: 10) notes that non-standard variants have characteristically evolved within contexts of involved face-to-face interaction and that 'it is not surprising [...] if these environments tend to regulate variation.'

than female adolescents, even though the females use negative tags slightly more frequently than the males overall.

By contrast, tags with past *BE* are less frequent in the preadolescent data, although there is one example of invariant *weren't it* used by a ten-year-old boy:

(47) we were watching that on the sleepover *weren't it* [5M10/11]

TABLE 3.8

COMPARISON OF MALE AND FEMALE USE OF NEGATIVE TAGS IN COLT

Speaker sex	Non-standard <i>weren't</i> tags	Standard negative tags (past <i>BE</i> )	TOTAL
FEMALE			
N	4	15	19
%	21	79	
MALE			
N	6	8	14
%	43	57	

It is not clear whether the limited number of negative invariant tags of past *BE* in the preadolescent data can be attributed to age-related effects or, possibly, to constraints associated with the sociolinguistic interview which was used to elicit speech samples from the preadolescents. According to Stenström *et al.* (2002: 185), the use of invariant tags is considerably higher among adolescents than other age groups. Andersen (2001) also documents the prominent role played by the teenagers in the use and dissemination of tags such as *innit*, which may be used to fulfill sociosymbolic functions such as indexing membership in vernacular culture.

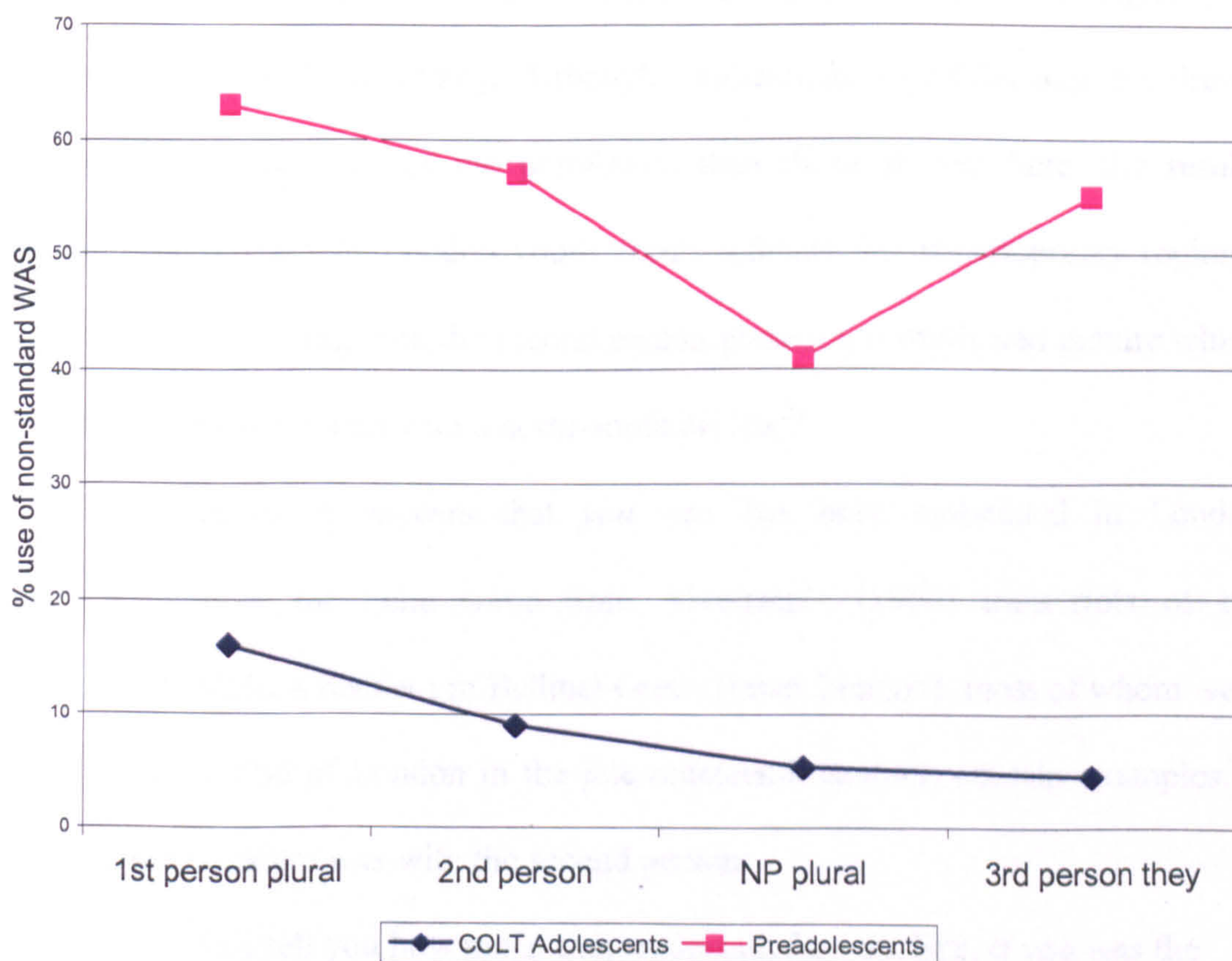
Where the preadolescents do use the negative tag *weren't it* in the interview context, it is apparently employed as a marker of positive politeness to seek

confirmation from a friend about a recently expressed utterance, often in association with the recount of a jointly experienced event, as in (47) above.<sup>28</sup>

Other aspects of the data reveal important similarities in preadolescents' and adolescents' patterns of variation in past *BE*. Figure 3.6 below compares rates of levelling to non-standard *was* by grammatical person in the preadolescent corpus and in COLT.

FIGURE 3.6

COMPARISON OF RATES OF LEVELLING TO NON-STANDARD *WAS* BY GRAMMATICAL PERSON IN THE PREADOLESCENT AND COLT ADOLESCENT CORPORA



<sup>28</sup> See also examples 42-44 above, where the negative tag forms are similarly used to elicit confirmation or corroboration of an utterance from a co-participant in the discourse.

Although rates of levelling to non-standard *was* in contexts of standard *were* are clearly much lower in COLT than in the preadolescent data, the first person plural and the second person singular have comparatively higher rates of levelling to non-standard *was* than other grammatical subjects in both the adolescent and preadolescent data sets.

Rates of levelling to *was* with the second person shown in Figure 3.6 above (see also Figure 3.4) suggest that the use of non-standard *was* with this grammatical person is not (as is sometimes implied in the research literature) a reliable diagnostic for contrasting contemporary northern and southern British dialects. Tagliamonte (2002a: 744), for example, explains contrastive rates of *was* levelling with the second person singular in Buckie (northeast Scotland) and Devon (southwest England) in terms of a north-south dichotomy. Although Tagliamonte's (2002a) data are drawn from much older sections of the population than those studied here, the results presented above based on London youth norms indicate that contemporary regional patterns of *was* levelling with the second person present a complicated picture which is not necessarily consistent with a north-south divide.<sup>29</sup>

Furthermore, it appears that *you was* has been embedded in London vernacular English for quite some time. Sivertsen's (1960) transcripts of the recordings of speakers resident in Bethnal Green (inner London), most of whom were born in the East End of London in the late nineteenth century, contain examples of the use of non-standard *was* with the second person:

- (48) they tell you how much they collected, for you, like, if you *was* the  
wife, or the husband (Sivertsen 1960: 245)

---

<sup>29</sup> Further evidence of the internal complexity or heterogeneity of the south of England with respect to patterns of *was/were* variation can be found in Tagliamonte (2002a). Tagliamonte (2002a: 745) notes that levelling to non-standard *was* in Devon is almost as frequent in negative contexts as in affirmative ones. Thus, in Tagliamonte's (2002a) results for Devon, there is no evidence of the polarity effect reported for southeastern dialects (i.e. a split paradigm based on positive *was* and negative *weren 't*).

The use of non-standard *was* with the second person, as in (48) above, is consistent with the findings reported in Cheshire and Fox (2006), who found relatively high rates of levelling to non-standard *was* with the second person in the vernacular of inner London elderly speakers (see Figure 3.4 above). The fact that levelling to non-standard *was* with the second person is also found in the speech of outer London preadolescents, but not in the vernacular usage of the outer London elderly speakers investigated by Cheshire and Fox (2006), may offer some support for the possible diffusion of traditional inner London vernacular patterns to outer London, perhaps as the result of population movements, as Cheshire and Fox (2006) suggest.<sup>30</sup>

As far as the rates of levelling to non-standard *was* with the first person plural are concerned, the time-depth of this pattern in London vernacular English is unknown, although it by no means seems to be a recent innovation (see also Matthews 1972: 196). Map 22 in the *Linguistic Atlas of England* (Orton *et al.* 1978) shows that *were* occurs predominantly with the first person plural in London, although *we was* is reported as being pervasive in adjacent geographical areas. Once again, Sivertsen's (1960) transcripts are a useful source of non-standard usage:

(49) as we *was* going along, we started singing (Sivertsen 1960: 238)<sup>31</sup>

Turning to rates of levelling to non-standard *was* with plural noun phrases and the third person plural pronoun, *they*, the data displayed in Figure 3.6 above show that although levelling to non-standard *was* is only marginally higher with plural noun phrases than with the third person plural pronoun in the adolescent corpus, the reverse

<sup>30</sup> Altendorf and Watt (2004:182-3) note that population movements have resulted in increased levels of admixture of the population in the southeast. They also briefly discuss internal migration patterns within the London metropolitan area, and mention the trend for young families to move out of London into the suburbs or the neighbouring Home Counties since World War II.

<sup>31</sup> *We was* appears to be socially salient in the community where I recorded the preadolescent data. Evidence for this comes from an observation that I recorded from a teacher who was participating in a training session on grammar in the school where my fieldwork was undertaken. The teacher in question specifically cited the children's use of *we was* as characteristic of their non-standard English. This remark is consistent with Wolfram and Schilling-Estes' (2003: 149) observation that levelling to *was* is 'an icon of prescriptivism,' whereas levelling to *weren't* appears to attract little overt comment.

state of affairs obtains in the preadolescent corpus, where the rate of levelling to non-standard *was* with the third person plural pronoun exceeds levelling to non-standard *was* with plural noun phrases. On the basis of these findings, it appears that the NP/PRO constraint, which is reported to be operative in northern dialects as well as some southern varieties (Godfrey and Tagliamonte 1999; Tagliamonte 2002a: 741), has relatively little impact on *was/were* variation in the adolescent data, and none at all in the preadolescent data, where the reversal of this constraint is consistent with the findings reported in Britain (2002a: 26) for the Fens. The hierarchy of subject constraints on *was* levelling in the preadolescent data, namely 2<sup>nd</sup> person > 1<sup>st</sup> person > third person plural pronoun > noun phrase (plural), differs from that reported for northern dialects and exhibits similarities with the subject hierarchy reported by Britain (2002a: 26), indicating that the findings for the preadolescents are not idiosyncratic to this age group, but may reflect a broader, possibly supraregional pattern that encompasses not only parts of the London metropolitan area, but extends into other parts of the southeast, as Britain *et al.* (2005) suggest.

Before turning to a consideration of the influence of social factors on patterns of *was/were* variation in the preadolescent data, I briefly summarise the findings relating to the effects of syntactic configuration, as well as characteristics of the subject noun (e.g. collective nouns) on variation in past *BE*.

With regard to the operation of a subject proximity constraint, whereby the non-adjacency of the subject and verb has been found to favour the use of non-standard *was* (see Tagliamonte and Smith 2000: 154), the preadolescent data remain inconclusive owing to a severely limited number of relevant contexts in the corpus. Smith and Tagliamonte (1998: 120) report that the relative marker *that* exhibits 100

percent use of non-standard *was* in Buckie English, but this is not the case in the preadolescent data, as the following examples illustrate:

(50) all the people *that were* on the Earth they *were* aliens in their private time [3M10/11]

(51) it happened to all the aliens *that were* here [23F7/8]

(52) he saw golf clubs *that was* in the house [4M10/11]

Other cases in the preadolescent corpus where syntactic configuration does not conform to the canonical structure of SUBJECT + VERB appear to be favourable to non-standard *was* (see also Pietsch 2005a: 129):

(53) inside *was* some false nails that I didn't see [19F10/11]

(54) she said '*Was* you scared?' [2M10/11]

(55) *was* they crucified? [16M7/8]<sup>32</sup>

Examples (53)-(55) may simply reflect the persistence of diachronically attested patterns which were widespread in earlier varieties of English (see further below; and Fischer 1992: 366).

There were few examples of collective nouns used in the preadolescent corpus, and those which were used predominantly consisted of the collective noun *people*. Once again, it is difficult to draw any firm conclusions on how such lexical items may pattern with past *BE* allomorphs on the basis of restricted data, other than noting that *people* occurs with both *was* and *were* in the preadolescent data:

(56) and these people *was* going on the plane [15M7/8]

(57) she couldn't understand what the English people *were* saying [23F7/8]

---

<sup>32</sup> The use of non-standard *was* in inversions is also found in the COLT adolescent corpus; e.g. *How long was you going out with him for?* [B132611/F]. The use of singular concord in these examples apparently differs from certain other non-standard varieties such as Belfast English, where it is claimed that 'singular concord is generally impossible in structures showing inversion' (Henry 1995: 42). However, Pietsch (2005a: 182) claims that the prohibition of verbal *-s* in inversions in Belfast English posited by Henry (1995) is not supported in his data from Northern Ireland. Contrasting findings for Northern Ireland may well be due to the divergent methodologies employed by researchers operating within very different theoretical frameworks.

Again, examples such as (56) may simply be remnants of earlier historical patterns (see Labov 1980: xvii) that have persisted in non-standard varieties of English. Fischer (1992: 365) points out that there was a great deal of variation in earlier varieties of English with regard to patterns of verbal concord with collective nouns, with, for example, *people* usually being treated as a singular noun in Chaucer, whereas *folk* was treated as a plural noun. Fischer (1992: 365) adds that it was only in the Late Middle English period that the tendency to treat collective nouns as requiring plural agreement became stronger, although even at this time, this tendency still depended on which particular item was being used, suggesting, as Levin (2006:340-1) argues, that variable agreement patterns with different collective nouns can be accounted for with reference to various lexical factors and processes characteristic of lexical diffusion.

Conjoined NPs also occur with non-standard *was* in the preadolescent data, as exemplified below:

(58) my real mum and my dad *was* having an argue (*sic*) [20F10/11]

(59) it was because Julie and my other friend *was* calling her names  
[16F10/11]

(60) Shaggy and Scooby Doo *was* running away [18M7/8]

The absence of subject-verb agreement in past *BE* with conjoined singular noun phrases is a longstanding phenomenon, and is attested as far back as Old English (Traugott 1992: 179; see also Visser 1963: 80; Pietsch 2005a: 129). In examples (58)-(60) agreement with the closest conjunct appears to be preferred, a pattern which is also encountered in vernacular varieties of English in several other syntactic constructions which may contain conjoined noun phrases, such as expletive *there* constructions or locative inversion constructions (e.g. *in the driveway was/were a fire*



truck and two police cars; Schütze 1999: 472). The occurrence of non-standard *was* with conjoined noun phrases may be explicable in terms of sequential processing constraints, or linear adjacency effects, resulting in agreement between the verb and the closest NP in a conjoined noun phrase construction (see Schütze 1999: 472-3).

Turning now to a consideration of social constraints on *was/were* variation in the preadolescent data, Figures 3.7 and 3.8 show the distribution of rates of levelling to non-standard *was* in contexts of standard *were* according to age and gender respectively.

FIGURE 3.7

PERCENTAGE OF NON-STANDARD *WAS* IN CONTEXTS OF STANDARD *WERE* BY AGE

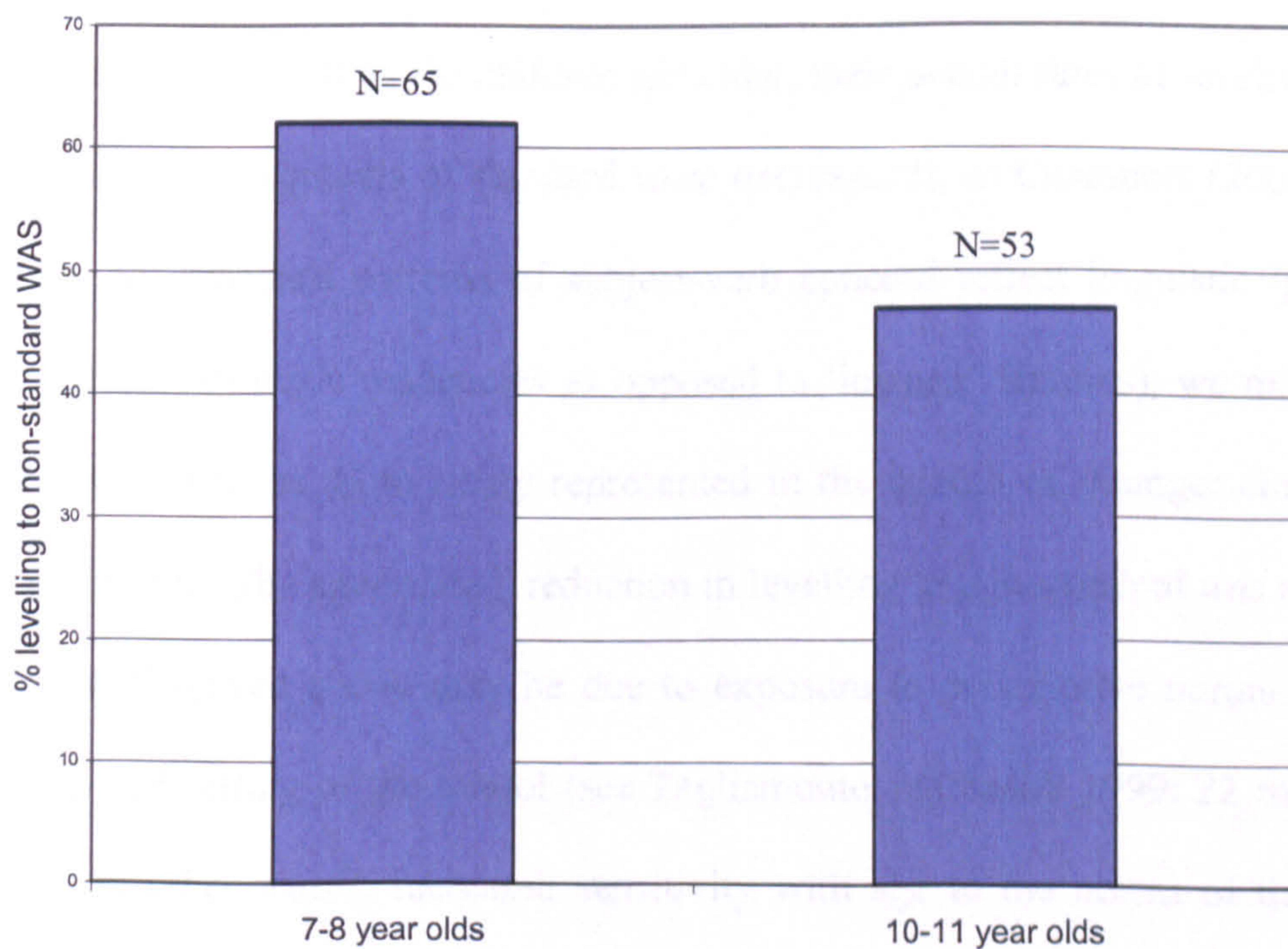


FIGURE 3.8

PERCENTAGE OF NON-STANDARD *WAS* IN CONTEXTS OF STANDARD *WERE* BY GENDER

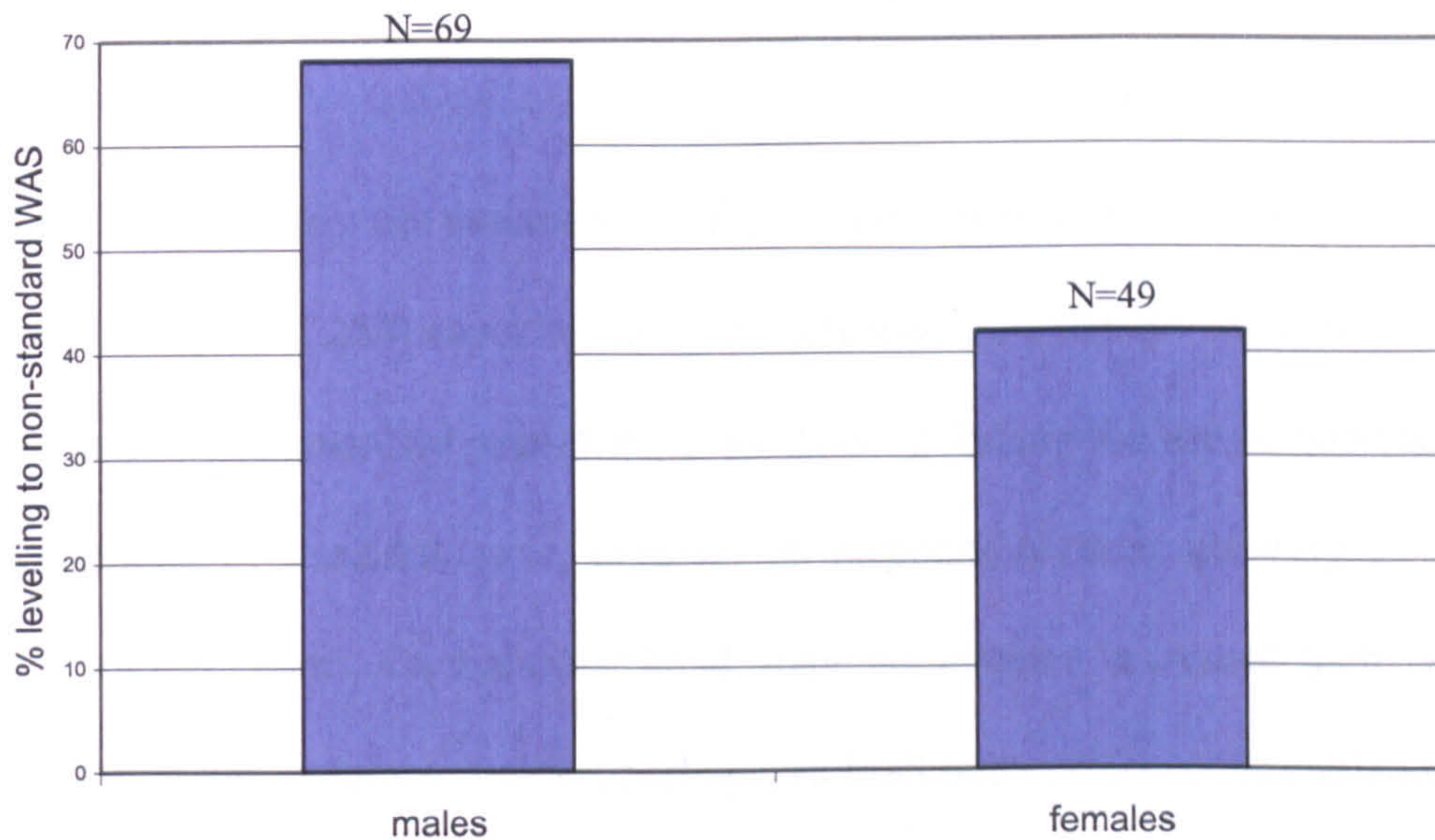


Figure 3.7 shows that as the children get older, their overall rates of levelling to non-standard *was* in contexts of standard *were* decrease. If, as Chambers (2003: 266-71) claims, non-standard patterns of subject-verb concord reflect linguistic ‘primitives’ (i.e. natural linguistic tendencies as opposed to ‘learned’ features), we might expect such features to be more richly represented in the speech of younger children than older children. The age-related reduction in levelling to non-standard *was* apparent in the data displayed above may be due to exposure to prescriptive norms within the institutional setting of the school (see Tagliamonte and Smith 1999: 22 for a similar point). In other words, increased sensitivity with age to the norms of the standard written language via the acquisition of literacy practices may have an overall disavouring effect on rates of levelling to non-standard *was*.

It is not just age which influences rates of levelling to non-standard *was* but also gender. The data displayed in Figure 3.8 are consistent with the canonical pattern

whereby a non-standard variant is used more frequently by males than females (Labov 1990: 205; Milroy and Gordon 2003: 171; Chambers 2003:116). On the basis of the data shown above, gender-affiliated differences in the use of non-standard *was* appear to emerge early, and are found in the preadolescent corpus among children as young as 7.

In her study of variation in subject-verb agreement in Inner Sydney English, Eisikovits (1991: 253) reports that with increased age, female adolescents decreased their use of non-standard grammatical variants, including the use of non-standard *was* in contexts of standard *were*, possibly in response to their growing awareness of grammatical norms. On the other hand, male adolescents increased their use of non-standard forms as they got older, perhaps, as Eisikovits (1991: 253) contends, because for males such forms embody sociosymbolic functions associated with a gendered discourse style.<sup>33</sup> However, in the case of the preadolescents examined here, although males and females exhibit contrastive rates of levelling to non-standard *was*, the females' use of non-standard *was* remains relatively constant across the two age cohorts (i.e. 39% for the 7-8 year old girls; and 44% for the 10-11 year old girls). The males, on the other hand, show a more dramatic difference in rates of levelling to non-standard *was*: the 7-8 year olds have the highest rates of levelling at 79%, which drops to 52% for the 10-11 year old males. Thus, it is not just differences between males and females that are of interest here, particularly in the youngest age cohort, but also differences within the same gender that are important too (see Cameron 2005: 487).

---

<sup>33</sup> Cheshire (1982b: 163) reports that Reading male adolescents used non-standard *was* more frequently than females. However, Cheshire (1982b: 163-4) notes that for certain groups of girls (i.e. girls who did not necessarily behave in conventionally sanctioned ways), non-standard *was* functioned as a marker of adherence to vernacular culture.

### 3.5.1 *Multivariate analysis*

In order to assess the relative contributions of internal and external factors in determining patterns of *was/were* variation in the preadolescent data, I now turn to a multivariate analysis of the data to investigate which factors exert statistically significant effects on the use of non-standard *was* in contexts of standard *were*. I also carry out a similar multivariate analysis of the COLT adolescent data for comparative purposes, focusing in particular on the hierarchy of constraints (i.e. ordering of factor weights within a factor group; Poplack and Tagliamonte 2001: 93) in order to locate patterns of convergence and divergence in the variable grammars underlying the adolescents' and preadolescents' use of non-standard *was*.

Owing to the paucity of negative contexts in the preadolescent data, which are insufficient to permit multivariate analysis, I focus here on the contribution of factors to the probability of non-standard *was* in contexts of standard *were* in affirmative contexts.

As in previous studies of past *BE*, the results displayed in Table 3.9 below show that grammatical person is statistically significant in both corpora, although the magnitude of this effect, as indicated by the range, is somewhat greater for the adolescents than the preadolescents. The constraint hierarchy for this factor group exhibits some similarities in both corpora: both the adolescents and the preadolescents favour the first person plural with non-standard *was*, and in both data sets, the second person singular is ranked in second place in the subject hierarchy after the first person plural. Note too that in both corpora, the first person plural is distinguished from the other grammatical persons by having a comparatively stronger favouring effect on the realisation of non-standard *was*. One feasible explanation for this finding could be, as

Smith (2000: 65) suggests for Buckie English, that *we was* has acquired the status of a sociolinguistic marker for London youth.

TABLE 3.9

TWO INDEPENDENT MULTIVARIATE ANALYSES OF NON-STANDARD *WAS* IN CONTEXTS OF STANDARD *WERE* IN THE PREADOLESCENT AND COLT ADOLESCENT CORPORA (AFFIRMATIVE CONTEXTS ONLY)

CORPUS	PREADOLESCENT CORPUS			COLT ADOLESCENT		
	FW	%	Total N	FW	%	Total N
<b>INPUT</b>	0.6			0.08		
<b>FACTORS</b>						
<b>Speaker Gender</b>						
Female	0.36	42	116	0.54	10	337
Male	0.66	70	99	0.45	7	283
<b>Range</b>	30					
<b>Age</b>						
7-8 years	0.55	62	105	***NOT INCLUDED***		
10-11 years	0.45	48	110			
<b>Grammatical Person</b>						
<i>we</i>	0.62	63	82	0.69	16	133
<i>you(singular)</i>	0.49	57	7	0.53	9	236
<i>full NP</i>	0.36	41	51	0.40	5	75
<i>they</i>	0.46	55	75	0.36	4.5	176
<b>Range</b>	26			33		

In other respects, however, the subject hierarchy is not entirely consistent across both data sets, as evidenced by the fact that the third person plural pronoun is ranked above full NPs in the preadolescent corpus, but not in the adolescent corpus. Nevertheless, although the NP over PRO ranking is sustained in the adolescent corpus, the factor weights allocated to these grammatical subjects are not sharply contrastive. In the preadolescent data, on the other hand, the third person plural pronoun is clearly ranked over third person plural NPs. The proximity of the preadolescent fieldwork site to other locations in Essex which instantiate a similar reversal of the NP/PRO constraint (see Britain 2002a; Britain *et al.* 2005) supports the possibility that this pattern is geographically widespread (see also the data for inner London adolescents in Figure 3.4 above abstracted from Cheshire and Fox 2006).

The data displayed in the multivariate analyses above, viewed in conjunction with the distributional data for adolescents in outer and inner London discussed by Cheshire and Fox (2006), suggest that if the Northern Subject Rule was once operative in the southeast, it is no longer a potent determinant of allomorphic variation in past *BE* in the vernacular of contemporary London youth.

The data displayed in Table 3.9 above also show that it is not just internal factors which constrain patterns of variation in the use of non-standard *was*: social factors are also strongly implicated in the preadolescent corpus. The difference between preadolescent males' and females' use of non-standard *was* is statistically significant.<sup>34</sup> In fact, the magnitude of the gender effect is even greater than that for grammatical person. This contrasts with the corresponding factor weights in the

---

<sup>34</sup> Like Romaine's (1984a) Scottish preadolescents, several of the preteen speakers I interviewed were keen to share their perceptions of the differences between girls' and boys' speech. They were also clearly aware of the social evaluations of certain variables in their community, and reported that they were corrected (most frequently by mothers or grandmothers) for using non-standard grammatical variants such as *ain't*, or non-standard phonological variants such as the glottal stop.

adolescent data in which gender fails to attain statistical significance, although there is a slight tendency in the adolescent data for females to use non-standard *was* more than males.

### 3.6 SUMMARY

The results presented above for the preadolescent speakers provide some confirmation of Anderwald's (2001: 14) tentative suggestion that there is a change towards a mixed system in past *BE* among the younger generation of Londoners. In line with the findings presented in Britain (2002a) for the Fens, the preteen data and the COLT adolescent data indicate that a more focused sub-system based on a paradigmatic split between positive *was* and negative *weren't* appears to be emerging in the vernacular used by London youth. However, this particular interpretation of the data needs to be qualified by the findings from Cheshire and Fox's (2006) recent study of past *BE* in London which suggest that other allomorphic variants such as non-standard *wasn't* compete quite vigorously with non-standard *weren't* in outer London.

By drawing on several data sets, and situating the synchronic data examined in this chapter in relation to evidence of patterns of *was/were* variation in older varieties of English, I have been able to show that differentiated rates of levelling to non-standard *was* across the verbal paradigm in the preadolescent and COLT adolescent corpora do not always conform to the patterns that can be extrapolated from the existing historical literature. Thus, whereas other researchers such as Smith and Tagliamonte (1998) have interpreted the internal grammatical factors that constrain *was/were* variation in certain northern dialects of English with reference to the persistence of diachronic patterns of variation, this is less easily accomplished in the

case of London English, particularly with regard to the subject hierarchy, where there is evidence that past *BE* appears to have undergone grammatical restructuring (see also Nevalainen 2006).

The reduction of variable phonetic forms, as well as the apparent simplification of allomorphic variation in past *BE* evident in the preadolescent and COLT adolescent data are consonant with levelling tendencies in other areas of the grammar that are claimed to be operative in London and the southeast. For example, Britain (2002b: 616-7) notes that contact between various London, RP and local southeastern varieties is resulting in the reduction of marked minority dialect forms while also fostering the concomitant spread of supralocal features such as the fronting of /θ/ and /ð/ to /f/ and /v/ as well as the vocalization of /l/ and the glottalization of /t/.

Restructuring phenomena in past *BE*, particularly in terms of the realignment of allomorphic variants to mark polarity distinctions in the speech of London preadolescents, adolescents and young Fenlanders point to evolving supraregional vernacular norms in this grammatical sub-system. However, the extent to which the remorphologization of *was/were* to mark polarity distinctions is a feature of other dialects spoken in urban (and rural) areas of Britain, especially among the younger generation, must await further investigation as systematic studies of morphosyntactic variation among adolescent and preadolescent age cohorts in a range of geographical locations in Britain are generally lacking (see however Cheshire *et al.* 1999).

The findings presented above suggest that important grammatical constraints previously cited in the research literature as conditioning *was/were* variation in British dialects are either inoperative or vitiated in the speech of children and adolescents in the southeast. The NP/PRO constraint, which is taken to be the prototypical feature of the Northern Subject Rule (Pietsch 2005a: 161), was not instantiated in the speech of



the preadolescents, who share the reversal of this constraint with other eastern /southeastern varieties of English (Britain 2002a; Britain *et al.* 2005; see also the distributional results for the inner London adolescents in Cheshire and Fox 2006). On the other hand, the statistical analysis of the COLT adolescent data showed that the NP/PRO ranking historically associated with the Northern Subject Rule was preserved, albeit marginally.

Can the different constraint rankings in the preadolescent and COLT adolescent corpora be reconciled with one another? Within the comparative sociolinguistic enterprise, as exemplified by Poplack and Tagliamonte (2001), parallels across constraint hierarchies are an important diagnostic for assessing the relationship of linguistic forms and for determining the extent to which underlying variable grammars are shared across varieties. According to the analytical procedures outlined in Poplack and Tagliamonte (2001), the absence of a parallel ordering for the NP/PRO constraint in the preadolescent and COLT adolescent data could point to divergent underlying variable grammars with respect to this grammatical feature. While it may well be the case, as I have already indicated, that London is a linguistically heterogeneous area with regard to past *BE* variation, I submit that it is not necessary to infer from the statistical analyses presented in Table 3.9 above that the preadolescents and the COLT adolescents are operating with two very different variable constraints (i.e. a Northern Subject Rule versus an East Anglian or Southern Subject Rule) as far as third person plural contexts are concerned.<sup>35</sup> Witness the fact

---

<sup>35</sup> Rupp (2004) has recently claimed that verbal *-s/* non-standard *was* (e.g. *the pupils is in school/ the pupils was in school*) may have a different function across dialects. On the basis of a diachronic analysis of the differential evolution and status of *-s* in northern and southern dialects, Rupp (2004) proposes that in northern dialects *-s/non-standard was* may mark the absence of agreement, whereas in East Anglian dialects, *-s/non-standard was* can mark agreement. Embedding her analysis within a generative framework, and assuming that agreement between subject and verb takes places in a separate agreement phrase (see Pollock 1989), Rupp (2004) argues that verbal *-s/non-standard was* forms that mark agreement raise to AGR where they can assign nominative Case to pronominal subjects in (Spec, AGRP).

that not only are the factor weights in the adolescent data for third person plural NPs and the pronoun *they* minimally different from one another, they both have a moderately inhibiting effect on the occurrence of non-standard *was*. This may possibly reflect the gradual attrition of third person plural NPs as a factor favouring non-standard *was* over time (for some quantitative evidence of this, see Nevalainen 2006: 363).

Tagliamonte (2006c) has recently posited two versions of the NP/PRO constraint for verbal *-s* to account for broad cross-variety differences in the distribution of this feature: a European-American version of this constraint which shows a strong subject constraint where NPs favour *-s* over pronouns, whereas the corresponding African American Vernacular English version of this constraint relaxes the noun phrase effect (see Wolfram 2000:54).<sup>36</sup> Extrapolating from this observation, it may be the case in the contemporary vernaculars of young London speakers that a similar relaxation of the NP/PRO constraint is taking place in past *BE* resulting in either the neutralization of this constraint or its reversal, possibly as a result of levelling processes as well as contact with other southeastern varieties. Of course, this hypothesis remains speculative, particularly as it is difficult to ascertain in the absence of more robust sociodiachronic evidence whether the reversal of the NP/PRO constraint in the preadolescent data and in other southeastern varieties is the result of recent developments, or whether this phenomenon has a more extensive heritage in London (and the southeast). The fact that the distributional data displayed in Figure 3.4 for the outer London elderly investigated by Cheshire and Fox (2006) do not

---

<sup>36</sup> The operation of a strong NP/PRO constraint in American dialects is debatable. Recent research by Trüb (2006) suggests that in southern American vernacular varieties of English the effect of the NP/PRO constraint has weakened over time. According to Trüb (2006: 262), the trend towards higher rates of *was* levelling in modern varieties of southern American vernacular English suggests that 'analogical leveling of the [past-tense *BE*] paradigm has increased for all persons and superseded the subject type constraint in the past 150 years.'

exhibit the NP/PRO distinction suggests that the diachronic embeddedness of alternative configurations to the NP/PRO constraint in the Greater London area is a fruitful area for further exploration.

Finally, returning specifically to the preadolescent corpus, the results of the investigation into *was/were* variation in this age group indicate not only that linguistic constraints on morphosyntactic variation are acquired early (see also Smith *et al.* 2007), but also that *was/were* variation in the group of preteen speakers investigated here is socially constrained, as evidenced by the statistically significant differences in levelling to non-standard *was* between males and females.

In the following chapters, I show how other aspects of variation from different levels of the grammar are similarly open to social intervention in this age group, sometimes in quite nuanced ways.

# VARIATION IN THE RELATIVIZATION SYSTEM

### 4.1 INTRODUCTION

This chapter builds on existing research relating to variation in the relative marker system in British English by investigating linguistic and social factors which constrain the choice of relative markers in the preadolescents' vernacular.<sup>1</sup> The system of relativization in contemporary varieties of English is fertile territory for the exploration of patterns of synchronic variation between competing forms (see for example Ball 1996; Tottie and Rey 1997; Tagliamonte 2002b; Tagliamonte *et al.* 2005; Herrmann 2005).

The examination of synchronic patterns of variation in the preadolescents' speech is contextualized by briefly examining the historical evolution of relativization in English, and by reviewing recent research relating to variation in the relative marker system in contemporary British vernaculars.

The focal point of the chapter is the quantitative analysis of relative marker usage in restrictive relative clauses where there is a great deal of variation in the patterning and distribution of competing forms (see Ball 1996: 243). Previous research has particularly highlighted the differential frequencies of *wh*-forms in the relative marker paradigms of different varieties of contemporary English which can be located at different points along the trajectory of linguistic change in terms of levels of *wh*-marker usage.

---

<sup>1</sup> Following Tottie and Rey (1997: 225), I use the term 'relative marker' or 'relativizer' as a cover term for the variants examined in this study.

The springboard for much of the recent variationist research on relativization in British English (see e.g. Tagliamonte 2002b; and Tagliamonte *et al.* 2005) is Romaine's (1982: 212) testable claim that 'the infiltration of *wh* into the relative system can be seen as completed in the modern written language [...] but it has not really affected the spoken language.' Several later studies (e.g. Geisler 2002; Tagliamonte 2002b; and Tagliamonte *et al.* 2005) have bolstered this claim. Nevertheless, not all dialect areas have been thoroughly investigated using an accountable statistical methodology. For example, Tagliamonte (2002b: 155) notes that there are reports that some southeastern dialects are surprisingly *that*-less, and Poussa (1994: 424) suggests that this could be diagnostic of a major east-west dialect division in southern dialects.

Through a systematic distributional and multivariate analysis of the data in the London preadolescent corpus, I aim to show that although the *wh*-relativization strategy has made inroads into the children's informal spoken language, it nevertheless competes with the older predominant relative marker, *that*.<sup>2</sup> I argue that the findings of the quantitative exploration of the data can furnish insights into preadolescents' acquisition of syntactic variation, and can also contribute to descriptions of regional grammatical variation in the southeast of England.

## 4.2 OVERVIEW OF THE HISTORY OF RELATIVIZATION IN ENGLISH

In Old English, the demonstrative pronouns *se*, *seo*, and *þæt* could be used in a relativizing function. Relative clauses could also be constructed using the

---

<sup>2</sup> I sidestep the question of whether *that* in a relativizing function is a pronoun or a complementizer as this issue is not central to the variationist analysis at the core of this chapter. For the record, Sag (1997: 462) notes that relativizer *that* is often thought to be the same as the complementizer as in *I thought [that you were sick]*. Hudson (1990) and Van der Auwera (1985) present evidence that relativizer *that* has pronominal status (see e.g. the use of possessive *that's* in example (9) below).

indeclinable particle *þe*, which could be placed after a substantive or pronominal element, as in the following example:

(1) sealde þæm munucum corn genog *þe* wæron æt Hierusalem

‘Gave enough corn to the monks who were in Jerusalem’

(*Or* 6 4.260.9; cited in Traugott 1992: 226)

During the Middle English period, the indeclinable particle *þe* was gradually lost, and the demonstrative pronouns *se*, *seo* and *þæt* were eventually replaced by indeclinable *that*. The emergence of indeclinable *that* appears to have originated in the north and subsequently diffused to other dialect areas. According to Fischer *et al.* (2000: 91), *that* was the pre-eminent relativizer in the thirteenth century and was used in both restrictive and non-restrictive clauses, as well as with animate and inanimate antecedents.

The major change to occur in the relativization system in the Middle English period was the use of *wh*-forms (*whom*, *whose*, *which* and *what*) in a relativizing function. Romaine (1980: 221) presents evidence which suggests that the *wh*-relativization strategy appears to have entered the language in the most complex written styles in the least frequently relativized positions. The subsequent spread of the *wh*-forms competed with the older primary strategy involving the relativizer *that*, and resulted in the addition of a new relativization strategy which was stylistically diagnostic of more formal styles. According to Mustanoja (1960: 191), the use of *wh*-forms in relative clauses arose from types of indirect questions where the interrogative character of the *wh*-pronouns became weakened leading to their use first as generalizing relatives, as illustrated in (2) below, and subsequently as pronouns that were co-referential with clearly identifiable antecedents in a strictly relativizing function.

(2) *Hwam* mai he luve troweliche *hwa* ne luves his broþer.

Whom may he love truly, who (ever) does not love his brother.

(*Wooing of Lord* 275; cited in Mustanoja 1960: 192)

However, the incursion of *wh*-forms into the English relativization system was a gradual process, with different pronominal forms infiltrating the system at different rates and times, entailing a functional reallocation of forms in the relative paradigm as time progressed. For example, Fischer *et al.* (2000: 92) note that *which* was initially infrequent, but gradually began to compete with relative *that*, and was used with animate and inanimate antecedents, as well as sentential antecedents (see Romaine 1982: 213). Ball (1996: 246) observes that by the sixteenth century, *which* was well established with non-personal antecedents, and the use of *which* with personal antecedents began to give way to *who*.

In contrast with the inflected forms *whom* and *whose*, which were in use since the early Middle English period (Romaine 1980: 223), *who* was a comparatively late arrival to the English relativization system, and was used at first in stereotyped valedictory formulae in letter writing, as in the following epistolary closing phrase from a fifteenth century letter (where square brackets denote the boundaries of the relative clause):

(3) and that knoweth God, [who have you in his blessed keypyng ]

(Stonor, before 1431; cited in Nevalainen and Raumolin-Brunberg 2002: 112)

The entrance of *who* into the relativization system appears to have been sensitive to an animacy hierarchy, with *who* being used initially when God was the antecedent, then with personal-name antecedents and finally co-occurring with entities lower on the scale of personhood (Nevalainen and Raumolin-Brunberg 2002: 112). According to Ball (1996: 246-7), the rise of *who* in a restrictive relativizing function in the early

sixteenth century laid the foundation for the contemporary distinction in the standard language between personal *who* and non-personal *which*.

During the Early Modern English period, Ball (1996: 250) claims that the most significant development in the relative paradigm involved the functional shift of *that* away from personal subject antecedents in conjunction with a growth in the frequency of *who* with personal subjects. This period also witnessed an increase in zero relativization, with the frequency of zero object relativization surpassing the frequency of zero subject relativization, thereby reversing the state of affairs which obtained in Middle English, where zero subject relatives were still the predominant type (see Tottie and Harvie 2000: 202). Examples such as (4) below, taken from a seventeenth century text, came to be proscribed in the standard written language, although zero subject relatives are still found in contemporary non-standard varieties of English.

(4) I bring him news [  $\emptyset$  will raise his drooping spirits]

( Dryden , *All for Love*, cited in Tottie and Harvie 2000: 202)

Contra Bever and Langendoen (1971), Romaine (1982: 78-80) claims that the demise of zero subject relatives in written English was not associated with perceptual difficulties, but was more plausibly attributable to external social influences associated with standardisation, which led to subject relative deletion being equated with informal registers and non-standard speech (Romaine 1982: 173).<sup>3</sup>

The preceding overview of the evolution of relativization in the history of English profiles the emergence of a tripartite division of relativization strategies (i.e. *WH*, *that* and  $\emptyset$ ) characteristic of the standard (written) language. However, less is

---

<sup>3</sup> Aitchison (1983: 636) claims that a sociolinguistic explanation for the demise of subject relative deletion (i.e. its association with non-standard speech) is not incompatible with a psycholinguistic explanation (i.e. perceptual difficulty). A combined sociolinguistic and psycholinguistic explanation could therefore account for a reduction in the frequency of deleted subject relatives.



known about diachronic variation in the use of regionally circumscribed forms which continue to participate in the patterning of relative markers in vernacular paradigms. The most relevant of these regionally circumscribed variants for the synchronic analysis presented in Section 4.3 below are the relative markers *what* and *as*. The latter appears infrequently in Middle English from the fourteenth century onwards, particularly in combination with *such* (Mustanoja 1960: 202), and is attested in twentieth century dialects in the Western Central region, judging from the evidence presented in the *Linguistic Atlas of England* (map S5, Orton *et al.* 1978). Edwards (1993: 229) also claims that *as* is used in a relativizing function in contemporary southeastern non-standard speech (e.g. *the food [as I bought]*).<sup>4</sup>

Mustanoja (1960: 194) claims that *what* has been used as a relative marker since the eleventh century, although it is infrequent in Middle English and occurs mainly after indefinite antecedents such as *all* or *nothing*.<sup>5</sup> Early evidence furnished by the nineteenth century dialectologist, Forby (1830/ 1970: 138), suggests that *what* was pervasive as a relative marker in eastern dialects. Several decades later, Wright (1905, §423) notes that '*what* can be used when it refers to persons as well as to inanimate objects in some of the north-midland counties and in nearly all the counties south of the north midlands.' Based on evidence from the *Survey of English Dialects* (Orton *et al.* 1963-71), Poussa (1988: 450) observes that *what* has been traditionally used as a relative marker in the rural catchment area of London as well as in the urban

---

<sup>4</sup> Herrmann (2005: 36, 57) points out that relativizer *as* is a recessive feature of contemporary English dialects, and notes that it plays 'a minor role' in the data emerging from her cross-dialectal analysis of relative clauses in the *Freiburg English Dialect Corpus (FRED)*. According to Herrmann (2005: 56), relativizer *as* is nominally instantiated in only three regions: the central Midlands, the central North [of England], and Northern Ireland.

<sup>5</sup> The co-occurrence of non-standard *what* with indefinite antecedents such as *all* is also attested in Early African American Vernacular English (see Van Herk and Poplack 2003: 251). Kautzsch too (2002:162) notes that non-standard *what* occurs at rates as high as 51% in Early African American Vernacular English.

traditional working-class district of east London.<sup>6</sup> I return to the status of non-standard *what* in contemporary British dialects in 4.3 below.

### 4.3 VARIATION IN THE RELATIVE MARKER SYSTEM IN CONTEMPORARY BRITISH VERNACULARS

#### 4.3.1 *The status of the wh-relativization strategy in contemporary British dialects*

Using data presented in the *Survey of English Dialects* (Orton *et al.* 1963-71), Nevalainen and Raumolin-Brunberg (2002: 112) observe that the regional distribution of subject relativizers suggests that there is a broad geographical division between two traditional relativization strategies: a particle strategy where *that* and, to a lesser extent, *at*, are the predominant forms in the North; and the *wh*-strategy which prevails in eastern and southern vernaculars. The dichotomy between a *wh*-strategy and a non-*wh*-strategy, which is claimed to be existent to different degrees in contemporary vernacular varieties of English (see Seppänen 1999: 15), has been broadly confirmed by recent large-scale quantitative studies. Tagliamonte (2002b: 153) found that the marker *that* was the predominant relativizer (54%) in her analysis of the overall distribution of relative markers in six varieties of English from different geographical locations in the north and south. By contrast, the proportion of *wh*-markers was markedly lower, with the most frequent *wh*-marker, *who*, accounting for only 13% of the relativizers used. Tagliamonte's (2002b: 162-3) results indicate that *wh*-forms have not generally infiltrated the non-standard British dialects she analysed, whereas *that* appears to be generalising as a supraregional relative marker. These findings partially confirm Romaine's (1982: 60) claim that 'there remains a significant residue

---

<sup>6</sup> See Poussa (2002b: 18), who notes that *what* is a 'stereotypical feature of the sociolect of the east-London working class,' and calls for more research among younger speakers to ascertain whether or not its use is receding owing to its apparent stigmatisation.

of *that* forms in the spoken language' in contrast with the modern written language 'which shows *wh*-forms in nearly all styles and all syntactic positions.'

Furthermore, Tagliamonte's (2002b) study adds considerable refinement to the putative geographical divisions between dialects which show higher levels of *wh*-relativizers in comparison with those in which the frequency of *wh*-markers is lower. Tagliamonte (2002b: 164) claims that rather than a north-south division being a strong conditioning factor in the use of certain relativizers in preference to others, a major determinant affecting relative marker usage is the relative proximity of dialects to mainstream norms. More peripheral dialects in the north and southwest show similarities in the distribution of relative markers, whereas the dialect of York in the north east, which is comparatively more permeable to mainstream norms (see Tagliamonte 1998b), reveals some interesting differences when compared with northern peripheral varieties. In terms of broad distributional patterns of usage, Tagliamonte (2002b: 161) found that York English favoured *who* as a subject relative marker markedly more so than other vernaculars she examined. However, in an apparent-time analysis of relative marker usage in York English, Tagliamonte (2002b: 161-2) found significant inter-generational differences in speakers' use of the relative marker *who*: whereas older speakers (36-60, and 60+) favoured *who* as a subject relative marker, the youngest generation of speakers did not, and showed an increase in the use of *that* in subject position with personal antecedents. One conclusion that might be drawn from this apparent-time trend is that the use of *who* is recessive when viewed cross-generationally. This finding receives additional support from other studies such as Geisler's (2002) on Ulster English. On the basis of data derived from the *Northern Ireland Transcribed Corpus of Speech*, Geisler (2002: 135) concludes

that Ulster English is predominantly *wh*-less, and where *wh*-forms do occur, they are restricted to a small group of speakers.

Other studies which discuss the use of *wh*-forms in northern dialects (e.g. Beal and Corrigan 2002), highlight the essentially conservative nature of northern dialects vis-à-vis southern ones (see also Trudgill 1990: 65-78). Beal and Corrigan (2002:125) note that although *wh*-forms have made inroads into Tyneside English, this variety still remains more resistant to *wh*-infiltration than non-standard varieties in the south such as Reading English (see Cheshire 1982a). However, in a recent study, Beal and Corrigan (2005) mention that there are differences between northern varieties in their use of *wh*-forms which warrant further investigation. I return to this issue below.

In comparison with the evidence of competition between relative *who*, the principal *wh*-contender, and *that* in regional dialects, there is generally less consensus about the degree to which non-standard *what* is instantiated in contemporary dialects, although there are some indications that its robustness varies according to geographical location. Cheshire *et al.*'s (1993) questionnaire survey of variation in the grammatical usage of British schoolchildren revealed that *what* was reported frequently as a relative marker in British urban centres.<sup>7</sup> However, Tagliamonte (2002b: 148) comments on the extreme infrequency of this form in the six varieties of English she analysed based on data from the lowlands of Scotland, as well as northern and parts of southern England. In Tagliamonte *et al.*'s (2005: 87) recent large-scale

---

<sup>7</sup> Some researchers (e.g. Flynn and Lust 1980) have claimed that children's use of *what* as a relativizer is motivated by developmental factors on the grounds that *what* is commonly used in headless relative clauses, which children appear to acquire before they master headed relative clauses. I do not pursue this argument further, mainly because the preadolescents examined here seemed to have no difficulty in producing a number of headed relative clauses. Furthermore, in comparable data sets from preadolescent speakers recorded in other geographical locations, the overall frequency of *what* is lower than the proportion reported here. Romaine (1984a: 48) notes that *what* only accounts for 6% of the relative markers she examined in the speech of Edinburgh preadolescents aged 6-10 years. The higher incidence of *what* in my own data set, coupled with the presence of this feature in the speech of adults in the area where I recorded the data, suggests that the children are acquiring a productive vernacular feature of their community.

quantitative survey of relative markers in four geographically peripheral communities in Lowland Scotland, northwest England and Northern Ireland, non-standard *what* only accounts for 0.4% of the total number of relative markers used (N= 1922). Similarly, Beal and Corrigan (2005: 222) report that they found only one example of *what* in the *Tyneside Linguistic Survey* recorded in the 1960's, and no examples from the later *Phonological Variation and Change* project based in Newcastle in 1994.<sup>8</sup> By contrast, Beal and Corrigan (2005: 217) found *what* to be more frequent in vernacular data from Sheffield than in Tyneside dialects, even though the *wh*-relativization strategy was more predominant overall in Tyneside than in Sheffield.<sup>9</sup>

In southern and eastern England, the frequency of non-standard *what* appears to be higher (Herrmann 2005: 58). Van den Eynden Morpeth (2002: 182) points out that *what* is found in Dorset English in the southwest of England, where it is used in both restrictive and non-restrictive relative clauses. Furthermore, the dialects of East Anglia, considered by Herrmann (2005: 58) to be the 'heartland' of *what*, appear to be a particularly rich repository of examples of its non-standard use as a relativizer. According to Poussa (2002b: 17), *what* is robustly represented in the traditional rural dialects of East Anglia and Cambridgeshire. Similarly, Peitsara's (2002: 176) study of the Suffolk dialect leads her to conclude that *what* is 'the most useful all-purpose relativizer, while *that* tends to be favoured in the subject function.' Some of the strongest claims about the robustness of non-standard *what* based on an analysis of its occurrence in the *Freiburg English Dialect Corpus* are found in Herrmann (2005: 58, 95), who suggests that *what* is evolving as a supraregional non-standard marker which appears to have diffused from East Anglia and is 'radiating out through the adjoining

---

<sup>8</sup> Both these databases comprise *The Newcastle Electronic Corpus of Tyneside English* (NECTE).

<sup>9</sup> In spite of the greater overall prevalence of *wh*-forms in Tyneside than in Sheffield in Beal and Corrigan's data, the authors (2005: 223) state clearly that 'there is certainly no evidence of a steady northward drift of *wh*-infiltration.'

Midlands and the Home Counties, especially London, to the Southwest' and is spreading via large urban centres 'across Britain geographically, both numerically and in terms of syntactic environments.'<sup>10</sup> Ball (1996: 240-1) points out that if we classify relativizer *what* as a *wh*-pronoun, then dialects which contain relatively high proportions of non-standard *what* furnish additional evidence that the *wh*-strategy has infiltrated varieties of colloquial English.<sup>11</sup>

#### 4.3.2 *Subject relative deletion in British dialects*

Another important aspect of regional relative marker paradigms in contemporary British dialects concerns the status and distribution of zero relative markers. One of the salient features that distinguishes informal spoken varieties from the standard written language is the occurrence of zero subject relatives in the former but not in the latter. Zero subject relatives are encountered in many vernacular varieties, as illustrated in the examples below :

(5) and there was a select number of us [Ø became members]

(Ayr, Scotland; cited in Macaulay 1991: 68)

(6) tell us the one about the lady [Ø couldn't get a lift ]

(Reading English; cited in Cheshire 1982a: 74)

(7) it was a lot of schoolchildren [Ø were buried in Wales ]

(York English; cited in Tagliamonte 2002b: 158)

---

<sup>10</sup> Herrmann's (2005: 58) claim that *what* is most frequent in the south of England, and least frequent in the north warrants further investigation. Cheshire *et al.* (2005a: 151, Table 5.4) report lower frequency indices for non-standard *what* in the speech of working-class adolescents in Milton Keynes (frequency index =3.2) and Reading (frequency index =3.8) in the south of England than in Hull in the north of England (frequency index =25.5).

<sup>11</sup>This is not an uncontroversial point. Other researchers (e.g. Radford 1988: 522-3) analyse non-standard *what* as a complementizer on the basis that it is not inflected for case (see, however, Seppänen 1999 for further discussion and counterarguments).

(8) I've got something humorous [ $\emptyset$  happened to me], one thing I'll never forget (Norwich English; cited in Hughes and Trudgill 1979: 46)

Dasinger and Toupin (1994: 468, f.n. 12) claim that subject relative deletion in presentational constructions such as *there was a farmer [ $\emptyset$  had a dog]* can be explained in discourse-pragmatic terms: the relative clause in presentational constructions has the pragmatic status of an asserted independent clause and so retains the grammatical characteristics of one (see also Section 4.5)<sup>12</sup>

Henry (1995: 131) observes that constructions which permit zero subject relatives exhibit the characteristics of topic-type structures. Henry (1995: 134) further claims that zero subject relatives are root-type clauses which are able to have 'root null subjects,' and are constructions found in non-pro-drop languages (see Rizzi 1991).

Whatever the precise status of constructions permitting zero subject relatives, it is clear from the research literature that these structures are pervasive in many varieties of non-standard English. However, what appears to be difficult to establish are the grammatical constraints operating on zero subject relative constructions (see Tottie and Harvie 2000: 225). I address this issue in more detail in the following section.

---

<sup>12</sup> Dasinger and Toupin (1994: 468, f.n.12) observe that there is a tendency in German presentational relatives for the relative clause to exhibit verb-medial order rather than the verb-final order characteristic of subordinate clauses in standard German. Compare colloquial German: 'es war einmal ein Mann, der *verstand* allerlei Künste' 'there was once a man who knew all kinds of tricks' with standard German: 'es war einmal ein Mann, der allerlei Künste *verstand*'

## 4.4 METHODOLOGY

I turn now to a discussion of the factors selected for the investigation of variation in the use of relative markers in restrictive relative clauses in the preadolescent corpus. I describe the methodology used for circumscribing the variable context, and briefly discuss the claims and observations generated by previous research which motivate the selection of particular factor groups for the multivariate analysis of the data.

### 4.4.1 *Deciding what to count*

Most variationist accounts of relativization strategies have focused on restrictive and, to a lesser extent, non-restrictive relative clauses. Keenan (1985: 169) claims that the syntax of non-restrictive relative constructions largely resembles that of restrictives apart from a few minor differences. Restrictive relative clauses specify a subset within a larger set of referents, whereas non-restrictive relatives can be used to comment further on or predicate something about an identified referent (see Dasinger and Toupin 1994: 460). Keenan (1985: 169) also notes that in English and certain other languages, a comma in the written language and parenthetical intonation in the spoken language are used to separate the relative clause from its head in non-restrictive constructions. In standard English, there are also constraints on the type of relative marker used: *wh*-markers, *that*, and the zero variant are employed in restrictive relative clauses, whereas only *wh*-forms are conventionally encountered in non-restrictive constructions. However, this is not categorical in the preadolescent data I examined, as illustrated in the following example from the corpus where *that* occurs in a non-restrictive relative clause:

(9) I've grown up with Aaron, [*that's* his birthday today]



According to Denison (1998: 279), *that* was common in non-restrictive relatives in earlier varieties of English (see also Strang 1970:142). Such usages in contemporary spoken dialects may therefore be synchronic remnants that can be traced to previous patterns in the history of English (see also Beal and Corrigan 2002: 128).

For the purposes of this study, however, I focus on the quantitative analysis of restrictive relative clauses only. Ball (1996: 228-9) stresses that the distinction between restrictive and non-restrictive relative clauses is a crucial methodological point as non-restrictive relative clauses tend to favour *wh*-forms. Thus, the inclusion of non-restrictive relatives in a quantitative analysis of relative markers could potentially skew the data by inflating the number of *wh*-forms in the results while depressing the percentage of *that* and *zero* relatives (see also Tagliamonte *et al.* 2005: 85).

Even after the exclusion of non-restrictive relative clauses from the data base, there still remain, however, certain methodological problems with the delimitation of the variable context. The identification of restrictive relative clauses is complicated by the fact that relative clause constructions in speech and writing do not necessarily share the same typological affiliations (see Pawley and Syder 1983: 564; Miller and Weinert 1998: 372). While loose paratactic structures are frequently encountered in informal spoken language, formal written language characteristically draws on tight hypotactic structures where embeddedness is signalled by overt connectives. In speech there may be few clear-cut morphosyntactic indications that one clause is syntactically incorporated into another. Consider the following example from the preadolescent corpus:

- (10) there's this comedy programme well it's these funny people [*they're*  
on it] [10M10/11]

I have analysed the construction in square brackets as a restrictive relative clause on the basis that the matrix and the embedded clause were bounded by the same intonation contour.<sup>13</sup> Peitsara (2002:168) also notes the existence of similar constructions in the dialect of Suffolk where a personal pronoun may introduce a relative clause (see also Trudgill 1974: 41; and Ihalainen 1980: 190-191). To make my analysis comparable to other quantitative investigations of relative markers in restrictive relative clauses, I decided to exclude constructions such as (10), which were relatively few in number, from this study. Furthermore, to facilitate comparison with other studies, unfinished relative clauses as well as adverbial relative constructions (e.g. where the relative marker is *when*, *why*, *where*) were excluded from the data base (see Tottie and Harvie 2000: 206, and Tagliamonte 2002n: 152).<sup>14</sup> This left 183 restrictive relative clauses for quantitative analysis.

#### 4.4.2 Selection of factor groups

Previous research has illuminated a number of internal and external factors which constrain the distribution of relativizers within the relative marker system

---

<sup>13</sup> In cases where overt morphosyntactic markers are absent, hearers probably deduce the embeddedness of the relative clause from the sequential order of constituents aided by prosodic features such as stress and intonation (Romaine 1982: 78; 1992: 304).

<sup>14</sup> The omission of adverbial relatives from the study was by no means a straightforward matter. In certain cases, some speakers used the relative marker *where* along with a resumptive pronoun in a non-locative relativizing function; for example:

you know one of them bouncy things [ *where* you hang *it* up and you like sit in *it* ] (26F7/8)  
 Diessel and Tomasello (2005: 895) report that German children in their study used *wo* 'where' as a relativizer as 'a strategy to simplify the grammatical structure of [oblique] relatives without changing their meaning.' Guasti and Cardinaletti (2003: 63) also note that indirect object relatives are often introduced by the relative marker *où* 'where' in the speech of French children. Similar usage is also occasionally encountered in adult English speech (see e.g. Mufwene 1986: 21, and Romaine 1984c: 467 for examples from American English). Aitchison (1992:309) suggests that the use of a word meaning *where* to form relative clauses may be a universal tendency. The use of relativizers deriving from the locative adverb *where* is also widely attested in English-based creoles (see Romaine 1992: 300); and, in other Germanic languages such as Swiss German, *wo* 'where' is used to introduce all kinds of relatives (e.g. *dr Bueb wo isst* [lit. 'the boy where eats'] 'the boy who eats' (cited in Guasti and Cardinaletti 2003: 73).

(see e.g. Sigley 1997b: 214). Ball (1996: 230-1) notes that the type of antecedent and the syntactic function of the relative marker in the relative clause are major determinants conditioning the choice of relative marker (e.g. in standard British English, *who* is used with animate antecedents, and the zero variant is used in non-subject relatives). The grammatical function of the relative marker as well as characteristics of the antecedent head were accordingly factored into the coding protocol. In terms of antecedent type, separate factor groups were established for the animacy and definiteness of the antecedent. Within the animacy factor group, antecedents were coded as to whether they were human, non-human animate, or inanimate. With regard to definiteness, each token was coded as either a pronoun (11), an indefinite NP (12), or a definite NP (13):

(11) it wasn't me [who had the fight] [26F7/8]

(12) there was a ghost dragon [that went thro... through him ] [29F7/8]

(13) the person [who he made ] threw him into the locker [2M10/11 ]

According to Tottie and Rey (1997: 233), definite NP heads have been considered to be a factor favouring zero relativization, although in a quantitative analysis of relativization strategies in Early African American Vernacular English, Tottie and Rey (1997:233) found that the zero relative marker more frequently occurred with an indefinite NP head or a pronoun. In terms of definiteness, different varieties may therefore have different contextual constraints operating on the distribution of zero relatives.

Another factor group was established to take account of the possible effects of the structure of the matrix clause on the selection of relative markers. Ball (1996: 257) gives a useful overview of construction types which typically favour the occurrence of zero subject relatives. These constructions include existentials; possessives with

*have/get*; and *it/that* clefts.<sup>15</sup> There are several examples of existentials and clefts in the preadolescent corpus which co-occur with zero subject relatives:

### EXISTENTIAL

(14) there's like other animals [Ø lives there ], a duck and  
all that [10F10/11]

(15) there's this boy [Ø had this bracelet on] [ 13M10/11]

### CLEFT SENTENCE

(16) it was just one arm [Ø was really damaged ] [13M10/11]

(17) cos it was his girlfriend [Ø went to go and find him ] [ 19F10/11 ]

These construction types were factored into the analysis to determine whether there was a significant correlation between matrix clause type and the choice of relative marker.

Another important linguistic factor selected for coding was the length of the relative clause. Recent research (see Sigley 1997a: 214; Tagliamonte *et al.* 2005: 97) has highlighted the correlation between clause length and the use of zero markers, with longer relative clauses being less favourable to the use of zero markers, possibly

---

<sup>15</sup> It has been pointed out to me that there may be an interaction between pronominal status and construction type (see (11) above). Interaction effects potentially have important consequences for multivariate analysis (see Tagliamonte 2006a). As pointed out in 2.3.3, although interactions between social factors are widely acknowledged in the variationist literature (see e.g. Milroy 1992: 78), significant interaction effects between linguistic factors are common too (Sigley 2003: 230). Tagliamonte (2006a: 139) points out that while such interactions are 'very real aspects of naturally occurring language [...] sometimes the interactions are so severe that they obscure underlying effects on the variable.' As explained in 2.3.3, there are a number of systematic procedures that can be adopted to facilitate the identification of interactions in the data. In the case of potential syntactic dependencies in the data, one way to pinpoint interaction effects is to run the variable rule program on those factor groups which are suspected of interacting. I performed such a procedure in order to investigate the potential existence of a strong interaction between pronominal status and construction type. In order for the variable rule program to run, existentials, clefts and possessives had to be collapsed and treated as an aggregated factor (see 4.5.1) in order to permit comparison with other construction types. Several different runs of the data revealed that there was only a mild interaction between pronominal status and construction type in relative clauses containing non-standard *what*. However, caution needs to be exercised in interpreting this result as the number of *what* tokens is low. Furthermore, statistical fluctuations are generally greater in smaller data sets than in larger ones (Poplack and Tagliamonte 2001: 93). Nevertheless, in future research, it will be important to investigate the existence of such potential syntactic dependencies in larger data sets.

owing to processing constraints. To test for the possible effects of clause length on relative marker selection, I coded for the number of words in each relative clause: clauses which contained five or fewer words were coded as short, whereas clauses which had more than five words were coded as long.<sup>16</sup>

Previous research has also highlighted the potential effects of social factors on the choice of relative marker, including the educational background as well as the age and gender of the speaker. For example, Tagliamonte (2002b: 161) found that more educated speakers were more frequent users of *who* than less educated speakers. Based on their analysis of a corpus of Tyneside English collected in 1994, Beal and Corrigan (2002) also report a gender difference related to the choice of relative marker. According to Beal and Corrigan (2002: 128), whereas females in the 1994 Tyneside corpus have a fairly even balance between *wh*, *that*, and zero relatives, the males in the corpus use the zero relative 50% of the time. Similarly, gender proved to be a sensitive variable in Tottie and Rey's (1997) analysis of relative marker usage in Early African American Vernacular English. Tottie and Rey (1997: 242) also found that men had a higher proportion of zero markers than women: 65% compared with 45%.

Turning to age, Romaine (1984a: 48) reports that Scottish preadolescents showed increasingly less reliance on the zero strategy as they got older, with ten-year-old children making greater use of the *wh*-strategy and *that* than six-year-old children. In order to examine possible correlations between social factors and relative marker usage in the preadolescent corpus, I coded each token with regard to the age and gender of the speaker.

---

<sup>16</sup> Sigley (1997a: 216) calculates the mean clause length of restrictive adnominal relative clauses in conversation to be 4.59 words based on his analysis of relativization in the *Wellington Corpus of Spoken New Zealand English*.

## 4.5 DISTRIBUTIONAL ANALYSIS

The data include a number of different overt relative markers as well as zero forms:

- (18) it's about this man [*that* kills a lady ] [19F7/8]
- (19) there was this man [*who* rang up ] [29F7/8]
- (20) it tells you how many ghosts [Ø you have ] [28F7/8]
- (21) cos he's got a bed [*which* you can go under ] [16F10/11]
- (22) it's you the one [*what* wants to fight ] [9M10/11]

Noticeably absent from the entire database is the use of *as* in a relativizing function, which suggests that this variant is no longer a productive feature of the relative marker paradigm used by the youngest generation. This finding concurs with research undertaken in other southern areas of England; for example, van den Eynden Morpeth (2002: 185) claims that relativizer *as* is obsolescent in the greater part of the southwest (see also Herrmann 2005: 36, 40 on the decline of *as* in its capacity as a non-standard relative marker).

There are also no instances in the data of pied-piping (i.e. where the preposition has been optionally fronted along with the *wh*-NP to the beginning of the relative clause as in 'the house in which I used to live'). According to McDaniel *et al.* (1998: 309), preposition pied-piping is not a natural option in English, but is a prescriptive artefact acquired during schooling.<sup>17</sup>

There are also no instances of *wh*-markers in the corpus which have overt case (e.g. *whose*). There is only one instance of a genitive relative in the entire preadolescent corpus (see example (9) above), which occurs in a non-restrictive relative clause. In this instance, however, the relative marker used is not *whose*, but consists of the relative marker *that* used in conjunction with a resumptive pronoun to

---

<sup>17</sup> Guasti and Cardinaletti (2003: 48) make a similar remark in relation to French and Italian children's acquisition of relativization strategies, noting that 'pied-piping in relatives emerges during school years and may be the result of explicit teaching.'

signal the case of the relativized NP. Weinert (2004: 10) also mentions the use of periphrastic constructions to replace genitive *whose* in spoken English (e.g. *the students that their work was excellent*), and Herrmann (2005: 61) claims that dialect speakers often make use of alternative constructions (e.g. left dislocations as well as *and* co-ordinations) in the place of the standard relative pronoun *whose*.<sup>18</sup> The strategy employed in (9) is also consonant with the predictions made by Keenan and Comrie (1977) based on the Accessibility Hierarchy according to which there is an order of difficulty which correlates with the syntactic position of the relativized NP (i.e. subject is the easiest position to relativize, but other positions lower down the hierarchy such as the genitive position are harder). According to Keenan and Comrie (1977), resumptive pronouns are at least optional when the relativized NP is lower down on the Accessibility Hierarchy. It is also important to stress that constructions such as (9) above in the preadolescent corpus do not simply reflect developmental immaturity in linguistic competence *per se*, but are options that are additionally exploited by adults (see Suñer 1998: 341), as well as being attested in earlier stages of the English language (see e.g. Fischer 1992: 308-9).

---

<sup>18</sup> A more holistic variationist approach to the study of relativization might arguably encompass alternative relativization strategies, including the use of relative markers in conjunction with resumptive or shadow pronouns, as well as constructions such as the one cited in (10) in section 4.4.1, where the use of a personal pronoun may function like a relative marker. Romaine (1988: 231; 1992: 243-4) notes that in Hawai'i Creole English and Tok Pisin, relative clauses often make use of no special marker of subordination, but instead use a personal pronoun in a relativizing function. However, as Romaine (1988: 229) points out, it is difficult to provide a 'universal set of defining properties shared by all constructions which syntacticians have defined under the heading of *relative clause*.' Furthermore, as in the case of example (10), it is not always easy to distinguish simple conjunction from subordination with embedding (Romaine 1992: 290). Nevertheless, many previous variationist studies have adopted an orthodox approach in their analyses of relativization strategies, which have been heavily influenced by formal structural accounts of relativization. Future variationist research dealing with child language needs to pay closer attention to alternative relativization strategies which may embody important pragmatic functions, and represent important intermediate stages in the child's acquisition of fully syntacticised relative clause constructions (see Romaine 1984a, 1988, 1992).

Turning now to a proportional analysis of the relative marker variants in the corpus, Table 4.1 below shows the overall distribution of relative markers in restrictive relative clauses produced by the preadolescents.

TABLE 4.1

OVERALL DISTRIBUTION OF RELATIVE MARKERS

Relative marker	Number	Percentage
<i>who</i>	39	21
<i>what</i>	19	10
<i>which</i>	2	1
<i>that</i>	74	40
<i>zero</i>	49	27
TOTAL	183	

The first thing to note about the distribution of the data is that although *wh*-markers collectively account for a substantial part of the preadolescents' relative paradigm, the predominant marker is *that*, which accounts for 40% of all forms. The data displayed offer corroborative evidence that relative marker *that* is the preferred form among the youngest generation of speakers (see Tagliamonte 2002b: 162). Nevertheless, as we shall see below, the environmental constraints on relative marker *that* can reflect subtle inter-variety differences once the configuration of internal conditioning factors operating on its selection is examined in more detail.

Although the *wh*-markers do not occupy the pre-eminent position in the relative marker paradigm, they are the second most popular strategy when considered as an aggregate (32%), and are slightly more frequent than the zero strategy (27%). The aggregated frequency of the *wh*-forms in the preadolescent data differs markedly from comparable results obtained in other geographical locations. Romaine (1984a:



47) found that Edinburgh schoolchildren tended to use *that* (43%) and zero (42%) in approximately equal proportions, whereas *wh*-forms only accounted for 16% of the relative markers used. The contrast between the frequency distributions displayed in Table 4.1 above and Romaine's (1984a) data support Beal and Corrigan's (2002: 132) view that northern British dialects lag behind southern ones in terms of their adoption of *wh*-forms.

On the other hand, however, the frequency of *wh*-forms in the preadolescent data is lower than the frequency of these forms in standard British English (Quirk 1957). The relative marker *which* is only negligibly instantiated in the data, which supports Ball's (1996: 251) claim that this form has found little place in the restrictive relative marker paradigms of non-standard varieties of English.

Other differences relate to the distribution and frequency of *who*. For example, in an analysis of relative marker usage in educated spoken language, Quirk (1957: 97-8) found that *who* accounted for 91% of all subject relatives (see also Tottie 1997). By contrast, Table 4.2 below shows that *who* accounts for a more marginal proportion of subject relatives in the preadolescent corpus. Furthermore, *who* is almost entirely confined to subject position in the data. The most syntactically versatile *wh*-marker is the non-standard form *what*, which occurs in all three grammatical positions.

TABLE 4.2

DISTRIBUTION OF RELATIVE MARKERS BY GRAMMATICAL CATEGORY

	<i>who</i>	<i>what</i>	<i>which</i>	<i>that</i>	<i>zero</i>	TOTAL
Subject	38 (34%)	9 (8%)	0 (-)	55 (50%)	9 (8%)	111
Object	1 (2%)	8 (13%)	1 (2%)	18 (30%)	33 (54%)	61
Oblique	0 (-)	2 (18%)	1 (9%)	1 (9%)	7 (64%)	11
TOTAL	39	19	2	74	49	183

In terms of distribution by syntactic function, we can see that whereas *that* is the predominant form in subject position, zero prevails in both object and oblique positions, although admittedly there is a small number of tokens in the latter category. Sigley (1997b: 230) also notes the use of zero as the default restrictive object relativizer in New Zealand English. We can also observe that, unlike in standard English, the zero variant is used in subject position, although it occurs relatively infrequently. This finding is in line with historical evidence indicating that the frequency of zero subject relativization began to decline in the Early Modern English period in contrast with its robustness in Middle English (see Tottie and Harvie 2000: 202).

Further differences become apparent when the animacy of the antecedent co-referential with the relativized NP is analysed. The data in Table 4.3 below show that *who* is the dominant relativizer with human antecedents. On the other hand, *that* is used more frequently with animate and inanimate (i.e. non-personal) antecedents. This finding contrasts with Tagliamonte's (2002b: 155) results, where *that* in subject position is more frequently co-referential with human antecedents than *who*.

TABLE 4.3

DISTRIBUTION OF RELATIVE MARKERS BY ANIMACY OF THE ANTECEDENT NP

	<i>who</i>	<i>what</i>	<i>which</i>	<i>that</i>	<i>zero</i>	TOTAL
Human	39 (53%)	2 (3%)	0 (-)	22 (30%)	10 (14%)	73
Non-human animate	0 (-)	5 (28%)	0 (-)	10 (56%)	3 (17%)	18
Inanimate	0 (-)	12 (13%)	2 (2%)	42 (46%)	36 (39%)	92
TOTAL	39	19	2	74	49	183

In the preadolescent corpus, the humanness of the antecedent NP appears to be a decisive factor affecting the choice of *who* in preference to other variants. The strong association between *who* and human subjects reflects its historical conditioning as the relative marker which came to be used with animate and personal antecedents by the end of the Early Modern English period (Romaine 1982: 69). By contrast, as far as animacy effects are concerned, non-standard *what* occurs more frequently with non-human animate and inanimate antecedents than with human antecedents (see also Herrmann 2005: 45). Similar effects of the animacy of the antecedent on the choice of *wh*-relative markers are reported in other dialect studies. Cheshire (1982a: 72) reports that in the vernacular of Reading teenagers, *what* occurs more often with a non-personal antecedent than with a personal one. Reading English showed high levels of *who* usage with personal antecedents too.

Interesting patterns of distribution are also evident when the definiteness of the antecedent NP is taken into account. The distributional results displayed in Table 4.4 below exhibit a partial overlap with the results obtained by Tottie and Rey (1997: 233).

Recall that Tottie and Rey (1997: 233) found that the zero relative marker in Early African American Vernacular English more frequently occurred with a pronoun or an indefinite NP. The frequencies displayed below show that there is a similar tendency for the zero variant to be selected when the antecedent is a pronoun, but in contrast with Tottie and Rey's (1997) results, definite antecedents rather than indefinite ones preferentially occur with zero.

TABLE 4.4

## DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO THE DEFINITENESS OF THE ANTECEDENT NP

	<i>who</i>	<i>what</i>	<i>which</i>	<i>that</i>	<i>zero</i>	TOTAL
Pronoun	2 (10%)	4 (20%)	0 (-)	3 (15%)	11 (55%)	20
Definite NP	13 (20%)	9 (14%)	1 (2%)	21 (32%)	22 (33%)	66
Indefinite NP	24 (25%)	6 (6%)	1 (1%)	50 (52%)	16 (16%)	97
TOTAL	39	19	2	74	49	183

Table 4.5 below summarises the distributional results for the effect of matrix clause structure on the choice of relative marker.

The zero relative marker occurs comparatively more frequently with cleft sentences than with other matrix clause types. The proportion of existentials favouring *that* is also relatively high, although clefts and possessives show lower proportions of *that* than other matrix clause construction types. There is less fluctuation in the relative frequencies of *who* in these sentence types, although it is notable that existentials, clefts and possessives exhibit a tendency to favour *who* more than other clause types.

Another facet of variation in the data which is notable concerns the aggregated frequency of existential, cleft and possessive constructions in the restrictive relative clause data base: cumulatively, these constructions account for over one third of matrix clause types. Moreover, the proportion of existential constructions is high in comparison with clefts and possessives.

TABLE 4.5

## DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO THE STRUCTURE OF THE MATRIX CLAUSE

	<i>who</i>	<i>what</i>	<i>which</i>	<i>that</i>	<i>zero</i>	TOTAL
Existential	9 (26%)	3 (9%)	0 (-)	16 (47%)	6 (18%)	34
Cleft	6 (33%)	1 (6%)	0 (-)	4 (22%)	7 (39%)	18
Possessive	3 (25%)	2 (17%)	2 (17%)	4 (33%)	1 (8%)	12
Other	21 (18%)	13 (11%)	0 (-)	50 (42%)	35 (29%)	119
TOTAL	39	19	2	74	49	183

From a developmental point of view, it is interesting to note that, according to Diessel and Tomasello (2000:136), relative clauses in young English-speaking children frequently modify the predicate nominal of a presentational copular clause. In order to explore the role and frequency of these constructions in more detail, it is necessary to move beyond strictly syntactic considerations and take into account the pragmatic properties of relative clause constructions (see Ball 1996:253), particularly their role in the management of information flow in interactive discourse. Thus, in examples (14) and (15) above (repeated below in (23) and (24) for convenience), the matrix clause of the construction is propositionally empty, and new information is asserted in the restrictive relative clause.

(23) there's like other animals [ $\emptyset$  lives there ], a duck and all that  
[10F10/11]

(24) there's this boy [ $\emptyset$  had this bracelet on] [ 13M10/11]

According to Dasinger and Toupin (1994: 464), existential constructions and *it* clefts are prime examples of syntactic constructions used to introduce new referents in non-subject position. The relative clause allows the insertion of information about a newly

introduced referent in the matrix clause which is established as a topic of importance. As topic-elaborating constructions (see also Henry 1995), relative clauses embedded in existential clauses and cleft constructions can effectively contextualize new referents in a minimal sentential processing unit (see Cheshire 1999b: 71). As Fox and Thompson (1990: 315) observe, communicative factors can determine a preference for particular kinds of relative constructions.

Surface level processing factors are also implicated in the choice of relative markers. The data displayed in Table 4.6 below show that there is an evident tendency for zero markers to be avoided in relative clauses which consist of more than 5 words.

TABLE 4.6

DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO THE LENGTH OF THE RELATIVE CLAUSE

	<i>who</i>	<i>what</i>	<i>which</i>	<i>that</i>	<i>zero</i>	TOTAL
Less than 5 words	24 (17%)	15 (11%)	1 (1%)	53 (38%)	48 (34%)	141
More than 5 words	15 (36%)	4 (10%)	1 (2%)	21 (50%)	1 (2%)	42
TOTAL	39	19	2	74	49	183

This distribution ties in with Cofer's (1975:31) claims that an overt relative marker is employed when the relative clause is long.<sup>19</sup>

Turning now to extra-linguistic factors, Tables 4.7 and 4.8 below show the distribution of relative markers according to the social parameters of gender and age.

<sup>19</sup> Additional factors discussed in the acquisitional literature which relate to the complexity of relative clauses are embeddedness and focus (De Villiers *et al.* 1979). Embeddedness relates to the function of the antecedent NP in the matrix clause which is qualified by the relative clause, whereas focus relates to the function of the relativized element in the relative clause. Different combinations of embeddedness and focus give rise to four major types of relative clause: SS, OO, OS, and SO, where S stands for subject, and O stands for the object of a verb or a preposition (Romaine 1992: 289). In both age cohorts in the preadolescent corpus (i.e. 7-8 year olds and 10-11 year olds), the ordering of these relative clause types is ranked as follows: OS>OO>SS>SO. Romaine (1992: 291) explains the predominance of the OS and OO types in child language on the grounds that they are parallel in form to simple conjunction, and are therefore easier for children to produce than other types of relative clause (see, however, Diessel and Tomasello 2005: 883 on SS relatives in early child language.)

TABLE 4.7

## DISTRIBUTION OF RELATIVE MARKERS BY GENDER

	<i>who</i>	<i>what</i>	<i>which</i>	<i>that</i>	<i>zero</i>	TOTAL
Male	17 (18%)	11 (11%)	0 (-)	43 (44%)	26 (27%)	97
Female	22 (26%)	8 (9%)	2 (2%)	31 (36%)	23 (27%)	86
TOTAL	39	19	2	74	49	183

TABLE 4.8

## DISTRIBUTION OF RELATIVE MARKERS BY AGE

	<i>who</i>	<i>what</i>	<i>which</i>	<i>that</i>	<i>zero</i>	TOTAL
7-8 year olds	18 (22%)	8 (10%)	1 (1%)	36 (44%)	19 (23%)	82
10-11 year olds	21 (21%)	11 (11%)	1 (1%)	38 (38%)	30 (30%)	101
TOTAL	39	19	2	74	49	183

The results displayed in the Table 4.7 above show that there is no difference in the proportion of zero relatives produced by males and females. These results differ from findings in other communities (e.g. Tottie and Rey 1997), where males were found to use a higher proportion of zero relatives than females. Although the preadolescent females are ahead in their use of *who* vis-à-vis the males, the difference between the two sexes is not considerable. Similarly, there is not much difference between males and females in their frequency of use of non-standard *what*.<sup>20</sup>

In terms of speaker age, there is a slight increase in the rate of relativization as speakers mature, with 7-8 year olds using 2.4 relative clauses per 1000 words, and 10-

<sup>20</sup> There is much intra-speaker variation in the use of relativization strategies, with individuals using *wh*-relativizers, *that* and *zero*. Thus, the percentages presented in Tables 4.7 and 4.8 reflect variation which exists within individual speaker usage as well as variation which is shared more broadly across speaker groups. Non-standard relativization strategies, including the use of zero subject relatives and non-standard *what*, are also endemic to speaker groups rather than reflecting the idiosyncratic behaviour of a restricted number of individuals.

11 year olds using 2.8 relative clauses per 1000 words. However, as the data in Table 4.8 show, there is no increase with age in the proportion of *wh*-markers used. Increased exposure to literacy and formal education might lead us to hypothesize an increment in the use of standard *wh*-forms, and, possibly, a corresponding decrease in the use of non-standard *what*, but this hypothesis is not borne out by the data tabulated above.

#### 4.5.1 *Multivariate Analysis*

The data were also configured for multivariate analysis with the aim of determining which factors favour the occurrence of different relative markers when all the factors are considered simultaneously.

Preparation of the data for multivariate analysis necessitated some re-organization: firstly, because the relative marker *who* was almost exclusively used in subject position, and occurred categorically with human antecedents, it was necessary to exclude this marker to make a multivariate analysis of the data viable. *Which* was also omitted from the multivariate analysis owing to its negligible occurrence in the preadolescent corpus. These exclusions resulted in the multivariate analysis being focused on variation between *what*, *that* and zero.<sup>21</sup>

Finally, owing to a small number of tokens, existentials, clefts and possessives were collapsed into one factor to enable comparison with the choice of relative marker in other construction types.

---

<sup>21</sup> This analysis differs, then, from the multivariate analysis of the data presented in Levey (2006: 63), in which all the *wh*-forms were grouped together. On reflection, the aggregation of *wh*-markers which are individually sensitive to different internal constraints now seems inappropriate, and is not pursued here.



TABLE 4.9

THREE INDEPENDENT MULTIVARIATE ANALYSES OF THE  
CONTRIBUTION OF FACTORS TO THE PROBABILITY OF *WHAT*, *THAT*, AND  
*ZERO* IN THE PREADOLESCENT CORPUS

INPUT	0.11		0.52		0.24		Total N
	WHAT		THAT		ZERO		
	FW	%	FW	%	FW	%	
<b>FACTORS CONSIDERED</b>							
<b>Grammatical Function of Relative Marker</b>							
Subject	0.51	12	0.71	75	0.26	12	73
Object	0.48	14	0.33	31	0.73	56	59
Oblique	0.57	20	0.07	10	0.85	70	10
Range			64		59		
<b>Animacy of antecedent NP</b>							
Human	0.32	6	0.46	65	0.66	29	34
Non-human animate	0.79	28	0.26	56	0.41	17	18
Inanimate	0.51	13	0.57	47	0.46	40	90
<b>Grammatical category of antecedent NP</b>							
Pronoun	0.74	22	0.16	17	0.72	61	18
Definite NP	0.65	17	0.36	40	0.59	42	52
Indefinite NP	0.33	8	0.70	69	0.38	22	72
Range			54				
<b>Sentence Structure</b>							
Existentials, clefts and possessives	0.59	14	0.32	55	0.67	32	44
Other	0.46	13	0.59	51	0.42	36	98
Range			27				
<b>Length of relative clause</b>							
Fewer than 5 words	0.49	13	0.44	46	0.62	41	116
More than 5 words	0.57	15	0.74	81	0.11	4	26
Range			30		51		
<b>Speaker Gender</b>							
Male	0.50	14	0.44	54	0.54	33	80
Female	0.50	13	0.57	50	0.45	37	62
<b>Speaker Age</b>							
7-8 year olds	0.49	13	0.57	57	0.45	30	63
10-11 year olds	0.51	14	0.44	48	0.54	38	79

With the exception of the animacy of the antecedent NP, the multivariate analyses displayed in Table 4.9 above reveal that each of the other internal linguistic factors exhibits levels of statistical significance with regard to the choice of either *that* or the zero relative marker. From a methodological perspective, this underscores the importance of carefully selecting and coding for a variety of factors in order to carry out a comprehensive investigation of the range of contextual constraints on the selection of relative markers (see also Ball 1996: 227).

For four of the five internal factors, *that* attains levels of statistical significance. In terms of the comparative magnitude of statistically significant effects, it can be seen that the grammatical function of the relative marker exerts the most potent influence on the selection of *that*, with subject position showing a favouring effect, and the other two grammatical positions disfavouring *that*. The second strongest effect is exerted by the definiteness of the antecedent, with only indefinite heads favouring the choice of *that* as a relative marker (see also Huddleston 1971).

Although the range is smaller than in the case of the other two internal factors just mentioned, we can also see that *that* is favoured in clauses that are longer than five words, which corroborates the influence of processing constraints on the choice of an overt relative marker.

A statistically significant effect is also apparent with regard to sentence structure: the factor weights indicate that relative marker *that* is disfavoured in existential, clefts and possessives (when treated as an aggregate group), whereas it is moderately favoured in other sentence types.

Although none of the factor groups for non-standard *what* is selected as significant, possibly owing to the small number of tokens in the corpus, it is interesting that the constraint ranking for *what* according to the grammatical category

of the antecedent NP is not in line with the historical association of this variant with indefinite antecedents (see e.g. Mustanoja 1960: 194). In fact, in the preadolescent data, *what* is disfavoured with indefinite NPs and favoured with definite NPs and pronouns. However, this interpretation of the data must be guarded owing to the limited overall occurrence of this non-standard variant in the corpus. Nevertheless, the issue of whether similar or different constraint rankings obtain for non-standard *what* in other varieties of English where it is attested is worth pursuing.

Turning to the zero variant, there are only two factor groups which attain levels of statistical significance. The most powerful effect is exerted by the grammatical function of the relative marker, with subject position strongly disavouring the zero variant in marked contrast with the other grammatical positions which exhibit favouring effects.<sup>22</sup> The favouring effect of sentence structure on the zero variant is not statistically significant in these data, although the direction of the effect is consistent with that of other studies (see e.g. Tagliamonte *et al.* 2005). This is possibly a dimension of variation that warrants further cross-dialectal investigation. On the other hand, clause length contributes a statistically significant effect on the use of the zero variant: relative clauses containing fewer than five words favour the choice of this variant, whereas longer clauses have a strongly inhibiting effect. The strength of this effect is also highlighted in Tagliamonte *et al.* (2005), who found clause length correlated highly with the use of the zero variant in three vernacular varieties of English in northern Britain. Whether or not this is a pan-community effect requires further empirical investigation.

---

<sup>22</sup> These findings are not altogether congruent with the constraints associated with the Accessibility Hierarchy. They are at variance with the claims made by Keenan (1985: 154) that 'NP<sub>rel</sub> is most likely to be gapped if it is the subject of S<sub>rel</sub>, next most likely if it is the direct object, etc.'

Even though neither of the extralinguistic factors examined exhibits levels of statistical significance, the factor weights associated with age and gender raise a number of important points. Firstly, none of the relativization strategies is sensitive to gender. According to Tottie and Rey (1997: 245), relativization is an unobtrusive, covert variable which is not readily available for social evaluation (see also Cheshire 1999b: 61). Moreover, as Rickford *et al.* (1995: 106) point out, syntactic variables occur comparatively less often than phonological variables do in spontaneous conversation, and are therefore less likely to be conscripted by members of a speech community for creating social indices.

An important finding emerging from the factor weights associated with age is that, in contrast with Romaine's (1984a) data from Edinburgh preadolescents, there is no discernible age-related tendency for the London preadolescents to rely less on the zero strategy and show a correspondingly greater dependence on the other two overt relativization strategies. In fact, there is relatively little fluctuation in the factor weights for all three variants across the two age groups, suggesting that internal linguistic constraints on variation in relative marker usage are in place at a relatively early age, and that variation in this grammatical subsystem is stable, at least as far as a comparison between the two preadolescent age cohorts reveals.

#### 4.6 SUMMARY

The results of this small-scale study of relative marker usage in the preadolescent corpus have shown that while *that* is proportionally the pre-eminent variant in the corpus, in line with the trends reported for other vernacular varieties of English (see e.g. Romaine 1982; Tagliamonte *et al.* 2005), the *wh*-forms, when viewed as an aggregate, are far from being numerically insignificant in the corpus.

From a proportional perspective, then, the preadolescent vernacular data examined here do not support Romaine's (1982: 212) claim that the *wh*-relativization strategy has '...not really affected the spoken language.' Nor are the data in line with the apparent paucity of *wh*-forms reported for the recently compiled *British National Corpus* (Tottie 1997: 471).

The predominant *wh*-form, *who*, was found to be almost exclusively restricted to subject position, where it is the preferred variant with human antecedents. It is possible that *who* maintains a foothold in the preadolescents' relative marker paradigm because it is semantically motivated: it is the only overt relative marker which is consistently used with human antecedents. The sensitivity of *who* to the humanness of the antecedent can be traced to earlier constraints attested in the history of English, most notably the diachronic tendency to distinguish referents high on the animacy hierarchy from those lower down on the scale (Nevalainen and Raumolin-Brunberg 2002: 110). The social embeddedness of this change also has a bearing on the interpretation of the results for the preadolescents.

Nevalainen and Raumolin-Brunberg (2002: 119-200) show that the rise of *who* in Early Modern English was subject to regional and social variation, with members of Court using *who* at higher rates than Londoners of lower social rank, who showed a preference for *that*. When the rise of *who* is viewed from a diachronic perspective, we can see that what originated as a learned change from above (see Nevalainen and Raumolin-Brunberg 2002: 120) has percolated down the social hierarchy and, judging from the distributional data discussed here, has become embedded as a productive variant in the vernacular relative marker paradigm of the youngest generation of Londoners.

Given that *wh*-markers originally infiltrated the relativization system through more formal registers (Romaine 1982: 234), and that high rates of *wh*-forms are typically reported in studies based on more standard varieties of English (see Tagliamonte *et al.* 2005: 93), it might appear that in comparison with other studies of relative marker usage in vernacular varieties of English, the relatively higher frequency of *wh*-forms in the preadolescent corpus is the product of formality effects, or sensitivity to exogenous normative influences. However, this seems extremely unlikely, not only in the light of the generally informal nature of the conversational data recorded, but also given the frequency of other non-standard forms in the same corpus detailed in previous and ensuing chapters in this study. Furthermore, although it is a minority variant in the children's relative marker paradigm, the use of non-standard *what* in a relativizing function provides additional evidence that the preadolescents are not simply orienting towards any external prestige norm in their use of *wh*-variants.

With regard to the other variants used, the results obtained for the zero relative marker raise several points of interest, particularly in connection with the effects of grammatical function of the relative marker and clause length. Although zero subject relatives were shown to be in evidence in the data, and were traceable to antecedent forms that are long-standing in the history of English, subject position was found to be an environment that strongly dispreferred the use of the zero variant. On the other hand, object and oblique positions were favourable to the use of zero relatives: these results indicate that zero marking in this vernacular is not a unitary strategy, with zero relativization in subject position behaving quite differently from zero marking in other syntactic positions.

In spite of the fact that *that* was found to be the numerically dominant variant, a comparison of the conditioning factors affecting its selection in the preadolescent data with the environmental constraints on its use in other vernaculars reveals that, unlike northern varieties of English (see Tagliamonte *et al.* 2005), it is used at substantially lower rates with human antecedents than *who*. Although it is not selected as significant, the constraint hierarchy for the animacy of the antecedent NP shows that *that* is used more often to relativize antecedents that are inanimate rather than non-human animate or human. This effect is in keeping with the historical assignment of *that* to nonpersonal antecedents, a change which was underway by the second half of the seventeenth century (Ball 1996: 246)

Although the results presented here are based on a relatively small number of tokens used by a specific age cohort, the findings which emerge from the quantitative investigation of relative marker usage by preadolescents can be used as a foundation for pursuing additional questions relating to relative marker usage in southeastern dialects, and the evolution of vernacular norms in this geographical region. In particular, there is a need for substantial inter-generational vernacular data in order to conduct comparisons of variation in relative marker usage across different age cohorts in London and the southeast with a view to establishing whether *wh*-variants are increasing or receding in apparent time. Furthermore, such research would be instrumental in illuminating the vitality of non-standard *what* in different age groups, as well as exploring the contextual effects which constrain its variable use.

## CHAPTER 5

# TENSE VARIATION IN PREADOLESCENT NARRATIVES

### 5.1 INTRODUCTION

This chapter aims to examine tense-switching phenomena in narratives of personal experience collected from the preadolescents I recorded.

Tense alternation has been the subject of a number of discourse-analytical studies examining spoken and literary genres (e.g. Wolfson 1979, 1982, and Schiffrin 1981 on contemporary spoken American English; Silva-Corvalán 1983 on oral Spanish narratives; Fleischman 1990 on tense variation in medieval literature; and Leith 1995 on tense variation in a Scottish folktale). The focus of much of this research (see especially Schiffrin 1981; Wolfson 1982) has been on the discourse functions of the socially meaningful alternation between the simple past and the CHP in specific types of narrative to refer to events which take place before the time of speaking.

The lion's share of this research has focused almost exclusively on American English (although see Romaine 1985 and Leith 1995 for discussion of tense-switching in Scots; Engel and Ritz 2000 on Australian English; and Cox 2005 on New Zealand English), and less research has been undertaken from a variationist perspective on tense-switching in narrative discourse in other varieties of English.

This chapter specifically targets tense variation in the domain of past temporal reference in complicating action clauses in order to explore whether the patterned organisation of tense alternation highlighted in previous research on narratives of



personal experience (e.g. Schiffrin 1981) is apparent in the narratives of the preadolescent speakers examined below. As Schiffrin (1981: 45) points out, the regular internal structure of narrative makes it readily accessible to systematic and controlled variationist analysis.

Although this chapter attempts to build on previous foundational research on tense variation, its approach differs from earlier sociolinguistic studies of this topic in a number of aspects. Firstly, my analysis is based on the speech of preadolescents, rather than on the speech of adult age groups, which have been the primary focus of previous studies of tense variation. Secondly, given that the findings I discuss below raise a number of questions pertaining to the long-standing variation between the simple past and the present perfect in English (see Tagliamonte 2000a), as well as cross-dialectal differences in the grammaticalisation of the present perfect, I embed my analysis within a synchronic and diachronic framework in order to investigate current trends in the evolution of the present perfect in contemporary vernaculars. Bearing in mind that the results discussed here are based on the usage of children and not adults, I attempt to contextualize some of my findings by drawing on evidence from other varieties of English used by older speakers.

In conducting a quantitative analysis of tense variation in complicating action clauses in preadolescent narratives, I seek to address the following questions arising from previous research into tense-switching phenomena:

1. Which tense forms alternate in this variety of English?
2. How is the distribution of these forms constrained by internal and external factors?
3. What are the discourse-pragmatic motivations underlying tense-switching between the simple past, the CHP, and the present perfect?

4. How does tense variation in the preadolescent narratives analysed here compare with tense variation in the narratives of older speakers?

In attempting to provide some answers to the above questions, I draw heavily on the framework for narrative analysis developed by Labov and Waletzky (1967) and Labov (1972). Labov (1972: 359-60) has defined a narrative as a method of 'recapitulating past experience by matching a verbal sequence of clauses to the sequence of events which (it is inferred) actually occurred.' The main criterion which a text needs to satisfy in order to qualify as a narrative is that of sequentiality, with a minimal narrative being defined as a sequence of two clauses which are temporally ordered (Labov 1972: 360).<sup>1</sup>

## 5.2 A NOTE ON TENSE-SWITCHING AS AN HISTORICAL PHENOMENON

Fleischman (1990) presents a detailed exposition of tense-switching as an historical phenomenon and argues that it is amply attested in medieval 'performed stories' in which simple past-CHP alternations are an intrinsic feature of an oral narrative style. Similarly, Romaine (1984d: 120) notes that grammarians have discussed uses of the historical present in the earlier stages of many Indo-European languages (e.g. Latin, Greek, Old Norse).

In Old English, Traugott (1992: 182) notes that there are few authentic native uses of the historical present, although it is found in Latin writings translated into Old English. Fischer (1992: 242) claims that it was only in Late Middle English that the historical present began to gain impetus. Although some linguists believe that its use

---

<sup>1</sup> The basic functional components of a narrative described by Labov (1972: 370) are the abstract (what the story is about); the orientation (which generally provides specifications relating to the setting and the situation of activity in the narrative); the complicating action (which encapsulates the sequence of events that form the backbone of the narrative); the evaluation (which indicates the point of the narrative); the resolution (which explains what finally happened); and the coda (which signals that the narrative is finished).

in English reflects the influence of Latin or French (see Fischer 1992: 242 for discussion), Traugott (1972: 142) points out that the historical present could have arisen natively as it is also widely encountered in other Germanic languages. French literary models as well as Scandinavian influence (see e.g. Old Norse sagas in which the historical present was common) could have reinforced the use of the historical present in English as a stylistic device. According to Fischer (1992: 244-5), the fact that many of the verbs that are found in the historical present are inherently imperfective, whereas the past is used for punctual or perfective verbs, suggests that the historical present in Middle English could have possibly served a grammatical function that was later fulfilled by the progressive.

The extent to which switching between the historical present and the simple past embodied identifiable discourse-pragmatic functions in older varieties of English is unclear, and warrants further research. Visser (1963: 2198) refers to the 'frequent and at first unsystematic interchange of the preterite and the historical present' in earlier varieties of English. However, Richardson's (1991:344) analysis of tense-switching in Middle English poetry, drawing on tense usage in *Sir Gawain and the Green Knight*, reveals that 'switches from the preterite to the historical present tense foreground key events, descriptions, and characters'.

Poplack (2006: 458-9) points out that there is early acknowledgment in the English grammatical tradition that the present tense may variably substitute for the simple past. Miège (1688: 70) observes that the present tense was 'sometimes used for the Preter Imperfect. As, *having met with him, he brings him to his House, and gives him very good Intertainment*. There we say *brings* for *brought*, and *gives* for *gave*.'

Tense shifting involving the present perfect is also deeply embedded diachronically, although it is less extensively documented than tense alternations

between the simple past and the historical present. Although the semantic perfect was often rendered in Old English by the simple past (Traugott 1992: 190), there is evidence that the periphrastic construction consisting of *habban* + past participle was reanalysed as a verbal complex as early as the Old English period (Traugott 1992: 193). With regard to historical variation between the simple past and the present perfect, Lawrence (2001: 200) notes that extensive variability between the two forms is evident from the earliest stages of English (see also Strang 1970: 149; Traugott 1992: 190). Denison (1993: 352) cites examples from Old English where the present perfect appears to be interchangeable with the simple past; for instance :

(1) Annania, deofol bepæhte ðine heortan, and ðu hæfst alogen þam

‘Ananias devil seduced your heart and you have lied-to the

Halgan Gaste

Holy Ghost’

The rise of the periphrastic perfect paved the way for the complex interplay between past temporal reference forms in the Middle English period. Visser (1963: 2192) refers to the ‘unsettled, pell-mell uses of the preterite and the perfect’ in older English, and comments that it is only after the time of Shakespeare that the preterite and the *have* + past participle construction are used as they are in contemporary standard English.<sup>2</sup> Similarly, Mustanoja (1960: 504) claims that in Middle English and Early Modern English, the functional distinction between the simple past and the compound tenses of the past are less clear-cut than in Modern English.

Mustanoja (1960: 506) further claims that the present perfect was used in narrative poetry in Middle English to give emphasis and additional vividness to events or situations in contexts where the preterite or the pluperfect might be more usual.

---

<sup>2</sup> Lawrence (2001: 201), however, comments that variation between the preterite and the present perfect persists through the Middle English period to the Modern English period.

Visser (1963: 2198) also notes uses of the 'epic perfect' in place of the preterite or the pluperfect in Middle English poetry, but ascribes these to the exigencies of metre and scansion. However, while noting that fluctuation between the perfect and other past temporal reference forms is less common in Middle English than tense alternations involving the historical present, Mustanoja (1960: 506) claims that contexts in which tense shifting involving the present perfect occurs bear a resemblance to those in which tense switching involving the historic present is found. According to Mustanoja (1960: 506), narrative uses of the 'historical perfect' in Middle English literature include marking out an event in an objective description; introducing and ending a series of events; and calling attention to the main complicating action. Other researchers similarly highlight the textual functions of the present perfect as an organisational device in Middle English texts. For example, Zimmerman (1973: 536) notes that switches from the preterite to the present perfect in *Sir Gawain and the Green Knight* can mark transition points between narrative episodes.

Burnley (1983: 48) provides a discussion of variation between the preterite, the historical present, and the present perfect in some of Chaucer's works. Consider the following extract from the *Pardoner's Tale* (an interlinear Modern English gloss is provided):

### Extract 1<sup>3</sup>

Togidres *han* thise thre hir trouthes *plight*  
'Together did these three their pledges plight'  
To lyue and dyen ech of hem with oother  
'To live and die, each of them for the other'  
As thogh he were his owene ybore brother

---

<sup>3</sup> Tense forms of interest are italicised in the extract.

'As if he were his very own blood brother'  
 And vp they *stirte* al drunken in his rage  
 'And up they started, drunken, in this rage'  
 And forth they *goon* towards that village  
 'And forth they went, towards that village'  
 Of which the tauerner hadde spoke biforn  
 'Of which the innkeeper had told before'  
 And many a grisly ooth thane *han they sworn*  
 'And many a grisly oath they swore'  
 And Cristes blessed body they *torente*  
 And Christ's blessed body once more rent  
 (cited in Burnley 1983: 48)

Burnley (1983: 49) notes that the tense mix includes the perfect (*han...plight*); two presents (*stirte; goon*); another perfect (*han they sworn*); and finishes with a preterite (*they torente*). As Burnley (1983: 49) points out, all three tenses refer to the same narrative past. Moreover, in Burnley's view (1983: 49), such tense alternations are not merely motivated by metrical requirements, but appear to contribute to the vividness of the narrative.<sup>4</sup>

Although there are interesting comparisons to be drawn between tense-switching in medieval literature and modern oral narratives (see Fleischman 1990), Romaine (1984d: 120) makes the important point that as languages circumscribe the semantic space occupied by tense-aspect forms in different ways, the dislocation of tense-aspect values from their conventional referential functions for pragmatic purposes cannot be assumed to have the same expressive value cross-linguistically or, when

---

<sup>4</sup> Most discussions of tense-switching in earlier varieties of English focus on the semantic properties of particular tense forms rather than on the significance of the switches themselves (see further 5.3 below).

comparing varieties of the same language, in different time periods. Thus, the pragmatic functions of tense alternation phenomena can only be reliably analysed with reference to a particular system at a particular time. Furthermore, extrapolating from Romaine's claim (1984d: 120) relating to cross-linguistic variation in the use of tense-aspect categories for pragmatic purposes, it may also be the case that not all varieties of a language make use of identical tense-aspect forms in discursive patterns of tense-switching. I return to this issue below.

### 5.3 TENSE ALTERNATION IN PREADOLESCENT NARRATIVES

Previous research on tense alternation between the simple past and the CHP has highlighted the textual and pragmatic multivalency of this phenomenon. Tense-switching between the CHP and simple past has variously been described as a device for organising a story into chronological segments (Wolfson 1979: 174); an expressive device for highlighting the relative importance of various narrative events (Silva-Corvalán 1983: 774); a discourse mechanism which correlates with the marking of the narrative peak or climax (Fludernik 1991: 374); and a breakthrough into performance (Wolfson 1982: 91), where the use of tense alternation, concomitant with direct speech, asides, repetition, expressive sounds, motion and gestures, indexes a performed recount rather than a simple report of past events. Fleischman (1990: 260) comments that a major expressive function of tense-aspect oppositions in the structuration of narratives is to 'discriminate different focalizations or perspectival filters through which elements of a story are projected.'

Although the use of tense-aspect forms for pragmatic purposes has long been recognised, traditional approaches to tense alternation between the simple past and the CHP, which typically view present tense forms in past tense contexts as a device for

heightening events by making them appear to unfold in the 'present' (i.e. 'the past-more-vivid'), have been criticised for failing to account adequately for all the functions of the CHP in literary texts and oral narratives (Brinton 1992: 221). Recent discourse-analytical approaches emphasise that it is the switch itself rather than any inherent semantic property of particular tense forms that is significant (see Wolfson 1982: 116; Fludernik 1991: 390).

Although switching between the simple past and the CHP is the dominant pattern of tense variation in complicating action clauses in the preadolescent corpus, there are also cases where tense alternation involves shifts between the simple past, the CHP and the present perfect. This is illustrated in the following narrative recounted by a ten year old girl I recorded who was asked about childhood accidents she had experienced:<sup>5</sup>

## Extract 2

<b>ORIENTATION</b>		my little brother he was like just playing around on his scooter
<b>COMPLICATING ACTION</b>	a.	and he fell off
<b>EMBEDDED ORIENTATION</b>		and then cos I was riding around with him
<b>COMPLICATING ACTION</b>	b.	and I called out to see if he was alright
	c.	he got up
	d.	and he run up to me and as he run up to me
	e.	I fell backwards
	f.	and I really banged my head on the wall
	g.	and my brother, my brother ... then I got my brother
	h.	and he ran up to my mum
	i.	and said, 'Denise's really hurt herself, she's really, really hurt herself'
	j.	and I've gone, 'I'm really hurt, AAAH, I'm really hurt'

<sup>5</sup> Following Schiffrin (1981), only clauses which constitute narrative events (i.e. complicating action clauses) are lettered. Subordinate clauses as well as clauses belonging to other narrative sub-components (e.g. orientation) are unlettered.



- k. and then he go... he goes back out the garden
  - l. gets his scooter
  - m. and puts it right on top of me by an accident
  - n. and then he go... he comes back in
  - o. and he .. he cuddles me  
and then as I've moved
  - p. he's hit his head on the wall
  - q. and then my brother's jumped on top of us
- oh it was really funny  
I couldn't stop laughing

## EVALUATION

The above narrative contains a cluster of devices which co-occur with the CHP and conjunctively operate as a means of conveying affect and creating involvement in discourse. The repertoire of involvement strategies used to engage the listener in the unfolding narrative includes the use of expressive noises, repetition, intensifiers such as '*really*', constructed dialogue, as well as numerous tense switches. According to Wolfson's (1982: 105) criteria, the above narrative would count as a performance rather than a simple report of past events.

In terms of variation in tense usage in the above narrative, clauses a-i contain verbs in the simple past which foreground the series of events culminating in the girl's accident, and clauses k-o contain verbs in the CHP, which describe how the girl's brother exacerbated the situation by accidentally placing his scooter on top of her. Forms in the simple past and the CHP cluster together rather than rapidly alternating between each other. Building on Wolfson's (1982) insights, other researchers have claimed that the switch from morphologically marked past forms to the CHP signifies a narrative turn of events (see e.g. Fludernik 1991: 374), and marks experientially important episodes, as seen in the example above.

With regard to narrative macrostructure, it has been found (e.g. Schiffrin 1981: 51) that tense alternation between the past and the CHP follows a general pattern

where the verb forms in the initial sequence of the complicating action are in the past tense, with switches to the CHP occurring after several clauses, followed by possible additional switches between past and CHP forms, and finally ending with a switch back to the past tense towards the resolution of a narrative. The girl's narrative above shows some adherence to this pattern, but there are several noteworthy anomalies. Firstly, the use of the present perfect in the quotation frame in *j* is unexpected, and is used in place of either the simple past or the CHP. Secondly, towards the end of the narrative, there is no switch back to the simple past, rather the verbs in *p* and *q* preceding the evaluation are in the present perfect. In the next section, I turn to a more detailed consideration of this phenomenon.

### 5.3.1 *Variation in the use of the present perfect*

The examples cited above do not constitute isolated incidents of the use of the present perfect in complicating action clauses in the preadolescent narratives: although the present perfect is comparatively much less frequent than either the simple past or the CHP in the corpus, several narratives make extended use of the present perfect in plot-advancing sequences. Consider the following extract from a narrative told by a seven-year-old boy who animatedly recounts an adventure he had with a friend in the snow:

#### **Extract 3**

- a. .... we stopped
- b. and I went, 'Mark come back! Shall we make a run?'
- c. and then he's gone ... then he's gone, 'But what out...  
with what though?'
- d. and I've gone, 'Wood, you div!'
- e. and then ... we've made it
- f. and the snow's come on it
- g. and I've gone <SOUND EFFECT INDICATING SPEED>

h. and I couldn't land

In the above example, the majority of the verbs which constitute the backbone of the narrative are in the present perfect rather than in either the simple past or the CHP.

Next, consider an example from a ten-year-old boy who recounts a car accident in which he sustained minor injuries:

#### Extract 4

- it was the High Roads (*sic*) and ...
- a. well he come through the lights  
when it was red to try and turn right  
cos he didn't think we was coming down
  - b. and he went straight through the green light  
and as he's turned
  - c. he's hit us
  - d. and we've hit him
  - e. and I've gone forward ...

In the above narrative, complicating action clauses **a** and **b** are in the simple past, but rather than switching to the CHP to foreground the pivotal events of the car accident, the speaker switches to the present perfect in complicating action clauses **c**, **d** and **e** as he proceeds to relate the climax of his story. Fleischman (1990: 142) notes that the dramatic peak of a narrative is a major locus of 'linguistic turbulence' which is often marked by tense shifting. Strategic shifts between one tense form and another at narrative peaks of emotional intensity can function evaluatively to emphasise the significance of dramatic high-points (see Fleischman 1990: 5; Georgakopoulou 1994: 205; Georgakopoulou and Goutsos 1997: 107). Whereas previous research has noted the segmentational and evaluative functions of tense alternations between the simple past and the CHP in narrative discourse (Brinton 2001:143), what is novel about

Extract 4 is that similar pragmatic functions appear to be fulfilled by switches between the simple past and present perfect.

The use of the present perfect in foregrounded narrative clauses to describe discrete bounded events in a temporal sequence has not been commented on extensively in previous research, and warrants further investigation ( see, however, Rickford and Théberge-Rafal 1999 discussed below). Wolfson (1982: 4) refers briefly to the use of the 'CHP perfect' in American narratives, which she claims is referentially equivalent to the past perfect. However, it is clear that the use of the present perfect in the preadolescent narrative extracts cited above is not substitutable for the past perfect. According to Comrie (1985: 65), the past perfect is an absolute-relative tense which situates a reference point in the past and locates an event prior to that reference point. In the preadolescent narratives, the present perfect in complicating action clauses is not employed in an out-of-sequence function, but is used to describe punctual events in the past in a sequence of iconically ordered clauses.

It is important to stress, then, that the use of the present perfect in the children's narratives differs from its conventional referential functions in discourse. According to Elsness (1997: 49), an essential difference between preterite and present perfect forms in standard English correlates with the discourse-pragmatic functions of foregrounding and backgrounding information. Hopper (1979:215) relates the foreground to the backbone of a narrative which contains the main-line events, and is characterized by the use of past tense verb forms, or a special narrative tense such as the CHP, for narrating a sequence of temporally bounded discrete events (see also Bybee *et al.* 1994: 62). The English present perfect, on the other hand, is said to be relational: its fundamental characteristic is claimed to be its association with the

present time (Comrie 1976: 52; Tagliamonte 2000a: 341). The prototypical function of perfects is not to locate a situation at some definite point in time, but to relate past actions to present situations (see Bybee *et al.* 1994: 61).<sup>6</sup> In narrative discourse, the present perfect is ordinarily rare. This rarity has to do with the fact that the present perfect is normally limited to contexts where the narrator momentarily steps out of diegetic time to address the audience directly (Fleischman 1986: 212). Givón (1983: 232) explains the low frequency of perfects in narrative discourse on the basis of the fact that they are typically used to mark ‘out-of-sequence clauses.’

In the preadolescent narrative extracts above, the use of the present perfect to sequence discrete bounded events in the past suggests that it can be used with a foregrounding function.<sup>7</sup> The encroachment of the present perfect on semantic domains prototypically associated with the simple past or preterite is of course amply documented in a number of Indo-European languages such as French, Romanian and Northern Italian (see e.g. Fleischman 1990: 30; Bybee *et al.* 1994: 81).

The extensive variation between the present perfect and the simple past attested in both synchronic and diachronic varieties of English (see Tagliamonte

---

<sup>6</sup> Elsness (1997: 67) notes that traditional grammars of English commonly refer to notions of ‘current relevance’ in defining the meaning of the present perfect in relation to the simple past. The explanatory power of ‘current relevance’ to account systematically for differences in the use of the present perfect in relation to the simple past is, however, somewhat restricted (Klein 1992: 531). Other studies of variation in temporal reference sectors (e.g. Poplack and Turpin 1999) note that there is often a considerable discrepancy between the actual usage of variant forms and the values ascribed to these forms in the prescriptive literature. Furthermore, as Poplack and Turpin (1999: 140) note, distinctions in ‘referential value or grammatical function among different surface forms can be *neutralized* in discourse.’

<sup>7</sup> Elsness (1997: 48-9) notes that the present perfect occasionally has a foregrounding function in narrative where it occurs at the beginning of a sequence of past-time references and is not associated with any past-time anchor (e.g. *Mary has inherited a house from an uncle. It came as a complete surprise*; cited in Elsness 1997: 49). Cox (2005: 115) notes that such uses frequently occur in news reports (e.g. *President Kennedy has been shot*). These uses exemplify what is often referred to as the ‘Hot-News Perfect’ which occur in contexts ‘where recency or completion is being signalled’ (Cornips and Corrigan 2005a: 103) or where there is some attempt to concentrate the hearer’s attention on an event (Schwenter 1994: 1001). In Schwenter’s (1994: 997) view, ‘hot news’ functions of the present perfect are ones which come closest to the function of pasts or perfectives and can therefore be seen as an important link in the grammaticalisation of these forms to perfectives. However, there is some debate as to whether the Hot-News Perfect constitutes a distinct sub-type of perfect (see e.g. McCoard 1978: 189).

2000a: 330) bolsters Harris's (1984: 316) claim that 'the modern perfect has not established itself uniformly across all dialects of English.' Moreover, recent studies on cross-dialectal variation in other domains of temporal reference (see e.g. Poplack and Tagliamonte 1999 on future temporal reference) have shown that differences between dialects can be correlated with the relative degree of grammaticalisation of tense forms across communities, with different varieties located at different points along the trajectory of change (see Poplack and Tagliamonte 1999: 316). Of particular relevance here is Howe and Schwenter's (2003: 65) recent investigation of the use of the present perfect in European Spanish, where there is evidence that the present perfect is being extended into the semantic space of the preterite. This phenomenon appears to be more advanced in Madrid than in Alicante, and exhibits significant differences based on age, with the younger age groups in the vanguard of change. This change shows differential progress in different varieties of Spanish, and is percolating through successive generations.

Howe and Schwenter (2003:65) also note that in Alicante and Madrid the present perfect is used innovatively in place of the preterite to mark foregrounded narrative clauses, although this functional extension is strongly conditioned by temporal factors such as the distance of the past reference point from utterance time. Extrapolating from the findings discussed by Howe and Schwenter (2003) for Spanish, there is no reason to assume that the present perfect in contemporary varieties of English is evolving in the same way in different dialects, or has

grammaticalised to the same extent.<sup>8</sup>

Viewed against the backdrop of extensive historical variation in tense usage, patterns of tense-shifting in the preadolescent narratives discussed above provoke questions of a broader nature relating to the evolution of tense-aspect categories in contemporary varieties of English. Is variation between the simple past and the present perfect in preadolescent narrative discourse simply a remnant of diachronic patterns of variation that have persisted in contemporary non-standard vernaculars (see Denison 1993: 352; Labov 1980: xvii)? Or, given the trajectory of change from discourse to grammar reported in previous research (see e.g. Sankoff and Brown 1976; Sankoff 1990), are discourse-pragmatic uses of the present perfect in narrative indicative of its ongoing expansion and grammaticalisation in certain varieties of English? Alternatively, what evidence is there that preadolescent speakers' use of the present perfect is not simply an age-graded or developmental feature?

In the following section, I attempt to shed some light on these questions by drawing on evidence of patterns of tense-switching from older speakers in the south of England in addition to examining similar tense-switching phenomena in other varieties of English.

---

<sup>8</sup> A cursory glance at contemporary varieties of English suggests that different varieties may reflect different stages in the diachronic evolution of the present perfect. Harris (1984: 322) notes that Hiberno-English perfect forms such as *I haven't it written out properly yet* are retentions of an older 'split perfect' found in Early Modern English; for example: *Have you the lion's part written?* (Shakespeare, *Midsummer Night's Dream* I.ii.68). However, in Hiberno English, structures such as the 'split perfect' may well be the result of interconnecting internal and external influences. In the case of the 'split perfect', there is also the possibility that the constituent order of this structure has been influenced by an equivalent Irish form. In recent language contact research which focuses on variation in English in Ireland (see e.g. Corrigan 2000 on aspects of deontic modality in Northern Irish English), transfer effects from Irish dialects; the legacy of early northern British vernaculars; as well as aspects of second language acquisition are all invoked to explain variable grammatical structures.

#### 5.4 TENSE VARIATION FROM A CROSS-VARIETY PERSPECTIVE

A diagnostic feature of the modern present perfect, in contrast with the preterite, is its prohibited collocation with a time adverbial referring to a specific time point or period in the past; for example *'I have arrived yesterday'*. Comrie (1985: 78) comments that this type of collocation restriction appears to be idiosyncratic to English. However, when evidence from vernacular varieties of English is considered, a more complex state of affairs emerges. For example, Harris (1984: 315) points out that in Hiberno English, the present perfect is used in contexts where standard English would conventionally use the preterite. Harris (1984: 315) cites the following examples:

(2) I've done a course two years ago

(3) They've been here when we came

Similarly, Bauer (1994: 401) claims that sentences of the type *'I have seen it last week'* are increasingly found in New Zealand English. Trudgill (1984: 42) also claims that there has been an increase in the use of such constructions in southern British English, and Denison (1998: 192-3) observes that the present perfect has started to occur quite frequently in colloquial varieties of British English in the narration of recent past events. More recently, in their global synopsis of morphosyntactic variation in English, Kortmann and Szmrecsanyi (2004: 1154) cite the levelling of the difference between the present perfect and the simple past as one of the most frequent features of worldwide varieties of English.

In spite of reports that the present perfect may be extending its range of application in contemporary vernaculars, there are claims in the research literature that the present perfect has peaked in terms of its frequency in modern varieties, and is now tending to recede. According to Elsness (1997: 341), for example, after an



initial increase in the distribution of the present perfect in relation to the preterite up until the Early Modern English period, there is evidence that within the Modern English period 'the increase in the ratio between the present perfect and the preterite has not been arrested but reversed as far as American English is concerned.' Furthermore, Elsness (1997: 341) claims that this decline is apparent in British English. Sampson (2002: 17) argues that in Britain, there is marked regional variation in the use of the present perfect. Based on an analysis of spontaneous speech data collected in the 1990s, Sampson (2002: 24) claims that the present perfect is a marginal category in southern English vernaculars. According to Sampson (2002: 22), the relatively low rates of occurrence of the present perfect construction in the vernacular of southern English speakers reflect a recent evolutionary trend towards the loss of the perfect category in southern non-standard speech.

In view of the fact that both Elsness (1997) and Sampson (2002) base their conclusions concerning the apparent recessiveness of the present perfect on data drawn largely from adults, I now return to the question of whether the use of the present perfect in the preadolescent narratives cited above represents age-grading. This issue is also raised by Rickford and Théberge-Rafal (1999) in their discussion of preterite *had* + *V-ed* in the narratives of African American preadolescents. At this point, it is useful to compare variation in the present perfect in the London preadolescent corpus with their research, as there are a number of parallelisms between their findings and mine. In both studies there is evidence that the perfect serves foregrounding functions in narrative; however, in Rickford and Théberge's (1999) corpus, it is the past perfect rather than the present perfect which participates in this extended range of uses.

Using a corpus of narratives collected from African American preadolescents in East Palo Alto, California in 1988, Rickford and Théberge-Rafal (1999: 34) found that the past perfect was variably used by young speakers to mark the simple past in narrative recounts, as in the following extract adapted from Rickford and Théberge-Rafal (1999: 48).

#### Extract 5

- a. And then she hit kinda hard in his face, but not that hard ...
- b. And then so they *had walked off*
- c. And Gerald started walking off, said, 'I'ma bring my peoples!'
- d. Then Corey walked up to him
- e. Corey Corey walked up to him
- f. and decked him in the eye, I mean in the jaw
- g. And then he said, 'I'ma bring my people'
- h. And then and ... we *had came* around a corner, and then we *had came* around a corner
- i. We *had went* home
- j. And then Gerald mother and him come up

Of particular interest is Rickford and Théberge-Rafal's (1999) finding that 94% of their preterite *had* tokens occurred in narrative complicating action clauses, where, in a similar fashion to the present perfect in the London preadolescent narratives, it was found to mark the prelude to narrative peaks, in addition to marking dramatic points of intensity themselves (Rickford and Théberge-Rafal 1999: 48). Similar results have also been obtained from African American speakers elsewhere in the United States (see e.g. Cukor-Avila and Bailey 1995; Ross *et al.* 2004), and preterite *had + V-ed* is also reported to exist among young Puerto Ricans in New

York City (Rickford and Théberge-Rafal 1999: 53).<sup>9</sup> However, as Rickford and Théberge-Rafal (1999:55) cautiously remark, the social distribution of this construction warrants further detailed analysis across a wider range of age groups in order to distinguish between developmental phenomena and bona fide linguistic change.<sup>10</sup>

This caveat equally applies to the London preadolescent data: the extent to which older British speakers use the present perfect in narrative in similar ways to the London preadolescents remains a moot point. There is some provisional evidence, however, that unusual uses of the present perfect in the narrative clauses of younger speakers do not simply reflect developmental tendencies.<sup>11</sup> Consider, for example, the following extract from the *Queen Mary Narrative Corpus* (Cheshire 2005b). The speaker is a twenty-year-old British student who recounts his frustration at trying to print out an important document:<sup>12</sup>

---

<sup>9</sup> Romaine (2000: 182) notes that in Hawai'i Creole English, the creole form *had play*, which resembles standard English *had played*, is equivalent to the simple past. Sharma (2001: 344) also reports that in Indian English, *had + V-ed* can sometimes be used with either a preterite or present perfect meaning (see also Leitner 1991: 228 on the 'remote past' uses of the past perfect in Indian English).

<sup>10</sup> I have scrutinised my own corpus of preadolescent narratives to see if the speakers I recorded displayed any unusual use of the past perfect in their recounts. The only example I have found in the entire corpus where the past perfect does not appear to act as an absolute-relative tense (see Comrie 1985) is shown below in an extract taken from a story told by a seven-year-old girl:

<b>ORIENTATION</b>	it was a really high slide and he was walking up it
<b>COMPLICATING ACTION</b>	a. and he <i>had slipped</i> b. and he fell off the actual slide c. and broke his arm...

It is not clear that the form *had slipped* in the initial complicating action clause is functioning as a conventional past perfect or whether it is encoding a preterite function analogous with preterite *had + V-ed* forms discussed in Rickford and Théberge-Rafal (1999). Further evidence which argues in favour of the latter interpretation comes from the fact that *had slipped* occurs in the first complicating action clause and does not appear to have a 'flashback' function that can be encoded by the past perfect in English (see Comrie 1986: 18-19; and Rickford and Théberge-Rafal 1999: 36). Nevertheless, as this is the only apparent example of preterite *had* in the London preadolescent corpus, my interpretation must necessarily be treated with caution.

<sup>11</sup> Narrative uses of the present perfect by adult British speakers have been recently highlighted by Walker (2007), who discusses whether these constitute a newly emergent phenomenon. Walker (2007) discusses examples from Scotland, London and Newcastle.

<sup>12</sup> The speaker in question lives in Surrey in the southeast of England and was recorded in 2000. I am grateful to Jenny Cheshire for making this material available to me.

## Extract 6

- and it was quite a big document  
it was like thirty pages long
- a. and all of a sudden it's gone through
  - b. and the printer's just started printing
  - c. it's got to page thirty
  - d. and it's kept going
  - e. and I couldn't stop the printer ...

In the above example, clauses **a**, **b**, **c** and **d** all contain a verb in the present perfect which relate a sequence of past events. Interestingly, both Wolfson (1979: 174) and Fludernik (1991: 374) note that the CHP frequently co-occurs with new events that are introduced by the adverbial *all of a sudden*. In Extract 6, however, it is the present perfect rather than the CHP that co-occurs with this adverbial, which further suggests that the present perfect, like the CHP, can be strategically employed by speakers to indicate a turn of events in narrative structure.

Looking beyond Britain to other varieties of English, it appears that the present perfect participates in discourse-level variation in the narratives of adult speakers which is reminiscent of similar patterning in the London preadolescent data. Engel and Ritz's (2000) analysis of tense variation in contemporary Australian English presents a detailed exposition of the discourse functions of the present perfect in data from radio programmes and chat shows. As in the case of the London preadolescent narratives, Engel and Ritz (2000: 13) note that in their corpus the present perfect can be used in a foregrounding function in sequences which indicate narrative progression. Consider the following example adapted from Engel and Ritz (2000: 133) where the present perfect not only occurs in the orientation section but also in the complicating action clauses **a-c**, as well as in the resolution section in clause **e**.

## Extract 7

<b>ORIENTATION</b>	in the morning , he's stuck an 'I love Redman' sticker on her back so there she is, she's up there cleaning the Mount Hawthorne TAB, you know, with 'I love Redman'
<b>COMPLICATION</b>	a. and then she's left b. she's gone to Innaloo Shopping Centre c. she's gone to Warwick Grove Shopping Centre
<b>EVALUATION</b>	Right ? Everyone's looking at her and she's sort of really paranoid.
<b>RESOLUTION</b>	d. she finally got home to her husband and kids e. and they've just pissed themselves laughing .... ( 96 FM Radio Perth, September 1 1999)

Engel and Ritz (2000: 137) suggest that the shift to the present perfect in the complicating action sequences acts as an expressive device for highlighting the pivotal events of the story. Yet again, we see that a shift to the present perfect has a pragmatically similar effect to a shift to the CHP in narrative. Both types of switch can demarcate boundaries within narrative discourse as well as evaluate specific narrative sub-components ( see e.g. Schiffrin 1981: 59 on the CHP).

Similar discourse functions of the present perfect have also been documented in New Zealand English. Cox's (2005) analysis of 1600 past-time narrative verb forms collected from a police reality television programme showed that not only did the present perfect occur with past-time adverbials, it was also used as a narrative past tense in ways which exhibit interesting parallels with the data discussed by Engel and Ritz (2000) for Australian English. Consider the following example adapted from Cox (2005: 110):

## Extract 8

### ORIENTATION

The police vehicle was at the top of the road putting in a cordon.

### COMPLICATION

- a. The van has turned up Bruce Road
- b. turned left into Chivalry Road <sup>13</sup>
- c. and the police car has had to stop the van.

(Tonight TV One 6/12/04)

Cox (2005: 75) notes 20% of the past-time-referring narrative verbs in the corpus that she compiled consists of preterite-like uses of the present perfect, as illustrated above. Furthermore, Cox (2005: 76) observes that the proportion of CHP forms in her corpus is markedly lower (8.17%) than that reported in Schiffrin's (1981: 47) study of tense variation in narrative (see also Tagliamonte and Poplack 1988: 520). This leads Cox (2005: 76) to hypothesise that the lower rate of CHP usage may be a result of New Zealand English speakers using the present perfect in place of the CHP.

In order to investigate further how variation in British preadolescents' use of the present perfect compares with tense alternation in other varieties of English, I now turn to a distributional and multivariate analysis of the data to determine which factors contribute to the use of the present perfect in complicating action clauses, and to examine its patterning with respect to the other tense forms employed in the same narrative sub-component.

## 5.5 METHODOLOGY

56 narratives in total were extracted from the main preadolescent corpus for the specific purpose of analysing tense variation. Details of the composition of the narrative sub-corpus are given in Table 5.1 below.

---

<sup>13</sup> Although Cox (2005) does not point this out, the tense of the verb in b is ambiguous. The verb in b could either be a simple past form, or a present perfect with the auxiliary deleted from the conjoined clause (see also 5.5 below). In Schiffrin's (1981) terminology, such a case would count as a 'neutralization site' in a quantitative analysis of the data.

TABLE 5.1

## COMPOSITION OF THE NARRATIVE SUB-CORPUS BY AGE AND GENDER OF SPEAKERS

Age and gender of speakers	Number of speakers	Number of Narratives
7-8 year old males	6	12
7-8 year old females	7	13
10-11 year old males	7	14
10-11 year old females	8	17
TOTAL	28	56

The verb in each complicating action clause was initially coded for tense. Unclear or ambiguous contexts were excluded from the analysis, as in (4) (italicised below), which was considered to be anomalous, and in (5) where it was impossible to ascertain whether the second verb was either a preterite form, or a present perfect form with the auxiliary deleted in a clausal conjunct:

(4) I've gone, 'What?' and *they're started pushing me* [12M7/8]

(5) and *I've gone* off on to the road and *scraped* all my arm [13M10/11]

Other factors which needed to be taken into account at the coding stage related to the status of non-standard verb forms as in (6) and (7) below:

(6) well, he *come* through the lights when it was red to try and turn right  
[3M10/11]

(7) I was watching the TV so ... I *done* something back to him [8M10/11]

The use of non-standard *come* in past reference contexts exemplified in (6) above is a widely attested non-standard preterite form encountered in English dialects (see Chapter 2; and Tagliamonte 2001). Similarly, non-standard *done* in (7) above is

in widespread use in many vernacular varieties of English where it has assumed generalized past marking functions (see Chapter 2; and Hughes and Trudgill 1979: 15). These variants were accordingly coded as simple past forms.

571 tokens were extracted in total from foregrounded narrative clauses and retained for analysis. Each token was coded for the two external factors of age and gender (see Wolfson 1982: 78). In addition, each token was coded for a number of internal factors which previous research has claimed to influence the choice of tense form in narrative. I briefly summarise these in turn below.

Firstly, in order to explore the possible interaction between tense alternation and text-marking functions, it was necessary to code for the location of particular tense forms within the sequence of complicating action clauses. Each token was therefore categorised with regard to whether it occurred in a sequence-initial, medial or final location.

Secondly, it was important to account for potential discourse-level effects such as parallel processing (see Scherre and Naro 1991) whereby the choice of verb form is influenced by the form used in the preceding clause resulting in the clustering of verbs in the same tense in successive clauses (see Tagliamonte and Poplack 1988: 520). Each verb form in a non-initial sequence was accordingly coded with regard to the tense form used in the preceding complicating action narrative clause.

The grammatical person of each verb form was also incorporated into the coding system. Elsness (1997: 342) notes, for example, that there is a tendency for the present perfect to be used with first person pronouns, whereas the preterite often occurs more frequently with third person subjects.

Another major factor included in the coding procedure was temporal disambiguation. Previous research has suggested that the presence of adverbial



specification may influence the use of overt past-tense marking (see e.g. Tagliamonte 2001: 50). A clause that is already temporally disambiguated by the presence of an adverbial might be assumed to be more conducive to the use of a morphologically unmarked verb form in past temporal reference contexts.

Furthermore, as already indicated above, one of the characteristics of the present perfect in English is the apparent constraint on its collocation with a time adverbial referring to a specific point in time (see e.g. Comrie 1976: 54; Comrie 1985: 78; and Elsness 1997: 25). If this constraint is operative in the preadolescents' vernacular, then it is reasonable to expect that the presence of a temporal adverbial referring to a specific time point would have an inhibiting effect on the use of the present perfect.

Finally, in the process of coding the material, it was noticed that the quotatives *go* and *say* frequently co-occurred with the CHP (see Chapter 7), and that *go* also attracted the use of the present perfect (see also Engel and Ritz 2000: 136). In order to study the patterning of *verba dicendi* with particular tense forms, a separate factor group was established which distinguished between non-quotative contexts and environments where a lexical quotative (e.g. *say*, *go*) was followed by direct speech.

In the following section, I present a distributional analysis of the tense forms in the narrative data before preceding to the multivariate analysis in order to determine which factors are significant in conditioning the choice of tense forms.

## 5.6 DISTRIBUTIONAL ANALYSIS

Table 5.2 below provides an overview of the distribution of tense forms in complicating action clauses in the narrative data.

TABLE 5.2

## OVERALL DISTRIBUTION OF TENSE VARIANTS USED IN COMPLICATING ACTION CLAUSES

Tense Variants	N	%
Preterite	340	60
CHP	181	32
Present Perfect	50	9
TOTAL N	571	

From a proportional perspective, it can be seen that the preterite is the unmarked tense form in complicating action clauses; thus, the CHP and present perfect forms can be considered marked (see Fludernik 1991: 369). The frequency of the CHP at 32% is comparable with that noted in previous research. In her study of American narratives, Schiffrin (1981: 47) found that 30% of verb forms in narrative clauses were in the CHP, and Silva-Corvalán's (1983: 767) study of tense variation in Spanish narratives similarly reports that 156 (32.7%) out of 476 complicating action clauses contained verbs in the CHP. Tagliamonte and Poplack (1988: 520) also note that 32% of the complicating action clauses they analysed in Samaná English contained CHP forms. In contrast with the frequency data discussed in Cox (2005: 76) for New Zealand English, there is little evidence in the preadolescent data displayed above that the CHP is being used at much lower rates than those reported in previous studies.

Although the frequency of the present perfect is low at 9%, the very fact that it occurs at all in an environment where previous research has noted its absence (e.g. Labov and Waltezsky 1967 : 43, note 7) is significant.

Table 5.3 below gives a breakdown of the location of tense forms in the complicating action clauses.

TABLE 5.3

LOCATION OF TENSE FORMS IN COMPLICATING ACTION CLAUSES

Location	Preterite	Conversational Historical Present	Present Perfect	TOTAL N
Initial Clause	41 (73%)	12 (21%)	3 (5%)	56
Medial Clauses	261 (56%)	161 (35%)	42 (9%)	464
Final Clause	38 (75%)	8 (16%)	5 (10%)	51

In terms of discourse structure, the fact that certain tense forms cluster in specific locations in complicating action sequences corroborates the suggestion that tense-switching performs text-level organisational functions. Although the preterite is the dominant tense form in initial, medial and final clauses, it faces greater competition from the other two variants in medial clauses. The high frequency of the preterite in initial and final clauses is consonant with Fludernik's (1991: 373) claim that the initial and final sections constitute a macro-structural narrative frame which characteristically retains an overt deictic past tense. By contrast, it is noteworthy that the present perfect is at its least frequent in initial position, indicating that it rarely

initiates narrative progression in complicating narrative clauses and relies on other temporal indicators in initial sequences, most notably the preterite, to establish an overt temporal frame in order to be used to advance the plot line.<sup>14</sup> Tense-switching from the preterite to the present perfect in medial clauses is illustrated in Extract 9 below in which a seven-year-old boy recounts how he and his friend were trying to evade the attention of the boy's mother. This example is also significant because it shows that the narrative action closes with an untypical switch from the preterite to the present perfect.

### Extract 9

we were in the boot, we were in the boot  
a. and then she looked in  
b. and we gone, 'Paul!'  
c. I <unclear> shoved him  
d. I've gone like that <PERFORMED ACTION>  
e. and then he's gone, 'What?'  
f. and I said, 'Look out the window'  
g. and then we jumped out the boot  
h. and she's caught us

Figure 5.1 below shows that there is a pronounced tendency for each tense variant to occur in the immediate vicinity of a preceding form of the same type. Formal parallelism across successive clauses gives rise to a clustering effect, which confirms Schiffrin's observation (1981: 51) that rapid switching between tense forms is uncharacteristic.

---

<sup>14</sup> As I noted earlier, the present perfect has been observed to have a foregrounding function at the beginning of news reports. However, as the data in Table 5.3 show, present perfect forms are not frequent in initial complicating action clauses in the preadolescent narratives. As Cox (2005: 115) also observes in the case of her New Zealand English data, the present perfect can be used in the opposite way, with an event being initially reported in the preterite and then further elaborated in subsequent complicating action clauses using the present perfect.

FIGURE 5.1

DISTRIBUTION OF TENSE VARIANTS ACCORDING TO TENSE OF VERB IN PRECEDING COMPLICATION ACTION CLAUSE

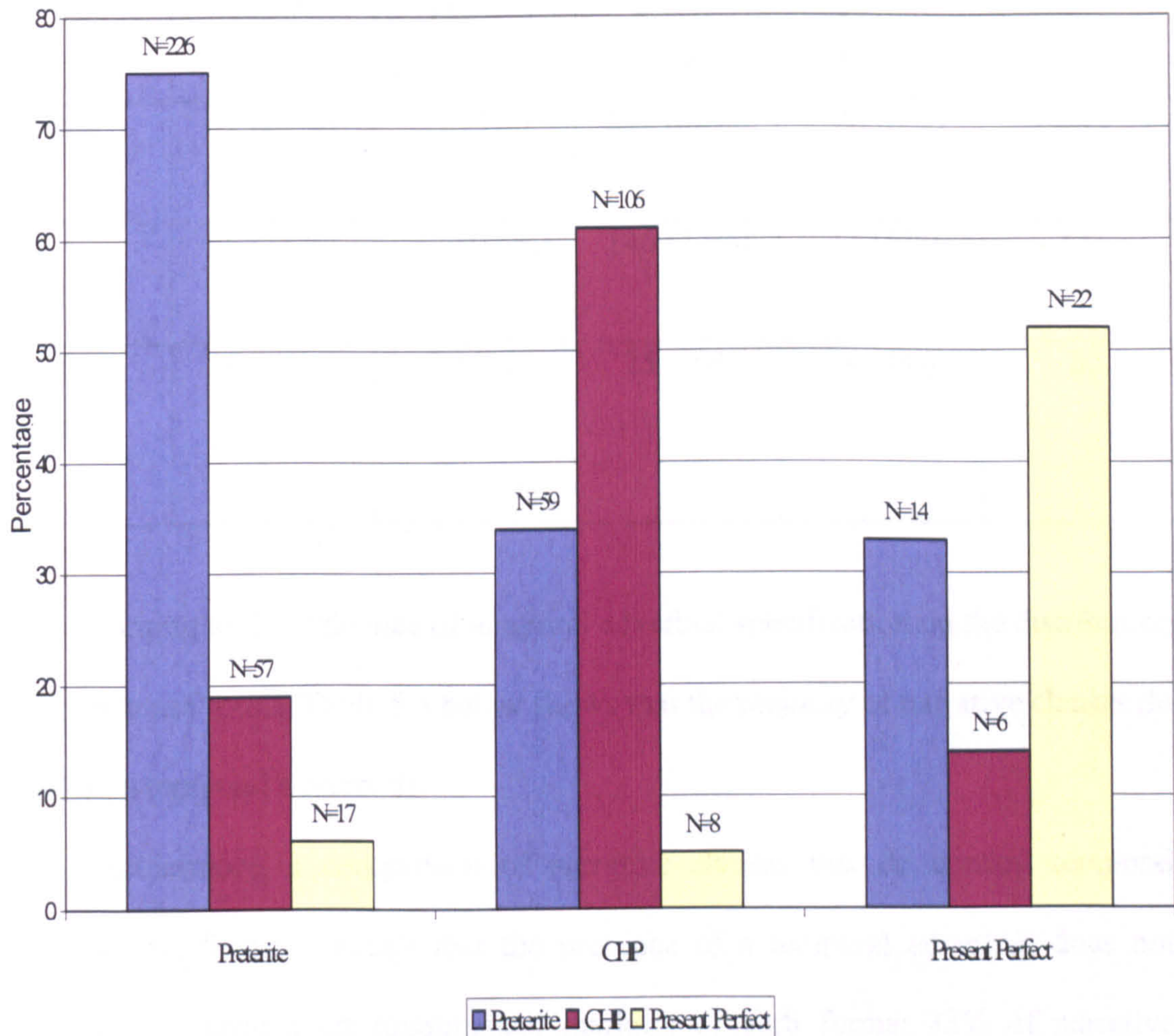


Table 5.4 below shows the distribution of tense forms by grammatical person. Recall that according to Elsness (1997: 327-332), there is diachronic evidence of a tendency for present perfect forms to occur with first person subjects. The data below show that the present perfect co-occurs with first person singular subjects somewhat more frequently when compared with the other tense variants.

TABLE 5.4

DISTRIBUTION OF TENSE FORMS IN COMPLICATING ACTION CLAUSES  
BY GRAMMATICAL PERSON

	1 <sup>st</sup> person singular	1 <sup>st</sup> person plural	3 <sup>rd</sup> person singular	3 <sup>rd</sup> person plural	TOTAL N
Preterite	91 (27%)	18 (5%)	211 (62%)	20 (6%)	340
CHP	29 (16%)	1 (0.6%)	140 (77%)	11 (6%)	181
Present Perfect	19 (38%)	2 (4%)	25 (50%)	4 (8%)	50

Turning to the influence of temporal adverbial specification on the distribution of variant tense forms, Table 5.5 below shows that the majority of narrative clauses do not contain temporal adverbials.

Furthermore, a comparison of narrative clauses that do contain temporal adverbial specification reveals that the presence of a temporal adverbial does not appear to influence overt grammatical marking of verb forms: 33% of narrative clauses containing the simple past co-occur with a temporal adverbial, which contrasts with 27% of narrative clauses containing a verb in the CHP, and 20% of clauses containing the present perfect. Thus, there is no marked tendency for the CHP, which, of the three tense variants, receives the least overt grammatical marking for temporal reference, to be disambiguated by the presence of a temporal adverbial (see also Tagliamonte 1999: 213). The most frequent marker of temporal specification in the data is the distal temporal adverb *then* (Schiffrin 1987: 246) which co-occurs with all three variants. The co-occurrence of the present perfect with a temporal marker of

successive event time (Schiffrin 1987: 249) offers further confirmation of the use of the present perfect in the preadolescent corpus to narrate a sequence of punctual and completed events (compare this with the occurrence of preterite *had* +Verb-*ed* in *then*-marked clauses discussed in Rickford and Théberge-Rafal 1999: 41)<sup>15</sup>

TABLE 5.5

DISTRIBUTION OF ADVERBIALS OF TEMPORAL SPECIFICATION USED IN COMPLICATING ACTION CLAUSES ACCORDING TO TENSE FORM OF THE VERB

	Preterite	Conversational Historical Present	Present Perfect	Total
<b>ADVERBIALS</b>				
then	107	46	9	162
yesterday	1	-	-	1
that night	1	-	-	1
last night	2	-	-	2
this morning	-	-	1	1
at eleven o'clock	1	-	-	1
then later on	-	1	-	1
then in the middle of the night	-	1	-	1
in the morning	1	-	-	1
Total N	113	48	10	171

<sup>15</sup> In New Zealand English, Cox (2005: 112) also notes that the present perfect can have a preterite-like function when it follows the temporal adverb *then*, as illustrated in the following extract from a police report: *My understanding is that she has left [the area] and then he has left.* (Detective Sergeant Brown, The Press, A6, 29/03/2005).

In terms of correlations between tense forms and quotative usage, Table 5.6 below shows that there is a marked difference in the patterning of tense forms in quotative and non-quotative contexts.

TABLE 5.6

DISTRIBUTIONAL COMPARISON OF TENSE VARIATION IN QUOTATIVE AND NON-QUOTATIVE CONTEXTS WITHIN NARRATIVE COMPLICATING ACTION SEQUENCES

	Preterite	CHP	Present Perfect	TOTAL N
Quotative	50 (33%)	92 (61%)	9 (6%)	151
Non-quotative	290 (69%)	89 (21%)	41 (10%)	420

The most noticeable difference is exhibited by the CHP, which occurs much more frequently in quotative contexts than in non-quotative ones. The preterite, on the other hand, is dispreferred in quotative contexts, even though it is the unmarked tense in foregrounded narrative clauses overall. Quotative contexts therefore constitute a major locus for tense-switching into the CHP in the preadolescent narratives (see also Wolfson 1982: 52, and Stenström *et al.* 2002: 122).

The tendency for *verba dicendi* to surface as present tense forms is well-documented in the history of English (Mustanoja 1960: 488; Poplack and Tagliamonte 2001: 138). Some authors (e.g. Romaine and Lange 1991) have claimed that one of the pragmatic functions of variation between verbs of saying in quotative frames is to help demarcate speaker roles. In conjunction with tense alternation between the simple past and the CHP (and vice versa), variation in the form of the lexical quotative may facilitate participant tracking across extended stretches of discourse. Consider the following example, taken from a narrative told by a seven-



year-old girl, where switches between tense forms, as well as lexical variation across quotative frames (*italicised below*), appear to help differentiate speaker roles in the narrative:

**Extract 10**

we all went out in the morning and *goes*, 'Oooh, it's a bomb, it's a bomb' and then my brother run downstairs and *he said*, 'Oh . . .' no, he went in his bedroom and *he said*, 'Oh look I've got some matches... I'll put it alight shall I?' and he put it on . . . and it went right . . . it went right to the bottom and then it just floated up in the air cos it was a balloon and *we all went*, 'Who was it?' and *he goes*, 'I dunno'

Previous research has also shown that extralinguistic factors such as gender and age constrain the use of the tense alternation in discourse. Wolfson (1982: 78) found that in same-gender groups, men were more likely to tell stories with CHP alternation than women. In mixed-gender groups, stories told by men more frequently contained more CHP alternation than stories told by women. Turning to the preadolescent data, Tables 5.7 and 5.8 below show the distribution in the of tense forms by gender and age respectively.

**TABLE 5.7**

**DISTRIBUTION OF TENSE FORMS IN COMPLICATING ACTION CLAUSES BY GENDER**

	Preterite	CHP	Present Perfect	TOTAL N
FEMALE	198 (64%)	94 (30%)	20 (6%)	312
MALE	142 (55%)	87 (34%)	30 (12%)	259

TABLE 5.8

DISTRIBUTION OF TENSE FORMS IN COMPLICATING ACTION CLAUSES  
BY AGE

	Preterite	CHP	Present Perfect	TOTAL N
7-8 year olds	154 (64%)	61 (25%)	27 (11%)	242
10-11 year olds	186 (57%)	120 (37%)	23 (7%)	329

In terms of frequency of use, the males make slightly more use of the CHP than the females, although the difference is not highly contrastive. With regard to the use of the present perfect, however, males make greater use of this form than the females.

When the data are considered in terms of age, it can be seen that there is an increase in the use of the CHP as the children get older, although the ten- to eleven-year-olds make slightly less use of the present perfect than the younger cohort.

### 5.6.1 *Multivariate analysis*

In the multivariate analysis of the data below, I assess the relative significance of both internal and external factors when all of these are considered simultaneously.<sup>16</sup> Table 5.9 below shows the results of the multivariate analysis of the data.

<sup>16</sup> In order for the multivariate analysis to operate effectively, I collapsed first person singular and plural subjects together as well as third person singular and plural subjects. The factor group relating to the potential disambiguating effect of temporal adverbs is also excluded from the results and ensuing discussion as initial multivariate analysis of the data revealed this to be insignificant.

TABLE 5.9

THREE INDEPENDENT MULTIVARIATE ANALYSES OF THE CONTRIBUTION OF FACTORS TO THE PROBABILITY OF THE USE OF PRETERITE, CONVERSATIONAL HISTORICAL PRESENT, AND PRESENT PERFECT FORMS IN COMPLICATING ACTION CLAUSES

INPUT	Preterite		CHP		Present Perfect		TOTAL N
	FW	%	FW	%	FW	%	
<b>Grammatical person</b>							
First person	0.61	68	0.32	19	0.59	13	160
Third person	0.46	56	0.57	37	0.46	7	411
Range	15		25				
<b>Location of tense form in complicating action clause</b>							
Initial clause	0.61	73	0.43	21	0.38	5	56 <sup>17</sup>
Medial clauses	0.47	56	0.54	34	0.51	9	464
Final clause	0.68	75	0.26	16	0.52	10	51
Range	21		28				
<b>Tense form used in preceding complicating action clause</b>							
Preterite	0.68	75	0.35	19	0.46	6	300 <sup>18</sup>
CHP	0.27	34	0.78	61	0.40	5	173
Present Perfect	0.21	33	0.29	14	0.94	52	42
Range	47		49		54		
<b>Quotation</b>							
Non-quotative	0.60	69	0.38	21	0.55	10	420
Quotative	0.25	33	0.80	61	0.37	6	151
Range	35		42				
<b>Gender</b>							
Male	0.45	55	0.52	34	0.58	12	259
Female	0.54	63	0.49	30	0.44	6	312
<b>Age</b>							
7-8 Year olds	0.55	64	0.42	25	0.54	11	242
10-11 Year olds	0.47	57	0.56	36	0.47	7	329
Range			14				

<sup>17</sup> The number of final clauses differs from the number of initial clauses because a small number of narratives lacked a resolution clause (Labov 1972: 370) owing to interruptions from other speakers, which prevented narrators from concluding their narratives.

<sup>18</sup> The total denominator for this group differs from that of the other factor groups as initial complicating action clauses were necessarily excluded.

Taking extralinguistic factors first, it can be seen that only age exerts a statistically significant effect, with the older cohort of children favouring the use of the CHP more than the younger age group. This comparison of two relatively compressed age cohorts reveals that the CHP is a discourse feature whose rate of usage appears to increase as the children mature. This acceleration in usage may conceivably be linked to a greater ability with age to manipulate structural devices to perform the discursive functions of foregrounding information in narrative. Berman and Slobin (1994: 600) claim that tense-aspect forms acquire broader and more discourse-motivated functions across different developmental phases, and that knowledge of both the structure and use of linguistic forms is intimately related to the growth in specific narrative skills.

In terms of internal linguistic factors, the tense form used in the preceding complicating action clause is the only factor which exerts a statistically significant effect on all three tense forms. Moreover, the magnitude of this effect is particularly salient. The factor weights within the constraint hierarchy clearly highlight the potency of the discourse-level tendency for similar forms to cluster together. With a contribution of .94, the present perfect is clearly favoured in contexts where it is preceded by another present perfect form.<sup>19</sup>

The results relating to the distribution of tense forms in specific locations within the sequence of complicating action clauses also reveal a number of noteworthy findings. The factor weights for the preterite shows that whereas this form is favoured in initial and final clauses, it is disfavoured in medial clauses. By contrast, medial clauses are comparatively more favourable to the use of the CHP. Although the results for the present perfect are not statistically significant, there is a mild

---

<sup>19</sup> Cox (2005: 89) similarly points out in her New Zealand data that out of all the verbs that had a present perfect used in a preterite-like function as their immediately preceding narrative verb, nearly 75% were present perfect forms.

tendency for this form to appear in medial clauses more so than in initial and final clauses. Thus, even within the tightly circumscribed domain of complicating action clauses, tense forms are differentially distributed, which underscores the importance of discourse-level constraints on tense-switching phenomena in the data.

Turning to the use of tense forms in quotative and non-quotative contexts, there is evidence of an association between specific tense forms and particular discourse-pragmatic functions. As indicated by the range in the results above, there is a particularly strong relationship between the CHP and lexical quotatives (e.g. *say* and *go*) used to frame direct speech in the preadolescent corpus. In contrast, the preterite is disfavoured in quotative contexts.

The factor weights for grammatical person for each tense form provide additional evidence of distinct tendencies exhibited by the variants. Grammatical asymmetries are most apparent between the preterite and the CHP: whereas the preterite is favoured with the first person, the CHP is favoured with the third person. There is much less fluctuation in the factor weights for the present perfect, although it is notable that the tendency for this form to occur with the first person is consonant with historical trends (see Elsness 1997: 197).

## 5.7 SUMMARY

The multivariate analysis of factors which constrain tense alternation patterns in the data reveal that there are important areas of divergence between tense forms in foregrounded narrative sequences. In particular, a factor-by-factor analysis points to distinct functional tendencies exhibited by the CHP and the preterite. Moreover, the data also indicate that there are differences in the environmental constraints on the occurrence of the present perfect and the preterite, especially with regard to their

tendency to surface in particular narrative positions in complicating action sequences, suggesting that these two tenses are not functionally isomorphic.

The distribution of the present perfect within the tense cohort used to structure foregrounded narrative events suggests that what we are dealing with here is primarily a pragmatic phenomenon rather than a grammaticalised substitute for the simple past. Like the CHP, it can be facultatively employed to help demarcate narrative scenes or episodes. A switch into the present perfect, like a switch into the CHP, constitutes an 'interpretive move' (Fludernik 1991: 30) which helps to set off events from each other and facilitates the internal evaluation of narrative episodes (see Silva-Corvalán 1983: 778). In brief, it appears, then, that alternations between the preterite, the CHP and the present perfect are used in the preadolescent narratives examined here to create a nexus of stylistically meaningful oppositions which speakers can exploit discursively for segmentational purposes and to signal different degrees of event saliency in their narrative recounts.

Clearly, a larger corpus incorporating a broader range of speakers and age ranges is desirable, particularly in order to investigate in more detail the social distribution of narrative uses of the present perfect. Using limited evidence from adult corpora, I have, however, tentatively proposed that the present perfect plays a similar discourse-structuring role in the narrative of older speakers, not only in British English, but also in New Zealand and Australian English.

The cross-variety evidence that I have drawn on in this chapter to examine discourse-pragmatic functions of the present perfect in narrative discourse suggests that the use of the present perfect in preadolescent narratives cannot be unequivocally

attributed to developmental or age-graded factors.<sup>20</sup> Nevertheless, in the absence of large socially stratified corpora of spoken vernacular data from British regional dialects, especially from the south of England, my interpretation of the role of the present perfect in preadolescent narratives must necessarily remain provisional.

The results generated within the confines of this small-scale study open up several other avenues of investigation which are worth pursuing from a cross-variety perspective. Firstly, as Fleischman (1985: 8) points out, there is no reason to suspect that discourse phenomena of the type examined in this chapter are consistently mapped on to the same grammatical categories cross-linguistically. From a cross-variety perspective, it would also be profitable to examine the extent to which tense-switching phenomena are differentially configured across varieties of English. Judging from the evidence that I have briefly reviewed above, the parallelisms between the British preadolescent data, the New Zealand data discussed by Cox (2005), and the narrative uses of the present perfect in Australian English investigated

---

<sup>20</sup> Hickmann (1995: 213) notes that evidence from a range of languages has been used to propose a defective tense hypothesis according to which children's verbal inflections do not initially encode tense, but mark viewpoint aspect (perfectivity) and/or situation aspect (results). However, most of this work is based on children in early rather than later childhood. Furthermore, as Hickmann (1995: 213-4) points out, only a minority of developmental studies examine how children deploy tense-aspect categories more globally to organise discourse. Those studies which do examine the role of tense and aspect in the development of discourse cohesion suggest that there are interesting differences between adults and children. For example, Bamberg and Marchman (1990) discuss how children make greater use of overt inceptive aspect (e.g. *he started to*) than adults to mark the initiation of event sequences. With regard to tense alternation, however, Hickmann (1995: 215) claims that children are able to tense-shift from the non-past to the perfective from the age of six to seven onwards in order to signal discourse relations associated with establishing the setting and the background in a narrative, as well as to discriminate overlapping events in the unfolding of a narrative. Also of relevance here is Bamberg's (1990) study of how German-speaking adults switch to the *Perfekt* (e.g. *er ist die Strasse entlang gegangen* 'he has gone/he went along the street') to indicate whether information is backgrounded or foregrounded in discourse, as well as to segment discourse into units. The acquisition of shifts characteristic of adult usage of the *Perfekt* emerge in later childhood (9-10 years), whereas younger children (3-4 years) use the *Perfekt* in more restricted ways to mark completion. While these studies do not necessarily disconfirm that children's use of the present perfect examined in this chapter is a developmental stage, they do provide some support for the view that children in later childhood have a more functionally differentiated control of tense-aspect categories than younger children, and are able to use tense alternations in adult-like ways to mark specific discourse relations and to establish discourse cohesion.

by Engel and Ritz (2000) constitute fertile territory for exploring possible convergence between varieties in discourse-pragmatic uses of the present perfect.

Within the broader domain of linguistic change, it would also be instructive to examine variation in the use of the present perfect with regard to the ongoing grammaticalisation of periphrastic constructions used to encode tense-aspect distinctions in contemporary varieties of English (see Mair and Hundt 1995; Elsness 1997: 347).<sup>21</sup> Of particular interest in connection with the evolution of the present perfect in varieties of English is whether discourse uses of this tense form in colloquial British, New Zealand and Australian English offer incipient indications of the grammaticalisation of evolving preterite-like functions.

Further exploration of these issues would evidently necessitate the application of historical and variationist methodologies in order to locate different varieties along the continuum of grammatical change (see Tagliamonte 2002a).

---

<sup>21</sup> As part of more detailed investigations into the grammatical evolution of the present perfect in contemporary English, it is also important in future research to pay greater attention to the lexico-semantic properties of the contexts in which tense variants appear. This would entail noting the lexical identity of each verb separately in order to ascertain whether there are associations between specific verbs and morphological exponents of past temporal reference (see e.g. Poplack and Tagliamonte 2001), as well as performing more detailed qualitative analyses of the relationship between the semantics of particular verb forms and temporal-aspectual markings. Cox (2005: 86) claims that if the present perfect has undergone a change that allows it to be used with preterite-like functions (i.e. to relate a sequence of narrative events), then we might expect to find a change in the aspectual classes with which the present perfect can be used to report an event. Thus, in Cox's (2005: 85) examination of preterite-like uses of the present perfect in New Zealand English, each present perfect form was coded as either an accomplishment, achievement, process, semelfactive or state.



## CHAPTER 6

# VARIATION IN THE USE OF DISCOURSE MARKER *LIKE*

### 6.1 INTRODUCTION

The rapid dissemination of discourse marker *like* in several contemporary varieties of English is frequently cited as an example of a burgeoning change in progress (see e.g. Dailey-O'Cain 2000; Andersen 2001; Tagliamonte 2005; D'Arcy 2005a; Cheshire *et al.* 2005a).<sup>1</sup>

Research on sociolinguistic variation in the use of discourse markers such as *like* was neglected until relatively recently (Ferrara 1997: 376). One of the reasons why there have been few rigorous quantitative studies of *like* is because it has frequently been viewed as used on an *ad hoc* or unsystematic basis. Siegel (2002: 64), for example, comments on the ability of *like* to 'occur grammatically anywhere in a sentence,' and Fleischman (1999) similarly notes that *like* can be 'inserted virtually anywhere in an utterance.' Nevertheless, Ferrara (1997: 366) states that if discourse markers 'conform to overall principles of language' they may be expected to 'exhibit orderly heterogeneity.' Traugott (1995) argues that discourse markers have highly constrained syntactic properties, and Brinton (2006: 308) claims that, like other components of grammatical

---

<sup>1</sup> Discourse markers are also referred to in the literature as pragmatic particles (Östman 1995) and pragmatic operators (Ariel 1994). The number of different designations for these forms reflects the lack of a widely accepted definition of what constitutes a discourse marker. Several researchers resort to providing descriptions of the properties exhibited by 'prototypical discourse markers' (Jucker 2002: 211). Descriptions of the key functional properties of discourse markers frequently include reference to their role in discourse connectivity; their syntactic optionality; their non-truth-conditionality; their weak clause association; and their primary, but by no means exclusive, association with speech (see Brinton 1996: 30-31; Schourup 1999: 230-4)

structure, their evolution in a range of languages shows evidence of ‘many of the morphosyntactic and semantic changes associated with grammaticalisation.’

Recent research by D’Arcy (2005a) on the sociolinguistic distribution and syntax of discourse *like* in the speech of English-speaking Canadians in Toronto indicates that *like* is far from random and is subject to internal grammatical constraints. To date, however, there have been few detailed quantitative studies of discourse *like* in British English (although see Andersen 2001; and Cheshire *et al.* 2005a).

In this chapter, I explore the sociolinguistic distribution of discourse *like* in the preadolescent corpus, and aim to contribute to the growing body of research findings indicating that this marker is not a grammatically unconstrained ‘filler.’ I present evidence challenging popular assumptions that *like* is a mark of inarticulacy or word-finding difficulties, and argue that in the vernacular of young speakers, it is used to fulfil a number of discourse-pragmatic functions, including marking relations between sequentially dependent units of discourse (Schiffrin 1987; Traugott 1995).

Examples of the use of discourse *like* taken from the preadolescent corpus are given below:

- (1) it looked all scary and people kept jumping *like* around [2M10/11]
- (2) cos *like* if he misses a shot, he starts punching people [4M10/11]
- (3) I got *like* a scar under my eyebrow but you can’t really see it [12M10/11]
- (4) there’s one minute *like* ..... *like* they’re sometimes they’re (*sic*) adults and sometimes they’re children [7F10/11]
- (5) they was *like* on this boat [15M7/8]
- (6) but then they found all *like* dinosaurs and all that [13M7/8]

These examples illustrate some of the functions of *like* which have variously been explained in terms of focus marking, hedging, and exemplification, as well as having been associated with hesitation phenomena, and word-finding difficulties (see Underhill 1988; Miller and Weinert 1995; and Andersen 2001).<sup>2</sup>

The plurality of discourse functions encompassed by *like* are not exclusive to this marker, but are similar in several respects to those of other expressions which have discourse marking roles. Thus, Hansen (1998) notes that not only do discourse markers frequently co-occur, they also often functionally overlap with one another (see Chapter 8). Furthermore, the multifunctionality of such markers means that they are typically used across a range of communicative domains including information management; speech management; textual organization; and turn-exchange systems, in addition to serving interpersonal functions relating to the expression of speaker attitude and tentativeness (see Fischer 2000; Brinton 2006).

In addition to the interest which the discourse-marking functions of *like* have generated, scholarly attention has focused on the related use of *be like* as a quotative marker used to foreground reported speech and thought, as well as non-lexicalized speech sounds (see Romaine and Lange 1991; Ferrara and Bell 1995; Tagliamonte and Hudson 1999; Tagliamonte and D'Arcy 2004). However, in keeping with previous variationist studies (see e.g. Tagliamonte and D'Arcy 2004; Tagliamonte 2005; D'Arcy 2005a), I treat this construction separately in Chapter 7, which focuses on variation in the quotative system.

---

<sup>2</sup> The rapid diffusion of discourse *like* has garnered considerable attention from the press. One British newspaper article (Robin Young, *The Times*, March 24, 2004) features *like* in a list of the top ten 'linguistic irritants' compiled by the Plain English Campaign. Cameron (2000: 131-2) notes that in the American press, similar complaints have appeared proscribing the discourse use of *like* on the basis that it is an unnecessary filler.

Previous research relating to the apparent increase in the use of discourse *like* has highlighted age and gender as important social parameters influencing the trajectory of change in the use of this feature in several vernacular varieties of English. Studies based on British, American and Canadian English have recurrently cited teenagers as being in the vanguard of change with regard to the dissemination of this feature (see e.g. Andersen 2001; Buchstaller 2001; Tagliamonte 2005). However, the focus on adolescent speakers has meant that less is known about the extent to which this variable is used and how it is acquired by younger speakers. As mentioned at the beginning of this study, it is often the case that when corpora do include data from younger speakers, the distinction between adolescents and preadolescents is not always systematically made, which makes it difficult to ascertain how particular changes may be diffusing to and are sustained by younger age cohorts (D'Arcy 2005b: 332).

In research conducted on British English, Miller and Weinert (1998: 320) make the claim that *like* appears to be acquired late, after the age of 10, on the grounds that it is infrequent in dialogues from the younger participants who took part in their study. Yule (2000: 378) states that this type of claim constitutes a hypothesis that could be tested by further research.

By presenting a quantitative analysis of the syntactic and pragmatic aspects of *like* in the speech of preadolescents, I specifically reassess and challenge the claims made by Miller and Weinert (1998) that discourse *like* is not acquired until early adolescence. I argue that not only are discourse uses of *like* frequently encountered in the vernacular of children below the age of 10, but that distributional differences in the use of this marker by males and females, as well as by different age divisions in preadolescence, also highlight important facets of its acquisition.

Before examining the distribution of discourse *like* in preadolescent speech, I widen the perspective of the analysis by first briefly considering cross-linguistic comparisons with discourse markers that have a similar functional inventory to *like*, and then reviewing the historical background associated with the evolution of discourse *like* in the English language.

## 6.2 CROSS-LINGUISTIC PERSPECTIVES ON THE GRAMMATICALISATION OF *LIKE*-LEXEMES IN DISCOURSE MARKING FUNCTIONS

There is accumulating evidence (Andersen 1997; Meyerhoff and Niedzielski 1998; Fleischman 1999; Maschler 2002) of analogous discourse uses of *like*-lexemes in a range of languages. These cross-linguistic analogues exhibit striking parallels: they can be used in a variety of syntactic positions to sequence chunks of discourse, to hedge propositions, mark focus, as well as, in some cases, foreground reported speech or thought (see Chapter 7). According to Meyerhoff (2002: 354) innovations with *like*-lexemes are highly functional, as the examples from a range of languages illustrate below:

Canadian French *comme* (*like, as*)

- (7) *Comme* au début tous les bureaux même pour les acheteurs c'était dans le magasin là-bas

'*Like* in the beginning all the offices even for the buyers they were in the store there' (Sankoff *et al.* 1997: 205 )

- (8) Hebrew *kaze* (*like [this]* )

veke'ilu haragláyim sh'xa nitka'ot bifním *kaze*

'and *like* your feet get stuck inside *like*'

(Maschler 2001: 296)

- (9) Bislama *olsem* (*like, as*)<sup>3</sup>  
 afta *olsem* hem i jas kambak, be afta i *olsem* long i tudak, eh  
 ‘and *like* he’d just come home, and it was *like* it had gotten dark eh’  
 (Meyerhoff and Niedzieslki 1998 : 239)
- (10) Finnish *niinku* (< *niin kuin as if*)  
 Ja sit mä olin *niinku* että herrajjumala et voi olla totta  
 ‘and then I was *like* oh my God, I can’t believe it’  
 (Fleischmann 1999)
- (11) Swedish *liksom* (< *lik + som like + as*)  
 å så tar han en *liksom* pizzasallad till midda  
 ‘and then he takes a *like* pizza salad for dinner’  
 (Kotsinas 1994: 87)
- (12) Italian *tipo (che)* (*type, kind, sort*)  
 cioè meno anni di Michele, *tipo che* sta con una che, che ha vent’anni più  
 di me  
 ‘that is, younger than Michele, *like* he’s got one [a woman] who’s twenty  
 years older than me’ (De Mauro *et al.* 1993, cited in Fleischman 1999 )

In addition to the languages exemplified in (7)-(12), Fleischman (1999) mentions that a variety of discourse functions associated with English *like* are encountered in German, Japanese, Lahu, Portuguese, and Russian. The fact that counterparts to *like* exist in a range of unrelated languages and have arisen from similar semantic sources suggests that the existence of these discourse markers cannot be easily ascribed to borrowing processes or common genetic inheritance; rather, they appear to have their

---

<sup>3</sup> Bislama is a Melanesian creole spoken in Vanuatu, and is related to Tok Pisin and Solomon Islands Pijin.

roots in fundamental communicative functions such as the construction of coherent discourse, utterance interpretation, and the negotiation of social meaning.

Fleischman (1999) observes that functional similarities between discourse markers raise important questions about diachronic parallelism. The possibility that discourse markers such as *like* (and its cross-linguistic analogues) have their origins in underlying principles of conceptualizing particular semantic domains which motivate trajectories of grammaticalisation has been the subject of recent discussion (see for example Heine and Kuteva 2002). Traugott and Dasher (2002: 156) note that although discourse markers can be highly language-specific in terms of their distribution and pragmatic function, there nevertheless 'seem to be quite similar paths of development at the macro-level', with the grammatical evolution of discourse markers proceeding along 'correlated diachronic continua' (Traugott 1995).

### 6.3 THE EVOLUTION OF *LIKE* AS A DISCOURSE MARKER

Bybee (1988: 351) argues that 'synchronic states must be understood in terms of the set of factors that create them.' Thus, in order to situate variation in the use of *like* within the broader perspective of possible ongoing change, I provide a synopsis of previous research on the grammaticalisation of *like* as a discourse marker, and explore the possible existence of historical precursors of this marker in older varieties of English.

#### 6.3.1 *The development of LIKE viewed from the perspective of grammaticalisation*

Romaine and Lange (1991: 261) claim that the syntactic functions of *like* as a preposition and a suffix (e.g. 'he looks *like* his brother', 'it was a tiny worm-*like* creature') provide the grammatical context for the extension of *like* to new focusing

functions in discourse. From a historical perspective, the use of *like* as a preposition and conjunction can be etymologically traced to Old English 'gelic' (< 'ga' corresponding to Latin 'com' + \*liko meaning 'body, form' originally meaning 'having the form of'; see Meehan 1991: 4).

In terms of its pathway of change, a comparison of *like* in its established grammatical contexts (e.g. as a preposition, conjunction and suffix derived historically from an earlier form with lexical content) with its more recently evolved discourse-marking functions highlights a number of distinctive features associated with the process of grammaticalisation (see Hopper and Traugott 1993).

If we concede that Romaine and Lange (1991:261) are correct in identifying prepositional uses of *like* as a possible point of grammatical origin for its later discourse functions, then its additional use as a conjunction, preceding its recent evolution as a discourse marker, not only exemplifies 'layering' within the grammar (Hopper 1991), but also illuminates a common semantic-pragmatic pathway of development associated with grammaticalising forms: namely, a shift from the propositional component (preposition), to the textual component (conjunction), to the interpersonal or expressive component (discourse marker) (see Traugott 1982: 255; Brinton 2006: 311). The precise nature of how *like* evolved multiple discourse-marking functions from earlier propositional uses remains, however, unresolved. Romaine and Lange (1991: 262) claim that the evolution of discourse and quotative uses of *like* from its earlier non-discourse uses cannot be accounted for by a simple linear model of grammaticalisation, and may instead involve 'a network of related meanings'.<sup>4</sup>

---

<sup>4</sup> See also Buchstaller (2001: 23), who proposes a radial model organised around a core meaning of comparison/resemblance to account for the relationship between the functional domains of discourse *like*.



### 6.3.2 *LIKE* in older varieties of English and regional dialects

From a diachronic perspective, the extent to which discourse-pragmatic uses of *like* can be traced back to patterns of variation in older varieties of English is an area which merits further research. Andersen (2001: 209), for example, notes that *like* has a history as a discourse marker in the traditional dialects of England. Examples of historical usage taken from Wright (1902) and the Oxford English Dictionary are given below:

- (13) Father grew quite uneasy, *like*, for fear of his Lordship's taking offence  
(1778 F. Burney *Evelina* II xxiii.222)
- (14) If your honour were more amongst us, there might be more discipline  
*like* (1838 Lyton *Alice* II. iii)
- (15) Why *like*, it's gaily nigh *like* to four mile *like* (1840-41 De Quincey  
Style II Wks 1862 X.224)
- (16) He is all mwopen *like* (Dorset / Barnes Gl, 1863)
- (17) He would not go *like* through that. They are *like* against one another as it  
is (Mid-Yorkshire)

[Examples 13-15 from OED; 16-17 from Wright 1902]

Andersen (2001: 222) draws a distinction between 'traditional' British uses of *like*, which occur clause-finally, and *like* in other 'non-traditional' syntactic positions in the *Bergen Corpus of London Teenage Language* (COLT) on which he bases his analysis of contemporary usage. According to Romaine and Lange (1991: 249), there is a major difference between British and American English with regard to the syntactic position of *like*: in American English, *like* appears before the phrase or constituent that it is intended to qualify, whereas in British English it is placed afterwards.

Andersen (2001:209) attributes the non-traditional uses of *like* in COLT to the influence of American English (see Underhill 1988) in which *like* occurs predominantly in non-clause final positions.

However, the assumption that non-traditional uses of *like* in contemporary varieties of British English should unequivocally be ascribed to the influence of American English is problematic for a number of reasons. Firstly, the historical record shows evidence of the use of *like* in non-clause final positions in traditional British English dialects (see example 17 above). Such examples may be construed as ‘embryonic variants’ (see Trudgill 2002: 41 ), possibly representing early attestations of clause-internal *like* which foreshadow its discourse-pragmatic uses in contemporary British vernaculars. Furthermore, synchronic evidence from British dialect speakers shows that non-traditional uses of *like* are productive in the speech of members of the oldest generation, suggesting that clause-internal variants have been embedded in vernacular speech for quite some time. D’Arcy (2005a: 6) cites the following examples from older speakers of regional British dialects discussed in Tagliamonte (*to appear*).

- (18) They were just *like* sitting waiting to die (Ayrshire, lowlands of Scotland)
- (19) We were *like* ready to *like* mutiny (York, northeast England)
- (20) I couldn’t stand it, *like* I just couldn’t (Maryport, northwest England)
- (21) We were doing *like* a nature study (Portavogie, Northern Ireland)

In other words, then, taking into account diachronic and synchronic evidence from regional British dialects, non-traditional uses of *like* (i.e. in clause-initial and clause-internal positions) in British English are not necessarily attributable to contact

with American English, but may have arisen (at least partially) by independent, parallel development (see also Meyerhoff and Niedzielski 2003: 545).

#### 6.4 PREVIOUS APPROACHES TO THE STUDY OF *LIKE* IN CONTEMPORARY VARIETIES OF ENGLISH

Although most of the previous research (e.g. Jucker and Smith 1998; Andersen 2001; Fuller 2003) discusses the pragmatic uses of *like* as being characteristic of discourse markers, *like* in the preadolescent corpus, and in other data sets, differs from prototypical discourse markers in that it does not invariably occur clause-initially, but can additionally occupy a number of intonationally integrated clause-internal positions.<sup>5</sup> Some researchers have consequently drawn a distinction between *like* used as a discourse marker in clause-initial position, and *like* as a discourse particle used clause-internally (see e.g. D'Arcy 2005a: 52). Thus, according to this categorisation, *like* in example (2) above, counts as a discourse marker, whereas *like* in example (3) counts as a discourse particle.

Yet another difficulty in categorising *like* as a discourse marker or particle arises from problematic contexts where *like* appears to contribute to the propositional content of an utterance (Andersen 2001: 260; Siegel 2002: 36; D'Arcy 2005a). Discourse markers are typically considered not to affect the truth-conditions of the proposition expressed by the utterance in which they occur (Schourup 1999: 232). However, consider the following example taken from the preadolescent corpus, where *like* is close in meaning to 'roughly', 'approximately':

- (22) and once you've beat *like* five yeah then you can get your place in  
history in the game [20M7/8]

---

<sup>5</sup> Hansen (1997: 156), for example, notes that discourse markers 'prototypically introduce the discourse segments they mark'. However, Schourup (1999: 233) notes that clause initiality is not necessarily considered criterial for discourse marker status (see also Brinton 2006: 308).

Andersen (2001: 260) argues that although 'roughly'/'approximately' are both truth-conditional adverbials, *like* in such contexts functions metalinguistically and is used to indicate that 'the utterance contains a loose interpretation of the speaker's thought.' The notion that speakers can use *like* to express a certain psychological distance from an expression is prominent in relevance-theoretic approaches to the study of *like* (see Andersen 1997, 2000, 2001). Within this framework, discourse markers (or what Andersen 2001 calls 'pragmatic markers') play a crucial role in guiding the hearer in utterance interpretation by 'constraining the selection of contextual background against which an utterance is to be interpreted' (Andersen 2001: 41). According to Andersen (2001: 230), *like* is a marker whose principal contribution to utterance meaning is to signal a relation of non-identical resemblance between an utterance and its underlying thought.

On the other hand, D'Arcy (2005a) argues that approximative uses of *like*, as in (22) above, cannot simply be categorised with its other discourse uses in a unitary analysis. D'Arcy (2005a: 9-10) argues that in contexts where *like* precedes numerical expressions, its status is controversial as it appears to function adverbially and competes with other expressions of approximation such as *about*. On the basis of an apparent-time analysis of a large corpus of Canadian speech, D'Arcy (2005a: 42) advances the claim that *like* is replacing the traditional adverb *about* in the speech of the youngest generation (10-16) of Canadians where it is becoming the preferred approximative adverb in numerical contexts. If, as D'Arcy (2005a) argues, the use of *like* in quantifying expressions is similar to that of other approximative adverbs, then its potential interaction with truth-conditional meaning effectively debar it from having discourse marker status in such environments because discourse markers

characteristically specify a procedural rather than a conceptual relationship between units of discourse (Fraser 1999: 944).

Others have argued that the use of *like* as a marker of approximation that indicates looseness of meaning has paved the way for new discourse-pragmatic functions. Fuller (2003: 369), for instance, claims that the use of *like* as a marker of approximation has given rise to its re-analysis as a focus marker, since information that is qualified is also the focus of an utterance. The persistence of the core meaning of 'similarity' and 'comparison' associated with *like* is strongly implicated in its metaphorical extension into discourse contexts associated with focus marking. Haiman (1988: 310), for example, notes a universal link between comparison and focus, with comparative constructions contrasting and hence focusing the elements compared.

## 6.5 THE SOCIOLINGUISTIC DISTRIBUTION OF *LIKE* IN CONTEMPORARY VARIETIES OF ENGLISH

### 6.5.1 *Age*

As already indicated, age figures prominently in discussions about the frequency and spread of *like* (see e.g. Romaine and Lange 1991; Ferrara and Bell 1995; Dailey-O'Cain 2000; Macaulay 2005; D'Arcy 2005a; Tagliamonte 2005). Cheshire *et al.* (2005a: 154) note that the spread of *like* in urban British vernaculars has been associated with a general youth culture. It is interesting that similar findings have been reported for other languages in relation to the use of discourse markers that exhibit functional parallels with *like*. Maschler (2002: 245) observes that the use of the discourse markers *ke'ilu* ('like', lit. 'as if') and *kaze* ('like', lit. 'like this') by young Israelis has led to this age cohort being stereotyped 'the *kaze ke'ilu* generation.'

Norrby and Wirdenäs (2003: 265) claim that Swedish adolescents' use of discourse markers such as *liksom* (cognate with English *like*) and *ba* (English 'just') are a 'resource for performing a joint discourse identity as young people'. Discourse markers such as *like* (and its cross-linguistic counterparts) constitute a potential pragmatic resource on which adolescents can draw to make sociosymbolic 'declarations of youth' (Chambers 2003).

Cheshire (2002b: 27) raises the issue of whether *like* is an age-graded feature, or represents an example of a more permanent linguistic change. Tagliamonte (2005) offers some purchase on this question in her study of the use of *like* in the Canadian English of Toronto youth. According to Tagliamonte (2005), differences in the frequency of *like* usage by different age groups in her data indicate a trend which resembles age-grading, with 15-16 year olds making the greatest use of *like* contrasting with the more conservative usage of 10-12 year olds and 17-19 year olds. D'Arcy (2005a: 140, 209), however, argues that speakers of all ages share a similar grammar for *like*, which appears to have generalized across syntactic structure in a step-wise fashion in successive generations of speakers.

Of particular interest in D'Arcy's (2005a: 85-6) discussion of generational differences in the use of *like* is her finding that the youngest Canadian speakers (10-12) use *like* in clause-initial position somewhat more frequently than older speakers. As D'Arcy (2005a) points out, this finding is significant because it is at variance with Miller and Weinert's (1995: 380) claim that discourse *like* is acquired relatively late, and is not a productive feature of the language of preteen speakers. I return to the issue of the age of acquisition of *like* and preteen speakers' frequency of usage below.

### 6.5.2 Gender

Given that gender-differentiation in language variation is often taken to be diagnostic of change in progress, previous research has sought to explore the existence of a correlation between discourse *like* and speaker gender. Blyth *et al.* (1990: 223-4) report that in their survey based on speakers of American English, *like* was used more by men than by women, even though middle-class teenage girls were perceived to be the most frequent users of this feature.<sup>6</sup> Dailey-O'Cain (2000: 69) found that similar perceptions were evident in her study, although, once again, men were found to use *like* more frequently than women, but the difference in usage was not statistically significant. Fuller (2003: 372), on the other hand, found that *like* was predominantly a feature of female conversation, with female speakers using *like* at a rate of 11.3 times per 1000 words, while males used it substantially less at a rate of only 4.1 times per 1000 words. In his study of London adolescent usage, Andersen (2001: 287) reports that *like* has mainly been adopted by white female speakers in late adolescence.

The lack of a consensus in the sociolinguistic literature concerning the relationship between the use of *like* and speaker gender has led some researchers to conclude that there is no simple correlation of this marker with either male or female speech (see Eckert 2003: 395). The absence of such a clear-cut connection bears testimony to the fact that gender is a socially situated construct which does not have 'a uniform effect on linguistic behaviour for the community as a whole' (Eckert 1989:253). Moreover, as a discourse-pragmatic change in progress, *like* may be

---

<sup>6</sup> Macaulay (2005) also notes social class differences in the use of *like*. Macaulay (2005: 82) found that middle-class Glaswegian adolescents used *like* more than five times as often as working-class adolescents. Macaulay (2005: 82) suggests that *like* is an interesting discourse marker because it is unusual for a stigmatised feature to appear more frequently in middle-class than in working-class usage. Similarly Cheshire *et al.* (2005a: 154) report that in Hull, middle-class adolescents used *like* more than working-class adolescents. This social class difference did not obtain, however, in the speech of adolescents recorded in Reading and Milton Keynes.

differentially sensitive to gender depending on where it is in the time course of its evolution (see e.g. Cameron 2005a: 48). Most important of all, as I show below, broad patterns of distribution may conceal subtle gender differences that only come to light when specific environments of use are taken into consideration (see Stubbe and Holmes 1995: 72).

Before exploring the existence of sociolinguistic patterns in the preadolescent corpus, I first explain how the envelope of variation was delimited in order to carry out a systematic quantitative analysis of *like* in the data.

## 6.6 METHODOLOGY

### 6.6.1 *Deciding what to count*

According to Dubois and Sankoff (2001: 283), ‘it is not feasible to contrast the presence of a discourse form to its absence as is done in phonological studies.’<sup>7</sup> I therefore analyse the distribution of *like* in terms of relative number of occurrences, and report its differential distribution across age and gender cohorts in terms of frequency per 1,000 words (see Stubbe and Holmes 1995). This procedure provides a normalising measure for accountable comparisons to be made across studies (see also Macaulay 2002: 299, f.n. 3).

Each context where *like* had a discourse-pragmatic function was retained for analysis. Non-discourse uses of *like* as a verb, preposition, conjunction, suffix or straightforward adverbial were excluded from the quantitative analysis, as were anomalous cases, in accordance with standard sociolinguistic practice. Following

---

<sup>7</sup> D’Arcy (2005a) is an exception. After excluding a number of ambiguous contexts, as well as a number of syntactic constructions which do not form part of the variable context for *like*, D’Arcy (2005a) carried out a wide-ranging variationist analysis of the syntactic behaviour of *like* in the speech of Canadians aged between 10 and 87. D’Arcy’s (2005a) analysis is based on more than 20,000 contexts which include not only actual occurrences of *like*, but also structurally defined contexts where *like* did not occur, but potentially could have done. D’Arcy’s (2005a) syntactic analysis of *like* is embedded within Minimalist Theory.



Tagliamonte (2005) and D'Arcy (2005a), quotative uses of *like* were excluded from the analysis, as were contexts in which *like* appeared to interact with truth-conditional semantics when it was used with numerical or quantifying expressions. I also excluded tokens of *it's like/it was like* (e.g. *it was like this monster in the sea*) on the grounds that the frequency of this construction in the preadolescent data suggests that we are not dealing with a single lexeme, but rather with the incipient grammaticalisation of a morphosyntactic string which is becoming automated as a single processing unit (Bybee 2002: 603), and which has its own collocational patterns (Tagliamonte and D'Arcy 2004).<sup>8</sup> After these exclusions, a total of 419 tokens of *like* were left for analysis.

### 6.6.2 Coding for syntactic position

Previous research findings relating to the syntactic mobility and pragmatic functions of *like* were operationalised to draw up a coding protocol. Given the reported syntactic versatility of *like*, it was necessary to code for a wide variety of syntactic locations including clause-initial and clause-final positions, in addition to a range of clause-internal positions. The scope of syntactic positions that were accounted for in the coding protocol is illustrated in the examples below:

#### (23) Preceding (and within) a noun phrase

- a. I've got *like* these cushion chairs [7M10/11]
- b. she started hitting him with this *like* crowbar [8M10/11]

#### (24) Preceding an adjective phrase

he was *like* nice to all the rest [2M10/11]

---

<sup>8</sup> See also Andersen (1997: 41), who notes that '*it's like* operates as a fixed string and it can be considered as a marker slightly different from *like* alone.'

(25) Preceding an adverbial phrase

he wasn't shouting *like* a lot [9F10/11]

(26) Preceding a prepositional phrase

they was *like* on this boat [15M7/8]

(27) With verb phrases

a. all her skin was *like* peeling off [26F7/8]

b. she accidentally *like* pushed him [26F7/8]

c. cos then you have to *like* move on with your life [15F10/11]

(28) Clause-initial

*like* you've got some of your friends at that school [9F10/11]

(29) Clause-final

when they're children *like*... their children act as if they're the parents

[7F10/11]

In this analysis, I assume that clause-internal *like* has scope over projections directly to its right.<sup>9</sup> There is only one example in the preadolescent corpus where *like* occurs clause-finally (see example 29 above), where it is assumed that *like* has retroactive scope over the preceding clause (see Miller and Weinert 1995). In cases where *like* occurs clause-initially, I follow Siegel (2002: 42) in assuming that *like* has wide scope over the entire clause (see also D'Arcy 2005a: 119).

In order to investigate the popular association of *like* with hesitations and syntactic reformulations (see Miller and Weinert 1995: 372-4), I also coded occurrences where it directly preceded false starts, pauses, and aborted utterances, as in (30) below:

---

<sup>9</sup> I am using the term 'projection' here in the sense in which it is used in Radford (1997:524) to designate a constituent containing a head word.

(30) Preceding an aborted utterance

and that night she was meant to be coming round my house

and she was *like* ... [27F7/8]

### 6.6.3 Coding for pragmatic function

The coding of the pragmatic functions of *like* was more problematic, not least because discourse markers are notoriously multifunctional and can operate on several different planes of discourse simultaneously (see Schiffrin 1987: 61). Macaulay (2005: 9) similarly argues that individual uses of a particular discourse marker have the potential to be multifunctional, and decisions about which function a specific occurrence illustrates can often only be made with reference to the discourse context in which it occurs. As Soares da Silva (2006: 2189) observes, the functional polysemy of discourse markers is a perennial problem encountered in studies of these forms. For instance, a discourse marker which has an exemplifying function can additionally be used for focus. This complicates attempts to classify the functions of a marker on a discrete basis. Andersen (2001: 265) tries to address the problem by classifying particular functions of *like* on the basis of which function seems salient in a particular context. However, this methodology, as Andersen (2001: 265) points out, is not intended to impose any infallible categorisation on the pragmatic functions of *like*, but rather constitutes a broad, heuristic approach for exploring its range of uses.<sup>10</sup> As Cameron *et al.* (1989: 85) comment in their discourse analysis of tag questions, this means that a certain degree of arbitrariness is inevitable in any functional

---

<sup>10</sup> Similarly, the pragmatic coding procedure I use here is not intended to impose any definitive or conclusive interpretation on the pragmatic functions of *like*. Future work based on a larger number of tokens gathered from a range of corpora will no doubt provide more nuanced insights into the discourse-pragmatic functions of *like*.

classification because examples cannot always be unambiguously assigned to one category or another.

The pragmatic coding system used in this study is based on a modified version of the categories used by Andersen (2001: 266) to investigate the functional distribution of *like* in the COLT corpus. Three broad categories were used for classificatory purposes: exemplification, metalinguistic focus, and discourse connectivity. I illustrate each of these in turn.

In terms of the first category, *like* often occurs in an exemplifying function in the preadolescent data in conjunction with a general discourse extender (see Chapter 8), as in (31) and (32) below:

(31) there was *like* all rats and things running around [13M10/11]

(32) ... my bestest subject at school is *like* art and stuff [9F10/11]

In these two examples, *like* is used together with a general discourse extender (underlined in the examples above) to highlight an illustrative example of a more general case; in other words, it has category-implicative or set-marking functions in these contexts (see Dubois 1992: 181).<sup>11</sup>

The second category employed in the functional classification of *like* concerns its use for metalinguistic focus or hedging. In this case, *like* functions as an epistemic hedge indicating the degree of speaker commitment to an expression in its scope. According to Andersen (2001: 243), *like* in this function indicates a relation of 'non-equivalence between a chosen expression and a potential alternative one' so that the chosen expression has a similarity relationship to what it is intended to refer to (see Jucker and Smith 1998: 186). Consider example (33), in which a child is describing

---

<sup>11</sup> The co-occurrence of *like* with other discourse markers is widespread in the preadolescent data. The pervasiveness of these co-occurrences highlights Schiffrin's (1987: 66) point that in oral discourse, messages are often multiply reinforced.

a race in the snow, where *like* appears to signal that there is a relationship of non-equivalence between the expression 'glasses' and the speaker's intended meaning :

(33) ... we was like James Bond cos we had *like* glasses on [14M7/8]

With regard to discourse connectivity, *like* occurs with self-repairs and the elaboration of preceding remarks where it marks transitions from one phase of the discourse to another. It can provide what Ferrara (1997: 359) calls 'macrolevel organizational cues' where it acts as a resumptive signal after a momentary departure from the main discourse thread. It can also function as a discourse continuity link between propositionally unrelated elements. Thus, it tends to occur in transitional locations where speakers seek to refocus their interlocutor's attention. Examples of these discourse-linking functions are given below:

(34) they crashed into a plank *like* a tree [15M7/8]

(35) he's an ogre *like* a monster [16M7/8]

(36) so then *like* ... my mum *like* he called my mum and then they called  
an ambulance [19F10/11]

## 6.7 DISTRIBUTIONAL ANALYSIS

I now turn to a discussion of the distribution of *like* in the preadolescent data. The first thing to note is the frequency of *like* in the corpus. Figure 6.1 below gives a frequency breakdown according to age and gender.

The results displayed in Figure 6.1 contrast markedly with the paucity of *like* tokens for the 8 and 10 year olds discussed by Miller and Weinert (1995) in their analysis of data from Scotland. The results obtained by Miller and Weinert (1995) are reproduced in Table 6.1 for comparison.

FIGURE 6.1

FREQUENCY OF *LIKE* BY AGE AND GENDER

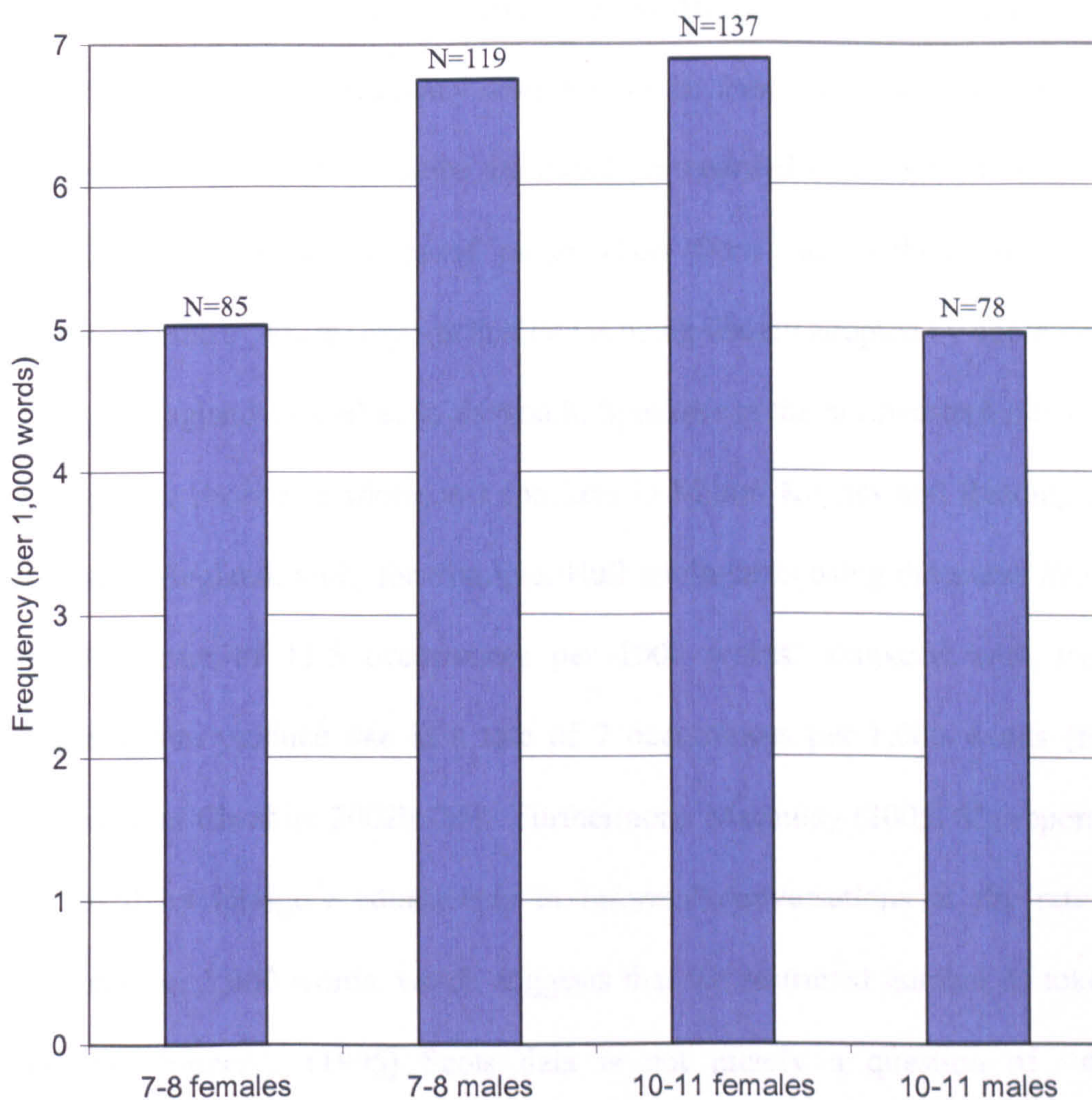


TABLE 6.1

FREQUENCY OF DISCOURSE *LIKE* BY AGE GROUP (ADAPTED FROM MILLER AND WEINERT 1995: 380)

Age Group	Occurrences of <i>like</i> (raw scores)
8 year olds	2
10 year olds	4
13 year olds	23

The difference between the frequency of *like* in the preadolescent corpus and the limited occurrence of *like* evident in Miller and Weinert's data set (1995) invites a number of explanations.<sup>12</sup> Firstly, differences in data elicitation techniques may be responsible for different frequency distributions in their data and mine. Secondly, Miller and Weinert's (1995) results are based on recorded conversations from Scots speakers, whereas mine are based on speakers from the southeast of England. However, Cheshire (2002b) reports that *like* is being widely adopted by adolescents in the north of England as well as in the south. Speakers in the northeastern city of Hull use *like* just as readily as adolescent speakers in Milton Keynes and Reading in the south-east of England, with, for example, Hull adolescents using discourse *like* at an approximate rate of 13.5 occurrences per 1000 words compared with Reading adolescents who produce *like* at a rate of 7 occurrences per 1,000 words (figures abstracted from Cheshire 2002b: 25). Furthermore, Macaulay (2005: 82) reports that *like* is used by Glasgow adolescents in informal conversations at the rate of 7 occurrences per 1,000 words, which suggests that the restricted number of tokens in Miller and Weinert's (1995) Scots data is not merely a question of dialect differences.

More important, Miller and Weinert (1995) base their discussion of *like* on data gathered in the 1980s.<sup>13</sup> Thus, higher rates of usage in my own data may reflect real-time changes in the spread and use of discourse *like* by younger speakers. In fact, there are several indications that this may well be the case. In the *Bergen Corpus of London Teenage Language* (COLT), recorded in 1993, Andersen (2001: 289) reports

---

<sup>12</sup> I do not have access to the total number of words in the corpus on which Miller and Weinert (1995) base their analysis, so I am unable to produce normalised frequency counts for their results.

<sup>13</sup> In future research on the spread of *like*, it will be important to consult more contemporary corpora such as the *SCOTS* corpus based at Glasgow University in order to examine how this discourse feature is used in Scotland, and to investigate more comprehensively how it has evolved over time in this geographical area.

that 10-13 year olds used *like* at a rate of 2.53 occurrences per 1,000 words, with speakers in late adolescence (17-19 ) year olds using it most frequently at a rate of 5.61 occurrences per 1,000 words. In other words, higher frequencies of *like* usage reported in Macaulay (2005), Cheshire (2002b), and my own preadolescent data may well be crude indicators of its ongoing grammaticalisation (Hopper and Traugott 1993: 103).<sup>14</sup> An increment in the frequency of an expression is considered to be one of the hallmarks of a grammaticalising change. Bybee (2003:602) argues that frequency is not just a result of grammaticalisation, it is also a primary contributor to the process.

In terms of the presence of social conditioning in the preadolescent data, there is no consistent pattern of usage according to gender as far as normalised frequency counts are concerned, nor is there any marked increase in the overall use of *like* between the two age groups. However, in apparent-time, we can see that by the age of 10, the girls have increased their use of *like* and have overtaken the 10-11 year old boys in terms of frequency of usage, whereas in the 7-8 age group, it is the boys who are the most frequent users of *like*. In the *Toronto Corpus of Youth English*, Tagliamonte (2005) found that gender differences in the youngest age cohort (10-12 year olds) were marginal, although the influence of gender on *like* usage was more pronounced in mid-adolescence, with peak usage found in the speech of 15-16 year old girls. Tagliamonte's (2005) findings suggest that gender differences associated with the use of *like* are developmental. I return to developmental issues implicated by the distribution of *like* in the preadolescent data in 6.7.2.

---

<sup>14</sup> It is difficult to specify a point in time when discourse *like* began to increase in frequency in Britain. Nevertheless, Cheshire *et al.* (2005a: 154) note that there is only one token of discourse *like* in 8,948 words recorded by Cheshire in adventure playgrounds in Reading in the late seventies. This contrasts with a rate of between 6 and 10 tokens of *like* per 1,000 words for Reading teenagers recorded in the mid to late nineties.



### 6.7.1 Variation in pragmatic function

In order to improve our understanding of the interrelationship of age and gender and the use of *like*, it is necessary to move beyond normalised frequency counts and examine more closely the broad range of pragmatic functions of *like* in the different age cohorts. Table 6.2 below gives a breakdown of the functional distribution of discourse marker *like* in the two age cohorts. The results in Table 6.2 show that both age cohorts have acquired a range of discourse-pragmatic functions for *like* which are in evidence by the time the children are seven years old.

TABLE 6.2

FUNCTIONAL DISTRIBUTION OF DISCOURSE *LIKE* BY AGE AND GENDER  
(UNBOLDED FIGURES ARE RAW SCORES)

Function	7-8 year olds				10-11 year olds			
	Male	Female	Total	% of Total	Male	Female	Total	% of Total
Exemplification	4	9	13	<b>6%</b>	10	24	34	<b>16%</b>
Metalinguistic Focus	65	44	109	<b>53%</b>	35	40	75	<b>35%</b>
Discourse connectivity	50	32	82	<b>40%</b>	33	73	106	<b>49%</b>
Total	119	85	204		78	137	215	

Discourse connectivity and metalinguistic focus are the most frequent functions in the two age cohorts, although the relative distribution of these differs proportionally according to age. In the older age cohort, *like* is used less often for metalinguistic focus than in the case of the 7-8 year olds. On the other hand, the older

children make greater use of *like* to fulfil discourse-linking and exemplifying functions than the younger cohort. The extent to which shifting frequencies in the use of *like* across different pragmatic functions reflects developmental differences in discourse management strategies is an area that warrants further research. Given the role of *like* in connecting sequences of discourse and marking the boundaries between them, it may be the case that the acquisition of its discourse-linking functions are dependent on the maturation of more general discourse abilities (for related discussion, see Berman and Slobin 1994: 600).

### 6.7.2 *Variation in syntactic position*

When the syntactic positions occupied by discourse *like* are taken into consideration, a more nuanced picture of differences in usage between the two age cohorts emerges. Table 6.3 below gives a breakdown of the syntactic positions of *like* in both age cohorts.

The first point to note is that there is only one instance in the entire corpus of *like* appearing in clause-final position, as mentioned earlier. I have already observed that, according to previous research, clause-final *like* is considered to be a feature of traditional British English dialects.<sup>15</sup> The question of whether clause-final *like* may be sensitive to regional variation, or other social parameters, is a matter for future investigation, but it is clear from the results in Table 6.3 that it is not a robust feature in the vernacular of the preadolescents I recorded.

---

<sup>15</sup> In particular, the use of clause-final *like* has been linked with northern British vernaculars (see e.g. Andersen 2001). However, even in the north of England, the findings are unclear. Cheshire *et al.* (2005a: 154) note that clause-final *like* is frequent in the speech of elderly speakers in Hull in the northeast, although it rarely occurs in young people's speech. Macaulay (2005:84) reports that working-class adults, mainly women, use clause-final *like* comparatively more frequently than adolescents in Glasgow, who hardly use it at all.

TABLE 6.3

SYNTACTIC POSITION OF *LIKE* (UNBOLDED FIGURES ARE RAW SCORES)

Syntactic Position	7-8 year olds				10-11 year olds			
	Male	Female	Total	% of Total	Male	Female	Total	% of Total
Noun phrase	50	21	71	<b>35%</b>	29	27	56	<b>26%</b>
Verb phrase	14	23	37	<b>18%</b>	16	33	49	<b>23%</b>
Adverb phrase	1	2	3	<b>1.5%</b>	1	2	3	<b>1%</b>
Adjective phrase	7	6	13	<b>6%</b>	2	3	5	<b>2%</b>
Prepositional phrase	7	2	9	<b>4%</b>	2	2	4	<b>2%</b>
Clause-initially	13	14	27	<b>13%</b>	14	42	56	<b>26%</b>
Clause- finally	0	0	0	<b>0%</b>	0	1	1	<b>0.5%</b>
Preceding either an aborted sentence (false start) or a pause	27	17	44	<b>22%</b>	14	27	41	<b>19%</b>
Total	119	85	204		78	137	215	

(Differences in the syntactic distribution of *like* between the two age-cohorts are statistically significant [occurrences of *like* preceding an aborted sentence or pause as well as *like* in clause-final position were excluded from the statistical analysis ],  $X^2 = 18.58$ ,  $df = 5$ ,  $p < 0.01$ )

On the other hand, the frequency of clause-initial and clause-internal uses of *like* in preadolescent speech is consistent with earlier research on urban adolescent vernaculars (see Andersen 2001), suggesting that preadolescents and adolescents are converging in their use of *like* in similar syntactic environments.

Another important point to emerge from the results displayed in Table 6.3 is that the frequency of *like* with aborted utterances or false starts is by no means the highest of the different distributions for both age cohorts. Contrary to popular associations of *like* with inarticulacy or conversational dysfluency, this finding indicates that *like* is not predominantly linked with poor syntactic planning or production problems (see also Miller and Weinert 1995: 372); in fact, the majority of occurrences of *like* are syntactically integrated.

Furthermore, the frequency distributions in Table 6.3 show that *like* is preferentially located in certain syntactic positions: before/within noun phrases, with verb phrases, and clause-initially.<sup>16</sup> These syntactic environments are the sites for 75% of the occurrences of all *like* tokens in the 10-11 age cohort, and 66% of all tokens in the 7-8 age cohort. Similar findings have been reported for other varieties of English. Tagliamonte (2005: 1901-2) observes that in the *Toronto Corpus of Youth English*, 30% of *like* tokens occur before a noun phrase, 23% occur at the beginning of a sentence, and 9% precede a verb, with these environments accounting in total for 62% of all occurrences of *like*. These findings indicate that *like* does not appear indiscriminately in discourse, and contrast with Siegel's (2002: 43) claim that *like* 'is best predicted not by structural or lexical factors, but by processing factors like lexical indecision.'

However, although it appears that certain syntactic environments favour *like* more than others, this assumption requires careful analysis and interpretation. As

---

<sup>16</sup> In recent syntactic frameworks, a noun phrase such as *a house* would be analysed as a nominal headed by the determiner *a* (i.e. a determiner phrase, DP). D'Arcy (2005a: 131) notes that out of 4000 DP contexts that she extracted from her Toronto data, only 47 contained *like* in a position other than on the periphery of the DP. Furthermore, examples such as *we had to go through these like slummy cities and stuff* (D'Arcy 2005a: 131) only appeared in the speech of speakers in their thirties, with further increments noted in the usage of younger speakers. In my much smaller preadolescent corpus, 13% of DPs contain occurrences of *like* following a determiner; e.g. *they put this like bandage stuff round it* [15F10/11]. Of crucial relevance here is D'Arcy's (2005a: 188) claim that discourse *like* first targets functional categories (i.e. the periphery of DP, preceding the determiner) before it spreads to lexical categories (i.e. directly preceding a noun).

D'Arcy (2005a: 23) observes, frequency calculations are often based on the occurrence of a discourse feature itself, rather than on an entire context which additionally takes into account where a discourse feature could have occurred but did not. When such adjustments are made, the frequency distribution of *like* in a given context (e.g. with noun phrases) may be markedly lower than frequency counts based on the presence of the discourse feature alone.

Closer examination of the syntactic environments in which *like* is used reveals that there are interesting gender- and age-related differences in the preadolescent data. The co-occurrence of *like* with noun phrases is more marked for the 7-8 year olds, where this environment accounts for 35% of all tokens in contrast with 26% for the 10-11 year olds. The 10-11 year olds, on the other hand, make greater use of *like* in clause-initial position, accounting for 26% of their total *like* tokens. This can be contrasted with the 7-8 year olds for whom clause-initial position represents only 13% of their total *like* usage. With increasing maturity, then, there is a proportional increment in the use of *like* in clause-initial position where it has scope over whole clauses.<sup>17</sup> Within the broader perspective of child language acquisition, it has been observed that changes in children's use of discourse markers correlate with scope expansion. For example, in his analysis of the use of several discourse markers in child language, Sprott (1992) reports that children used certain markers initially to mark local levels of discourse, but, with increasing age, new textual functions were acquired, with older children using discourse markers to indicate relationships at global levels of discourse that include larger units of talk. Similarly, Kyratzis and Ervin Tripp (1999: 1328) document age-related changes from local to global functions in children's (4-7 years) use of specific discourse markers. Developmental

---

<sup>17</sup> There are some interesting comparisons to be drawn here with diachronic change. Traugott and Dasher (2002: 40) describe the trajectory from 'scope within proposition' to 'scope over proposition' as an important step in the grammaticalisation of discourse markers in English.

shifts in discourse practices are not particularly surprising given that stylistic expansion is a concomitant of linguistic maturation and correlates with an increase in the use of devices for manipulating and structuring discourse (see Romaine 1984a).

Figures 6.2 and 6.3 below offer a proportional comparison by age and gender respectively of the occurrences of *like* preceding (and within) noun phrases and *like* in clause-initial position. The frequency of *like* with noun phrases peaks in the speech of the 7-8 year old boys. One of the reasons why *like* co-occurs with noun phrases is because it appears to have focus-marking functions which can be used to alert listeners to the introduction of new referents into discourse (see Underhill 1988: 234).<sup>18</sup> Evidence of focus-marking functions can be gleaned by considering the varying distribution of *like* with prenominal referential determiners (i.e. *a, the, this, that, etc.*): 72% of noun phrases preceded by *like* co-occur with a marker of indefiniteness (i.e. the indefinite article, or indefinite 'this'), which typically precedes the introduction of new referents.<sup>19</sup> *Like* may therefore be evolving functions which complement existing coding strategies in spoken English facilitating the identification and cognitive tracking of new information in discourse.

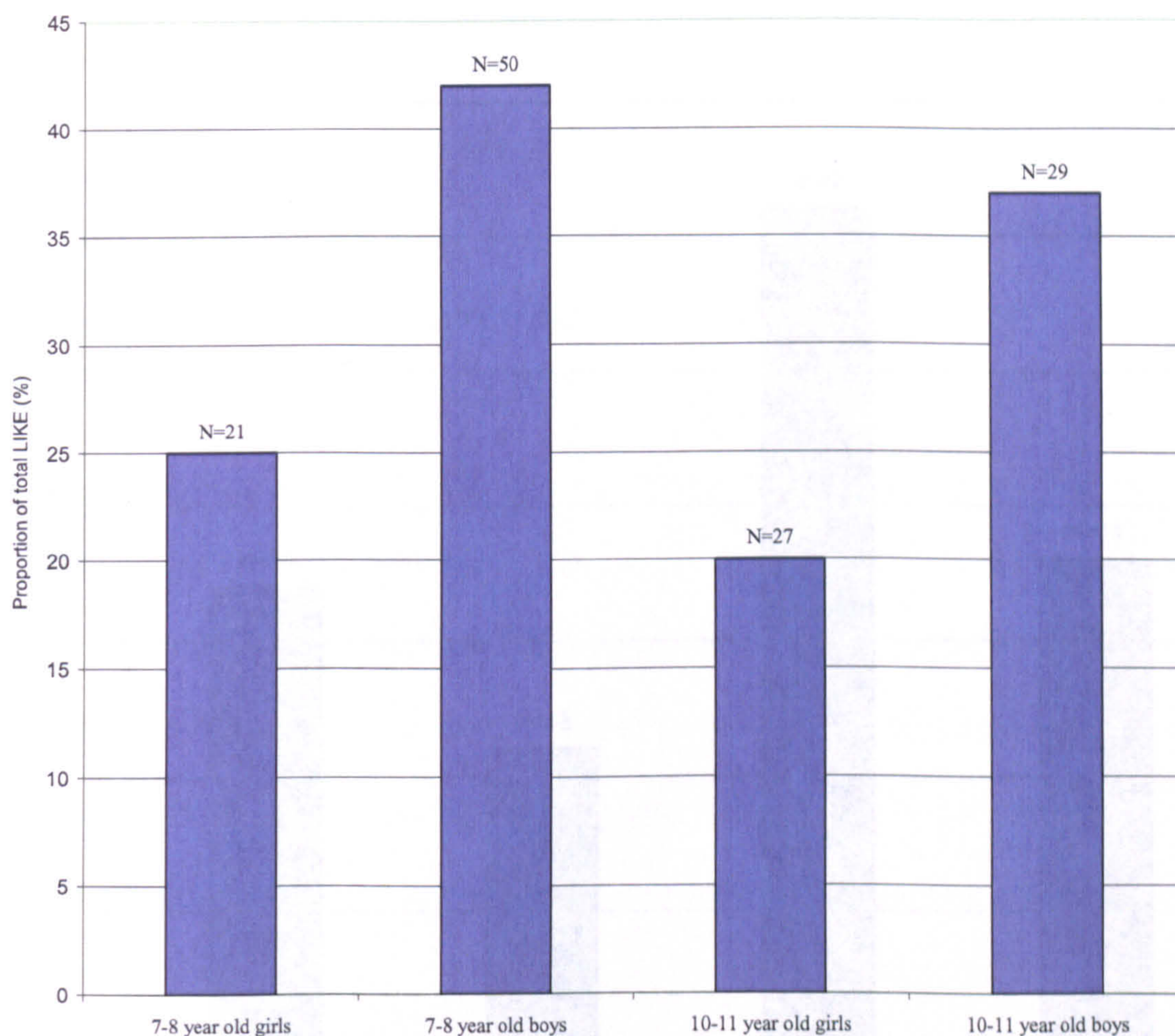
---

<sup>18</sup> Wennerstrom and Siegel (2003: 101) undertook a syntactic and prosodic analysis of discourse markers and report that *like* can occur turn-medially followed by a pause to form 'a jumping off point from which to launch a focused word.'

<sup>19</sup> Schiffrin (2006: 120) claims that although *this* can be used as a demonstrative, its status wavers between definite and indefinite. According to Cheshire (1989:54), *this* often exhibits addressee-oriented functions in narrative discourse, and can be used to give discourse prominence to referents.

FIGURE 6.2

COMPARISON BY AGE AND GENDER OF OCCURRENCES OF *LIKE* PRECEDING (AND WITHIN) NOUN PHRASES

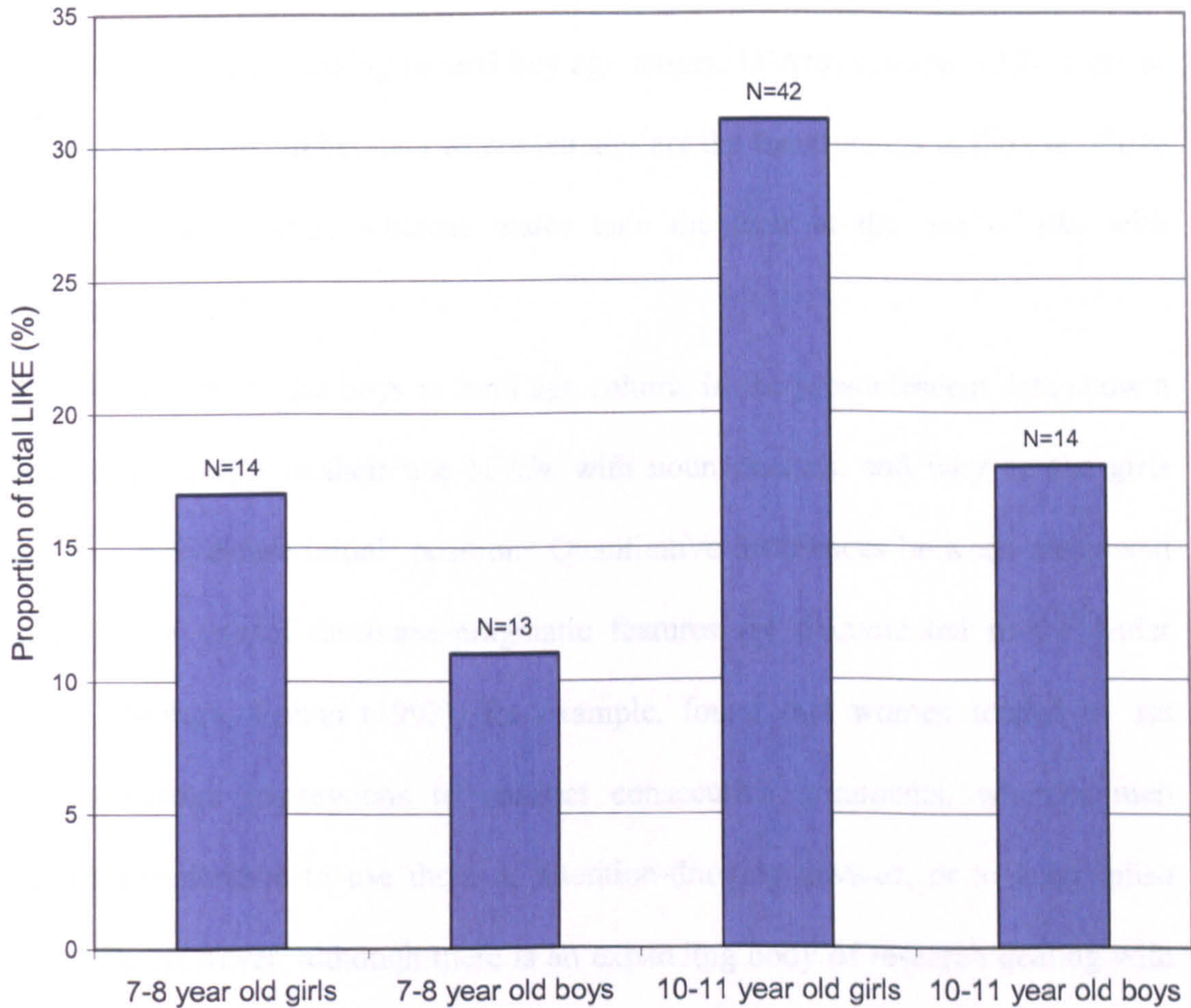


It is noteworthy that the co-occurrence of *like* with noun phrases is the preferred syntactic environment for *like* in both the speech of the 7-8 year old and the 10-11 year old boys where it accounts respectively for 42% and 37% of their total *like* usage. By contrast, the rate of co-occurrence of *like* with noun phrases is lower for the girls in both cohorts: 25% for the 7-8 year olds; and 20% for the 10-11 year olds.

Figure 6.3 below illustrates yet another dimension of variation in the use of *like* associated with gender.

FIGURE 6.3

COMPARISON BY AGE AND GENDER OF OCCURRENCES OF *LIKE* IN CLAUSE-INITIAL POSITION



This time, the girls in both age cohorts make greater use of clause-initial *like* than the boys. The female lead is marginal in the 7-8 year old cohort, but is much more pronounced in the 10-11 year old cohort. Note too that the boys increase their frequency of *like* usage in clause-initial position with age, but this increase is not nearly as marked as that between the younger and older girls. Thus, even within a narrowly defined age range, there is evidence of gender diversity which manifests



itself not only between female and male speakers, but also within speakers of the same gender (see also Cameron 2005a: 48).

What is particularly interesting about the data displayed in Figures 6.2 and 6.3 above is that the male-preferential use of *like* clause-internally with noun phrases, and the female preference for *like* in clause-initial position is consistent with similar gender-related patterns discussed by D'Arcy (2005a) based on a much larger corpus of Canadian English spanning several key age ranges. D'Arcy (2005a: 135) refers to 'two distinct patterns' in her data where females are the frontrunners in the use of *like* in clause-initial contexts, whereas males take the lead in the use of *like* with determiner phrases.<sup>20</sup>

Why, then, do the boys in both age cohorts in the preadolescent data show a consistent preference in their use of *like* with noun phrases, and why do the girls appear to prefer clause-initial position? Quantitative differences between males and females in the use of discourse-pragmatic features are documented in the wider research literature. Erman (1992), for example, found that women tended to use certain pragmatic expressions to connect consecutive arguments, whereas men exhibited a preference to use them as attention-drawing devices, or to accomplish repair work. However, although there is an expanding body of research dealing with pragmatic variation in adult populations, this remains a relatively unexplored area for preadolescent speakers.

Recent research additionally suggests that there is social variation in the packaging of information in discourse, which may further illuminate gender-related

---

<sup>20</sup> The association of the use of *like* in VP contexts with males reported in D'Arcy (2005a: 194) is not, however, apparent in the preadolescent data.

differences in preadolescent speakers' use of *like*.<sup>21</sup> Cheshire's (2005a) research on gender and social class variation in the use of discourse-new markers has a direct bearing on the results discussed above. Using recorded interview data from adolescents, Cheshire (2005a) explored a range of syntactic and pragmatic strategies used to mark discourse-new referents. Among the strategies examined, Cheshire (2005a: 487) points to the use of discourse forms such as *like*, *sort of*, *you know* which can be used to explicitly mark noun phrases as requiring the collaboration of the addressee in the identification of new discourse referents. Of particular importance is Cheshire's (2005a: 492) finding that male adolescent speakers, especially middle-class males, exhibited a specific preference to overtly mark noun phrases representing discourse-new entities, whereas female adolescents, especially working-class speakers, used bare noun phrases significantly more often than males. These differences, as Cheshire (2005a: 496) suggests, may derive from female and male speakers having different orientations to the expression of referential and affective meaning in discourse.

Although an investigation of the full range of strategies used by the preadolescents in the management of information structure is beyond the scope of this chapter, it is possible, in the light of Cheshire's (2005a) findings, that the preadolescent boys' preferred use of *like* with noun phrases is a consequence of their particular orientation to the construction of talk, with the boys paying greater attention than the girls to overtly marking the introduction of discourse-new entities in their discourse.

---

<sup>21</sup> Gender-related differences have also been documented for the discourse marker *you know*. Erman (1992: 231) claims that *you know* is used by men more frequently in contexts related to the decoding of information, whereas women use *you know* more frequently to mark transitions between structural units of texts.

## 6.8 SUMMARY

In the preceding sections, I have shown that, in contrast with the observations made by Miller and Weinert (1995), *like* occurs pervasively in contemporary preadolescent speech in the southeast of England. Comparison of normalised frequencies of *like* across a number of studies (e.g. Cheshire 2002b; D'Arcy 2005a; Macaulay 2005; Tagliamonte 2005) further suggest that it is far from being recessive (contra Underhill 1988), and appears to be on the increase.

The results presented in this study indicate that *like* is used for a broad range of pragmatic functions which have been similarly documented for adolescent speakers (see Andersen 2001). With regard to the syntactic distribution of *like*, both preadolescent and adolescent speakers in London appear to be converging in their use of *like* in clause-initial and clause-internal positions, at the expense of the more 'traditional' clause-final variant.

In terms of age, children as young as seven are active users of *like* indicating that, as far as the distributional data examined above are concerned, there are no grounds for claiming that *like* is only acquired in early adolescence, after the age of 10 (Miller and Weinert 1995). The participation of preadolescent speakers in contemporary discourse-pragmatic changes complements previous research that has demonstrated that preteen speakers are not merely 'acquirers' of the vernacular, but are also engaged in socially motivated patterns of language variation and change (Roberts 2002: 333; see also Tagliamonte 2005).

I have also argued that in terms of its social embeddedness, discourse *like* exhibits interesting patterns of gender-affiliated variation in preadolescent speech, particularly in clause-internal positions with noun phrases, as well as in clause-initial position. Furthermore, distinctive patterns of distribution in specific syntactic

positions show interesting parallels with the findings uncovered for other varieties of English (see D'Arcy 2005a). Whether or not similar patterns obtain in the vernacular usage of older speakers of British English remains to be investigated.<sup>22</sup>

---

<sup>22</sup> It would also be instructive to explore whether particular markers are being replaced by new forms as part of the continual lexical renewal of forms expressing enduring pragmatic functions (see Brinton 1996; Brinton 2001:151). Tagliamonte (*in press*) shows that this is a fruitful avenue to explore. In an apparent-time analysis of sentence tags (e.g. *eh*, *or whatever*, *you know*, *right*) in Canadian English, Tagliamonte found that while some forms were highly correlated with the youngest generation, other forms such as *you know* were recessive in the usage of the youngest speakers. Interestingly, *you know* is rarely used in a discourse marking capacity by the British preadolescents I recorded.

## CHAPTER 7

# VARIATION IN THE QUOTATIVE SYSTEM

### 7.1 INTRODUCTION

In this chapter, I explore variation in the quotative system in the preadolescent corpus. This is a domain of variation that lies at the boundaries of syntax and discourse (Cameron 2005a: 30), and has figured prominently in a number of recent sociolinguistic studies (see e.g. Romaine and Lange 1991; Ferrara and Bell 1995; Tagliamonte and Hudson 1999; Macaulay 2001; Singler 2001; Cukor-Avila 2002; D'Arcy 2004; Tagliamonte and D'Arcy 2004, 2007; Barbieri 2005, 2007; Buchstaller 2004, 2006a, 2006b). Of particular relevance to the data examined in this chapter are two recent studies based on quotative variation in Canadian English, D'Arcy (2004), and Tagliamonte and D'Arcy (2004), which include data from preadolescent and adolescent speakers, and allow insightful comparisons to be drawn between different age cohorts, as well as between different varieties of English.

Variation in the use of *say*, *go*, *be like* and *think* to foreground reported speech or thought is an area that is undergoing rapid change in contemporary English vernaculars, with young people, particularly adolescents, cited as being in the forefront of change (Tagliamonte and D'Arcy 2004).

A study of quotative variation in preadolescent speech affords opportunities for tracking the acquisitional history of variant quotative markers in later childhood; comparing the rates and conditioning of variant choices in preadolescent speech; and exploring the discourse-pragmatic functions of quotative variants in preadolescence.

Furthermore, given that the quotative system is claimed to be a locus for repeated innovative language use (Buchstaller 2004: 294), a quantitative investigation of quotative variation in preadolescence can also potentially illuminate how preteen children may actively engage in changes in progress.

Following Tagliamonte and Hudson (1999: 148), I offer an integrative approach to the study of quotative markers in preadolescence by examining how different quotative variants interact within the quotative system as a whole. This holistic approach not only enables me to explore the differences in the pragmatic functions of particular quotative forms, but also facilitates assessment of the extent to which preadolescents are participating in specific domains of change, particularly in connection with the use of the *be like* quotative, which is a vigorous innovation claimed to be spreading in contemporary British vernaculars (see Tagliamonte and Hudson 1999:158; Buchstaller 2004: 272) as well as rapidly diffusing in several other major varieties of English (Baird 2001; Winter 2002; Tagliamonte and D'Arcy 2004; Singler 2001; Barbieri 2005, 2007).<sup>1</sup>

## 7.2 DEVELOPMENTAL ISSUES RELATING TO THE ACQUISITION OF QUOTATIVE STRATEGIES AND THE USE OF REPORTED SPEECH

In order to have a clearer picture of how children acquire variation in the quotative system and engage in trajectories of grammatical change, it is necessary to disentangle developmental changes from socially motivated changes that are gradually diffusing through speech communities and successive generations.

---

<sup>1</sup> A comprehensive discussion of the diachronic embedding of the quotative variants examined in this chapter is problematic owing to the dearth of previous studies which treat quotative variation from a diachronic perspective, and which examine the evolution of different pragmatic functions of quotative variants from a historical standpoint.

### 7.2.1 *Age effects and direct speech*

As far as reported dialogue is concerned, previous research has documented in some detail maturationally progressive changes in children's acquisition of quotative strategies. Hickman (1993) examined some of the developmental aspects of reported dialogue in narrative discourse by focusing on the speech of children aged between 4 and 10 years old and by comparing her findings with adult strategies for reporting dialogue. According to Hickman (1993: 63), 4 year old children typically represented speech without using explicit framing devices to demarcate reported dialogue from the rest of their narrative. Instead of producing explicitly framed quotations, the youngest children re-enacted dialogue and tended to use distinct acoustic representations to animate different voices in their narratives. Hickman (1993: 64) adds that as children mature, they gradually pay more attention to the metapragmatic aspects of narrative organization and develop the linguistic skills for integrating reported speech into their narratives. Hickman (1993: 72) notes that children between the ages of 7 and 10 had a clear preference for framed quotations in contrast with 4 year olds. In terms of the overall frequency of quotative verbs used by preadolescents, Hickman (1993: 74) found that most of the children's frames that she analysed contained the verb *say* whereas the adults' inventory of quotative verbs was found to be more varied. *Say* accounted for 87% of the quotative forms used by 10 year old children in contrast with adults, whose use of *say* constituted only 56% of their entire quotative cohort.

Hickman (1993: 72) reports that unframed quotations were rare in the speech of the 7-10 year olds that she studied. Furthermore, this age group was also found to have a quantitative preference for direct speech in narrative discourse, whereas adults relied on the indirect as well as the direct mode when reporting the words of others. It

is unclear whether the 7-10 years old children's greater reliance on direct speech in comparison with adults is solely attributable to developmental factors, or whether this is a stylistic feature characteristic of young speakers. Romaine and Lange (1991:268) suggest that direct quotation in child language may be a simplifying strategy that helps the speaker to sidestep some of the complexities associated with deictic modification and tense-shifting which are conventionally required by indirect reported speech in English. On the other hand, the greater use of direct speech by children may also be more generally related to the 'high involvement' narrative style in which young people are typically reported to engage (see e.g. Nordberg 1987; Stenström *et al.* 2002: 107; Norrby and Wirdenäs 2004: 247). This stylistic orientation is broadly characterised by a clustering of linguistic resources such as changes in voice quality, direct speech, onomatopoeia, and implicit expressions all used for dramatic effect in recounts of past events

### 7.2.2 *Gender and direct speech*

Using data taken from a corpus of personal narratives produced by children between the ages of 4 and 9, Ely and McCabe (1993) and Ely *et al.* (1996) report a number of important correlations between the use of reported dialogue and children's age and gender. According to Ely and McCabe (1993: 690), as children mature, they pay greater attention to the content of past talk as opposed to simply stating that a speech event took place. With increasing maturity, Ely and McCabe (1993: 671) also note that in their depiction of past events, children more frequently tend to quote themselves than any other speaker.

With regard to speaker differences in the use of direct quotation, Ely and McCabe (1993: 688) report that girls foreground conversational interaction in



depicting past events more than boys, and that this pattern may be indicative of broader gendered differences in narrative style. In her study of verbal disputes among African-American inner-city children, Goodwin (1990) similarly found that girls were the primary participants in 'he-said-she-said' confrontations. The research literature further suggests that the greater attention that girls pay to reporting speech emerges at an early age. In a small-scale study of gender-related differences in children's narratives, Ely *et al.* (1996: 18) report that by the age of 5, girls' attention to speech, as reflected in their use of reported dialogue, was twice that of boys.

Other gender-related differences in the use of direct speech involve the selection of particular quotative variants. In Ely and McCabe's study (1993: 681), girls used the vernacular quotative *go* more than boys, and also employed graphic dialogue introducers (Tannen 1986: 323) such as *ask* more frequently than boys.

### 7.2.3 *Direct speech as a strategy of embedded evaluation*

An important narrative function of direct speech which appears to be acquired in later childhood relates to its use in embedded or internal evaluation (Labov 1972: 372-3). Maybin (1999: 462) argues that preadolescents can take on other people's voices in reported dialogue in order to explore different evaluative perspectives by prosodically and paralinguistically reframing and manipulating appropriated voices for specific purposes in conversational interaction. This manipulation of prosodic features can result in a 'polyphonic layering of voices' (see Günther 1999; and Bakhtin 1981a) which enables speakers to contextualize their own attitudinal alignment towards the reported dialogue (Levey 2003). As Goffman (1981: 128) observes, interlocutors in conversation are constantly altering the 'footing' in their

talk, or the way in which they manage their relationship towards others, and communicate affective stances to what is being said.<sup>2</sup>

### 7.3 PREVIOUS ANALYSES OF VARIATION IN QUOTATIVE MARKERS

Recent research on quotation in oral narratives has emphasised the constructed nature of direct speech. For example, Tannen (1989:110) prefers the term 'constructed dialogue' to designate what is conventionally described as direct speech. In most instances of reported dialogue, no attempt is made at rendering precisely the actual words of a quoted utterance; instead, quotations are often used by speakers as creative performances rather than verbatim reports. Speakers can draw on an array of quotative forms with different pragmatic functions in order to dramatise narrative recounts and create listener involvement (see Tannen 1989: 133).

As far as the partitioning of functions between variants within the quotative system is concerned, there is a dearth of research on the interaction between quotative variants at different points in the lifespan (see also Buchstaller 2004: 208). However, what seems clear from earlier research is that members of the quotative pool differ in terms of their grammaticalised status, and social distribution (see Ferrara and Bell 1995: 269).

#### 7.3.1 *SAY, GO and BE LIKE contrasted*

Romaine and Lange (1991: 235) claim that it is because *say* reports dialogue without encoding any particular pragmatic effects that it is unmarked and hence a highly frequent quotative.<sup>3</sup> Lucy (1993: 96) similarly claims that *say* offers only

---

<sup>2</sup> Goffman (1981: 227) defines 'footing' as 'the alignment of an individual to a particular utterance.'

<sup>3</sup> There are cross-variety differences in the pragmatic functions embodied by *say*. Labov (1972: 372, f.n. 9) observes that in African American Vernacular English, non-lexicalised sounds co-occur with *say* rather than *go*; for example: *the girl says, 'Powww!'*

token pragmatic anchoring when used in conjunction with direct speech and functions like a metapragmatic pro-verb which is substitutable for any verb of speaking. By contrast, whereas *say* appears to highlight the propositional content of a quote and generally frames verbally expressed quotations, the more stylistically marked vernacular quotatives, *go* and *be like*, can introduce constructed dialogue (see 1 and 2); interior monologues (3); quotable gestures (4); as well as sound effects (see 5 and 6).

(1) he goes to me, <MIMICKING A BOYISH VOICE>, ‘Danielle I like your shoes’ (cited in Stenström *et al.* 2002: 116)

(2) I was like, ‘Mom?’

she was like, ‘What!’ (cited in Romaine and Lange 1991: 253)

(3) and I’m like, ‘OK, how am I gonna get her ‘chief complaint’ out of her?’

(cited in Fleischman and Yaguello 2004: 136)

(4) I’m like, <MAKING SCARED FACE>, but it’s exhilarating !

(cited in D’Arcy 2004: 329)

(5) it was just like sort of, she just went, <BANGING NOISE>, and just, fell over

and I just .... (cited in Stenström *et al.* 2002: 128)

(6) yeah that was a sweet and I was like, ‘U:::rgh’

(cited in Buchstaller 2003: 4)

*Be like*, the newest addition to the quotative pool, has been described as functioning like a ‘wild card’ quotative which can be used to introduce any kind of reported discourse without necessarily indicating whether the discourse was actually uttered or not (Buchstaller 2004), allowing speakers to adopt a position of reduced commitment to what they are reporting (see Romaine and Lange 1991: 238). Furthermore, according to recent studies of quotative variation, *go* and *be like* frequently function

as mimesis markers in discourse (see Buchstaller 2003: 6). From a pragmatic perspective, *go* and *be like* often occur in frames foregrounding reported speech that can be broadly characterised as ‘quotation as demonstration’ (Clark and Gerrig 1990), and are used to dramatise quotations by giving discourse prominence to the voice, accent and emotional tone of reported speakers, in addition to framing non-lexicalised speech sounds and iconic gestures.

### 7.3.2 *Previous research on the sociolinguistic distribution of GO*

According to Hudson (1985: 235), *go* functions in a similar way to *say* (i.e. it serves as a cue for messages produced by the vocal channel; see also Romaine and Lange 1991), except that *go* can additionally introduce any kind of noise or noiseless action performed by the speaker (see Butters 1980: 305). As far as tense is concerned, the research literature has also highlighted a correlation between *go* and the conversational historical present (see Chapter 5; Schiffrin 1981: 58; Blyth *et al.* 1990: 216; Winter 2002: 11; Stenström *et al.* 2002: 121; Carey 2004: 7).<sup>4</sup> However, this correlation is not corroborated in all studies; for example, in Macaulay’s (2001: 10) analysis of quotative variation in the speech of Glaswegian adolescents, *go* was found to occur more frequently in the past tense (67%).

With regard to the social distribution of quotative *go*, Macaulay (2001: 6) suggests that it became popular in the United States among younger speakers about 25-30 years ago, although, as Buchstaller (2006: 4) notes, the diachronic evolution of *go* remains under-researched (see also Barbieri 2005: 223). Butters (1980) gives examples of quotative *go* that date back to the 1940s, although Buchstaller (2006a:

---

<sup>4</sup> In this chapter, the occurrence of quotative variants with particular tense forms is part of a broader investigation than the one presented in Chapter 5. Whereas the data relating to quotative variation and tense in Chapter 5 are based on a sub-corpus of preadolescent narratives, this chapter considers the total number of eligible quotative contexts in the entire preadolescent corpus, and also looks at tense variation with other quotative variants such as *be like*.

25, f.n. 20) reports that the use of quotative *go* to foreground sounds and gestures is attested in the *Oxford English Dictionary* as early as 1791.

According to Cukor-Avila (2002: 3), there is a general consensus in the research literature that *go* is used more often by teenagers and young adults. Stenström *et al.* (2002: 125) claim that the use of *go* in the *Bergen Corpus of London Teenage Language* (COLT) is a function of speaker age, with its frequency rising during early adolescence, then peaking in middle adolescence, followed by a subsequent sharp drop in frequency after late adolescence. Similarly, Buchstaller's (2006a) analysis of quotative variation in British English reveals a binary split with regard to the frequency distribution of *go* according to age, with younger speakers using this variant proportionally more often (19%) than older speakers (2%). However, Buchstaller (2006a: 22-3) points to the sporadic distribution of *go* across different age cohorts in American and British English, and suggests that its evolution 'cannot be fully captured by the notions of age grading or change in progress,' rather, *go* appears to recede at certain points in time, only to be 'recycled' by later generations of speakers, resulting in a wave-like patterning across different ages.

Previous research on the patterning of quotative *go* according to gender has produced conflicting findings. For example, in London, Stenström *et al.* (2002: 126) found that female rather than male adolescents were the main users of *go*, and Buchstaller (2004: 258) reports that *go* is associated with female speakers in the Derby/Newcastle corpus that she analysed. Macaulay (2001: 14) also found high rates of usage of quotative *go* among working-class women in Glasgow. However, differences emerge when British English is compared with other varieties of English. Tagliamonte and Hudson (1999: 160), for example, claim that *go* appears to be

associated with male speech in North American English (see also Singler 2001).<sup>5</sup> Different varieties also appear to favour different pragmatic uses. For instance, in colloquial Canadian English, *go* is favoured with internal dialogue, whereas in British English, direct speech rather than internal dialogue is a favoured context for *go* (Tagliamonte and Hudson 1999: 164).

### 7.3.3 Previous research on the sociolinguistic distribution of BE LIKE

Quotative *be like* also has contrasting sociolinguistic profiles in British and North American English. Particular interest has been paid to its rapid expansion within a relatively compressed time frame in North American varieties, which has been substantially documented in a number of recent publications (e.g. Ferrara and Bell 1995; Tagliamonte and D'Arcy 2004, 2007).<sup>6</sup> According to Singler (2001: 276), *be like* has established itself as the primary quotative in vernacular American English, particularly in the speech of the younger generation. Tagliamonte and D'Arcy (2004: 493) present evidence that there has been a tremendous increase in the use of *be like* in the speech of young Canadians over the past several years. By drawing comparisons between frequency data in two Canadian corpora from 1995 and 2002/3 respectively, Tagliamonte and D'Arcy (2004: 501) report that the proportion of *be like* usage by Canadian youth has increased by more than four and a half times in the

---

<sup>5</sup> Buchstaller (2006b) finds notable differences in the social stereotypes attached to *be like* and *go* in the U.S. and the U.K. (e.g. British respondents, in contrast with Americans, do not evince any strong covert attitudes with regard to the social distribution of *go*). Buchstaller (2006b: 375) claims that global innovations such as *be like*, which putatively diffused to other varieties of English from the U.S., are subject to social re-evaluation or re-analysis in the transmission process, and are not simply transferred 'wholesale' from one variety to another (see also Meyerhoff 2006: 246). Further evidence that global forms such as *be like* may be 'customised' in the process of diffusing from one variety to another can be found in Macaulay (2001), who notes the use of the quotative variant *be like that* in the speech of Glaswegian adolescents. According to Macaulay (2001: 13), '*be like* has not crossed the Atlantic in a straightforward manner.'

<sup>6</sup> North America, specifically California, is identified by Tagliamonte and D'Arcy (2007) as the geographical point of origin for the diffusion of *be like*. Tagliamonte and D'Arcy (2007: 199) claim that *be like* gained impetus in the 1980s due to its 'preppy associations'. However, this study does not specify the extent to which *be like* was used before the 1980s, and provides no attestations of this variant before this time. Thus, the early history of *be like* remains obscure.

last seven years, with the result that *be like* is now the predominant quotative variant in their 2002/3 corpus, accounting for 58% of quotative forms.

Other studies in Canada report similar results. Dion and Poplack (2005), who analysed quotative variation in the speech of 100 native anglophones in Quebec aged between 18 and 93, found even higher overall rates of *be like* in the speech of young Quebecers than those reported by Singler (2001) and Tagliamonte and D'Arcy (2004) for other cities in North America. D'Arcy's (2004) study of quotative usage in St. John's in Newfoundland establishes that *be like* has also clearly diffused to preteen Canadian speakers (aged 8-11), for whom it is the most frequent quotative variant.

Although Milroy (2002: 7) refers to the spread of quotative *be like* as a transnational phenomenon, Tagliamonte and Hudson (1999: 149) point out that most of the research on this form has been undertaken in relation to American English, and so less is known about the trajectories of change in which *be like* may be currently engaged in contemporary varieties of British English. Recent studies on quotative usage in British English vernaculars suggest that there has been an expansion in the use of *be like*, although it is not yet entirely clear whether reported increases in British English within the last decade are comparable with the magnitude of those documented for Canadian or American English.<sup>7</sup> Stenström *et al.* (2002: 118) note that *be like* amounts to less than 0.5% of all the quotatives used by London adolescents in the COLT corpus collected in 1993.<sup>8</sup> By contrast, using data collected several years later in 1997, Macaulay (2001: 10) reports that *be like* represents 14% of Glaswegian adolescents' quotative usage.

---

<sup>7</sup> Tagliamonte and Hudson (1999: 161) report that in the mid-nineties, *be like* was found more frequently and across many more speakers in their British corpus than in their comparable Canadian English corpus.

<sup>8</sup> The low frequency of *be like* in the COLT corpus dovetails with Buchstaller's (2006a: 8) observation that *be like* had only just begun to make inroads into the quotative cohort in British English in the early 1990s.

Additional corroborative evidence of an increase in the frequency of *be like* over time in British English can be found in Carey (2004), who compared the British quotative corpus analysed in Tagliamonte and Hudson (1999) with a comparable corpus drawn from the *York Narrative Database* (Smith 2003). Given that Tagliamonte and Hudson's (1999) British quotative corpus was originally collected in 1996, the comparison between the two data sets enabled the expansion of *be like* in spoken British English to be situated within a temporal interval of several years. Carey (2004: 6) found that whereas *be like* accounted for 18% of total quotative usage in 1996, it had substantially risen in frequency by 2003, representing 38% of total quotative forms used. Baker *et al.*'s (2006) recent analysis of quotative usage in the speech of 31 students aged 18-22 recorded in York in 2006 suggests that there has been a further expansion in the use of *be like*, which accounts for 60% of total quotative usage in their study. A synopsis of the shifting frequencies of *be like* in several British corpora spanning a period of twelve years is displayed in Figure 7.1 below.<sup>9</sup>

Ferrara and Bell (1995: 285) note that there is considerable variation in the use of *be like* from the perspectives of age, gender and ethnicity as well as from the standpoint of urban versus rural locations. If we take age as a starting point, there is a general consensus in previous studies that this correlates significantly with *be like* usage. In research on American English, Blyth *et al.* (1990: 219) report that the use of *be like* falls sharply after the age of 25, and in a later study of quotative usage in New York City, Singler (2001: 267) notes that speakers under the age of 35 are the

---

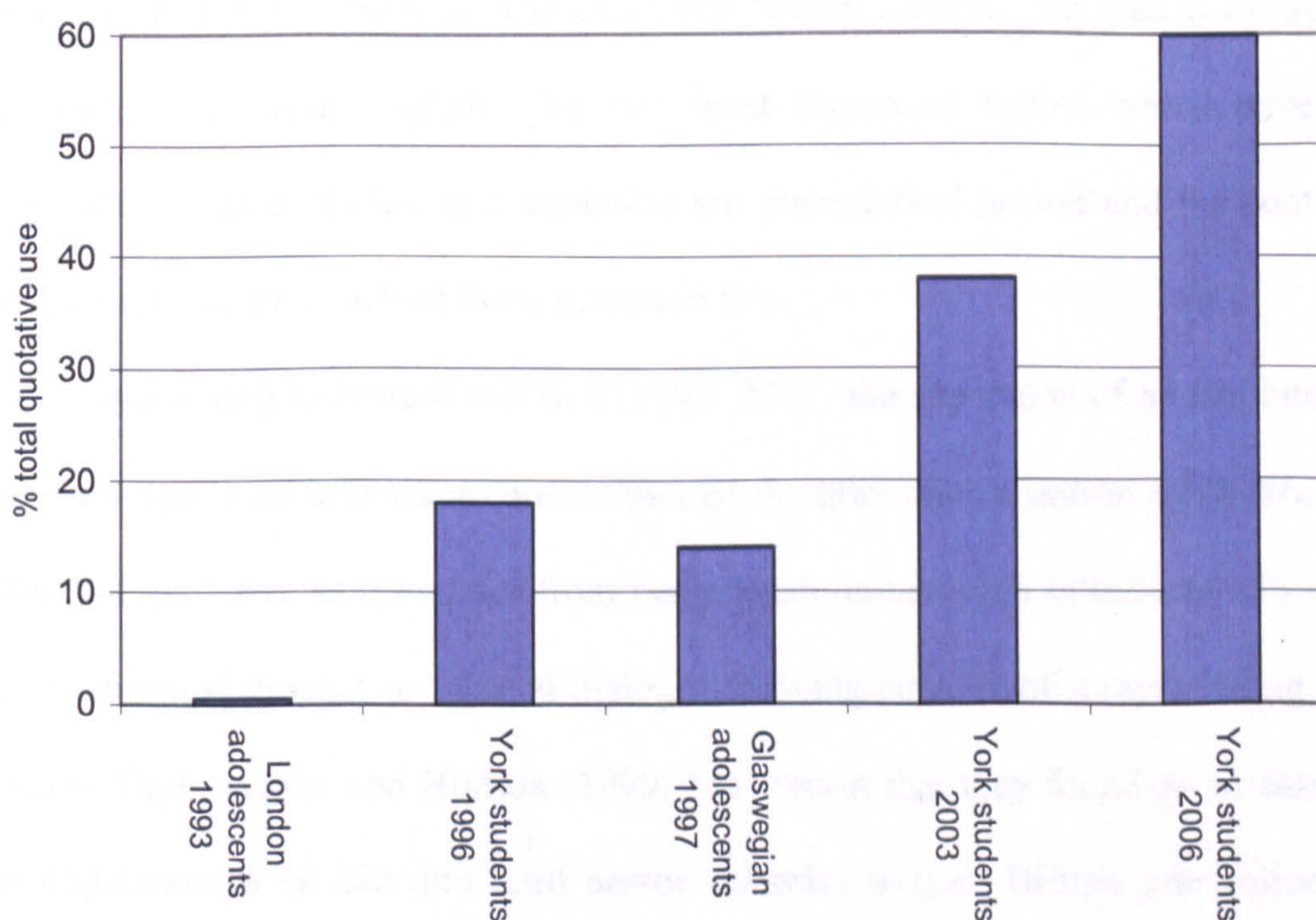
<sup>9</sup> A comparison of frequency distributions of *be like* in various British corpora is necessarily crude given that the corpora were not all collected in the same geographical location, and were compiled using different methodological approaches. Rates of occurrence can also be expected to fluctuate according to features of the situation (see Poplack and Tagliamonte 1991: 318). Two further caveats apply: firstly, although the corpora are drawn from speakers under the age of 30, the age range varies from corpus to corpus. Secondly, information about the socioeconomic background of speakers is incomplete.



dominant users of *be like*, whereas older speakers make comparatively less use of this form. In Britain, Buchstaller (2004: 214) observes that *be like* is distributionally biased towards younger speakers in the Derby/Newcastle Corpus, and Macaulay's (2001: 10-12) data on *be like* in the speech of Glaswegian adolescents and adults show clearly that *be like* is much more strongly favoured by young speakers.

FIGURE 7.1

FREQUENCY DISTRIBUTION OF *BE LIKE* IN FIVE DIFFERENT CORPORA OF BRITISH ENGLISH (FIGURES ABSTRACTED FROM STENSTRÖM *ET AL.* 2002; TAGLIAMONTE AND HUDSON 1999; MACAULAY 2001; CAREY 2004; BAKER *ET AL.* 2006)



Exactly at what age *be like* is acquired by younger speakers has not been comprehensively addressed (Tagliamonte and D'Arcy 2004: 494). Tagliamonte and D'Arcy (2004: 505-6) present evidence from Canadian English which indicates that

although the internal constraints on *be like* are acquired in preadolescence, the social constraints on *be like* are not fully operative until high school.

Furthermore, Tagliamonte and D'Arcy (2004) highlight the role of gender-differentiation in the ongoing grammaticalisation of *be like* in the speech of Canadian youth. In contrast with Ferrara and Bell's (1995: 285) finding that an initial female lead in the use of *be like* was later neutralised, Tagliamonte and D'Arcy (2004: 508) report that in their apparent-time study, the more frequently *be like* occurs, the stronger the effect of gender becomes, with females firmly at the forefront of change.

In terms of internal constraints on *be like*, Tagliamonte and Hudson (1999: 147) claim that the linguistic trajectory of change for *be like* exhibits remarkable parallelisms not only between Canadian and British corpora, but also with previous studies on American English. The two most important factors which have been implicated in its evolution as a quotative are grammatical person and the content of the quote. I consider each of these factors in turn.

According to Ferrara and Bell (1995: 270), the expansion of *be like* into third person usage is an important concomitant of the grammaticalisation of *be like* as an all-round quotative as it evolves from being predominantly an introducer of verbally non-committed thought or internal dialogue to being capable of foregrounding actual speech. Tagliamonte and Hudson (1999: 161) report that they found no evidence for the expansion of *be like* into third person contexts in their British data collected in 1996, and Carey's (2004: 8) later study indicates that *be like* is still preferred in first person contexts, with the third person disfavouring the use of this quotative. The most recent variationist analysis of *be like* in British English carried out by Baker *et al.* (2006) confirms the constancy of the grammatical subject constraint, with first person subjects continuing to favour *be like*.

As far as the relationship between the content of the quote and the ongoing grammaticalisation of *be like* is concerned, Ferrara and Bell (1995: 279) suggest that there is a developmental continuum, with the use of *be like* as an introducer of externally realised speech emerging later than its other pragmatic functions associated with foregrounding internal dialogue, non-lexicalised sounds and quotable gestures. Whereas Tagliamonte and Hudson (1999: 164) report that *be like* in British English is favoured with non-lexicalised sounds and internal dialogue more than with direct speech, Carey (2004: 10) observes that there has been a significant change in the hierarchy of *be like*'s pragmatic functions judging from her results based on the *York Narrative Database* (Smith 2003). In contrast with Tagliamonte and Hudson's (1999) study, Carey (2004: 9) found that *be like* is most favoured with direct speech, followed by non-lexicalised sounds, and internal dialogue. However, this re-weighting of discourse functions is not corroborated by Baker *et al.*'s (2006) recent analysis, which does not document an expansion of *be like* into direct speech contexts.

A third factor relating to the choice of *be like* which has received less systematic attention in the research literature is tense (see Barbieri 2005: 229). In contrast with *say*, which is typically found to occur in the past tense (see Blyth *et al.* 1990: 218; Singler 2001: 277), Barbieri (2005: 222) observes that in American English *be like* occurs more frequently in the present tense. By contrast, both Macaulay (2001:10) and Carey (2004: 7) report that in their British data, there is a greater tendency for *be like* to occur in the simple past, which points to possible differences between American and British English in relation to correlations between

vernacular quotatives and the use of particular tense forms.<sup>10</sup>

In the following sections, I look at the multiple constraints which influence the choice of quotative variants in the preadolescent corpus. I operationalise several of the factors highlighted in previous research discussed above, such as content of the quote, grammatical person as well as speaker age and gender in order to assess how variant forms interact in the quotative system. I also explore the extent to which distributional patterns in the preadolescents' quotative usage, particularly in connection with *be like*, reflect the participation of preteen speakers in 'fast-spreading global-scale innovations' (Buchstaller 2006a: 4).

#### 7.4 METHODOLOGY

Following Buchstaller (2006a: 5), the envelope of variation was initially delimited by setting up a (*say*) variable functionally defined as 'all strategies used to introduce reported speech, sounds, gesture and thought by self or other' (see also Cameron 2000: 256).

Every instance of a quotative verb introducing direct speech, internal dialogue, a non-lexicalised sound, or a quotable gesture was extracted from the corpus and

---

<sup>10</sup> The difficulty in characterising the evolutionary trajectory of *be like* in North American and British English respectively is compounded by the fact that no studies have been based on a 'large-scale socially stratified sample from a single, cohesive, speech community' (Tagliamonte and D'Arcy 2007: 200). This has resulted in a lack of information about the diffusion of *be like* at the community level, so it has not been entirely clear whether the use of this variant reflects age-grading, or a bona fide change in progress which is percolating through the speech community. However, on the basis of an in-depth community study in Toronto, Tagliamonte and D'Arcy (2007: 212) claim that the distribution of *be like* reveals the 'classic S-curve of linguistic change.' Furthermore, the authors establish a crucial link between *be like* and the conversational historical present. This association is claimed to be derivative of the earlier functional correlation of *be like* with reporting internal monologue (Tagliamonte and D'Arcy 2007: 211-212). Tagliamonte and D'Arcy (2007: 212) argue that the rising stylistic option of reporting internal monologue in narrative recounts has accelerated the use of *be like*. It is not clear what implications the association between narrative style and the use of *be like* has for the spread of this variant in Britain. Tagliamonte and Hudson (1999:166) highlight the impact of qualitative differences in narrative style between British and Canadian speakers on the differential distribution of quotative variants in British and Canadian English respectively.

retained for analysis.<sup>11</sup> Given the focus on the quotative cohort in its entirety, it was also essential that instances of direct speech which were not introduced by an overt lexical form (i.e. the so-called zero quotative; see example (17) below, and Mathis and Yule 1994) were extracted from the corpus (Meyerhoff 2006: 243). Zero quotatives were identified using contextual information recoverable from the audiotapes, such as a change in the reported speaker indicated by variations in the prosodic format of constructed dialogue (see Couper-Kuhlen 1998).

Other verbs which have no association with externally realised speech were necessarily included in order to provide an accountable analysis of quotative variation in the corpus. Although the verb *think* does not by its very nature foreground external dialogue, following Tannen (1986: 315), and Tagliamonte and D'Arcy (2004: 511, f.n. 2), it was methodologically necessary to include this verb within the inventory of variant forms, not only in order to draw accurate comparisons with previous studies, but also to investigate any potential distributional differences in the use of introducers of internal dialogue in the preadolescent corpus.<sup>12</sup>

Lexical forms used to preface indirect speech were excluded from the study as were instances of *say*, *think*, *go* and *like* used in a non-quotative function (e.g. *go* as a verb of motion and tense marker; and *like* as a discourse marker, see Chapter 6). In addition, following standard variationist practice, incomplete utterances were excluded, as were tokens where essential coding information such as grammatical person or number was not easily recoverable:

---

<sup>11</sup> Tagliamonte and Hudson (1999: 155) have pointed out that individual quotative verbs are not discursively equivalent, although it is only when all variants are included in the same analysis that these differences can be clearly assessed. In defining the envelope of variation, I operate with the notion of 'weak complementarity' of variant forms (Sankoff and Thibault 1981: 207). Sankoff and Thibault (1981: 208) note that in many cases 'the most we will be able to say is that the proposed variants can serve one, or more generally, similar discourse functions. We cannot even require that they be identical discourse functions.'

<sup>12</sup> Bolden (2004: 1081, f.n.14) argues that thoughts can be reported in ways that are similar to reporting somebody else's or one's own speech; thus, *think* should be included in the list of lexical quotatives.

(7) and I was like ..... [16M7/8] (incomplete utterance)

(8) cos the <...> said, 'You wanna cup of tea ?' [25F7/8] (grammatical person/number unclear)

A total of 514 tokens were coded for various factors that previous research has highlighted as constraining the choice of variants in the quotative system. Each token was coded with respect to the age and gender of the speaker. Every token was additionally coded for four main internal factors: quotative variant (i.e. *say, go, be like, zero, etc.*); grammatical person; content of the quote (i.e. direct speech, non-lexicalised sound, internal dialogue or quotable gesture); as well as for tense (e.g. present, conversational historical present, simple past, present perfect and past progressive).

Because some of the coding procedures involved important methodological decisions, they warrant brief consideration here. With regard to the content of the quote, the identification of non-lexicalised sounds and quotable gestures was straightforward as these had been carefully identified and annotated in the transcriptions of the audio-taped material shortly after the recordings were made. However, distinguishing between direct speech and internal dialogue was rather more problematic. In order to mitigate this difficulty, I adopted the tried-and-tested systematic procedures described in Ferrara and Bell (1995: 279), Buchstaller (2004: 53); Tagliamonte and D'Arcy (2004: 499); and Barbieri (2005: 235), who based coding decisions on contextual clues such as the role of framed quotations in the sequential unfolding of the complicating action (i.e. their contribution to the advancement of the storyline). Coding decisions were also informed by the principles underlying the normative framework of conversational organisation (Schiffrin 1994:

236). For example, '*he-said-she-said*' or '*I-was-like-she-was-like*' statement and response sequences were treated as instances of direct speech. Similarly, question and answer sequences, which are paradigm examples of adjacency pairs (Schiffrin 1994: 16) conventionally involving two or more conversational participants, were coded as instances of direct speech. On the other hand, if a quotative was interpreted as preceding verbally uncommitted thought, or a general feeling of a speaker, then it was coded as an instance of internal dialogue, as in the following example from the preadolescent corpus:

(9) then I had my nightmare which someone done that same thing to me (*sic*)

then I woke up, *I'm like*, 'I'm not thinking about that again' [27F7/8]

Finally, I coded for a broader variety of tense forms than those included in previous studies, as initial perusal of the data revealed that certain quotative variants appeared with tense forms (e.g. the present perfect; see Chapter 5) that have not been reported in studies of quotative variation based on North American varieties of English.

## 7.5 VARIATION IN THE LEXICAL-SYNTACTIC DEVICES FOR MARKING QUOTATION IN THE PREADOLESCENT CORPUS

Initial scrutiny of the data revealed that there are a number of strategies in the corpus used to frame quotation:

(10) they just *went*, 'Kkkkkkk' like that [15M7/8]

(11) he went in his bedroom and *said*, 'Oh look I've got some matches, I'll put it alight, shall I?' [23F7/8]

(12) we all went out in the morning and *goes*, 'Oooh, it's a bomb' [23F7/8]

(13) he *was going*, 'Where are you?' [25F7/8]

- (14) she *thought*, 'I better think of something' [26F7/8]
- (15) he *goes*, <MIMICS BITING> [16M7/8]
- (16) all the players have come up and *I've gone*, 'What?' [12M7/8]
- (17) he's *going*, 'No!'
- 'my leg's stuck... I can't move' [10F10/11]
- (18) *he's like*, 'Do it now!' like that
- and *I'm like*, 'No!' [6F10/11]
- (19) he were *all like this*, <MIMICS ANGELIC BEHAVIOUR> [15F10/11]
- (20) they're *saying*, 'What's tha:::t?' and all that [9F10/11]
- (21) my mum's here *saying like*, 'No, you can't have them as they're not ironed or washed or anything' [6F10/11]

Most of the examples above consist of a lexical quotative that frames a direct speech extract, a sound effect, internal dialogue or a quotable gesture. In example (17), however, there is a change of speaker to the right of the arrow where the following extract of speech is not preceded by an overt quotative verb, but is marked by a zero quotative signalled by prosodic variations in the reporting speaker's voice.

Several of the quotes above (e.g. 11, 12) also contain discourse markers of various kinds (see also Buchstaller 2004: 109; and Macaulay 2001: 13). In example (12) above, *ooh* at the beginning of the reported dialogue gives discourse prominence to an utterance with a strong emotional content that functions like a 'response cry' (see Goffman 1981). In addition to contributing to the dramatic replay of narrated events, it may also be the case that these discourse particles function on the textual plane by helping to demarcate the boundary between reporting and reported speech.<sup>13</sup>

---

<sup>13</sup> Schourup (1985) and Macaulay (1991) note that discourse markers such as *well* and *oh* can serve framing functions in discourse connected with marking the transition to reported speech.



Another example of a framing device can be found in the use of the general extender *and all that*, which is appended to the quoted dialogue in example (20).<sup>14</sup> According to Overstreet (1999: 118-9), general extenders frequently occur in the environment following quoted speech or thought, and serve to mark the preceding quoted material as an approximation in form and content to what could have been said in a specific instance. Textually, the general extender in example (20) may also help to mark the terminal boundary of the reported speech extract (see further Chapter 8).<sup>15</sup> Bolden (2004: 1072) points out that very few languages appear to have a fully grammaticalised unquote marker which signals the end of quoted talk, although many languages draw on discourse-pragmatic devices in conjunction with suprasegmental features to demarcate reported speech from subsequent non-reported speech (see e.g. Golato 2000 for a discussion of unquote markers in German).

It is also interesting to observe that in addition to the quotative variants *say*, *go*, and *be like* a number of the examples listed above contain mixed or 'hybrid' variants where *like* co-occurs with either another pragmatic form such as *all* in *all like this* in example (19), or with another lexical quotative such as *say* in example (21).<sup>16</sup> Similar collocations are attested in other vernacular varieties of English. Rickford *et al.* (2005) suggest that *all like* may be a reanalysed form, with *all* intensifying quotative *like*.<sup>17</sup> Rickford *et al.* (2005) further report that in their 2004/5 quotative corpus based on data from Californian speakers, quotative *all like* was particularly favoured by younger speakers.

---

<sup>14</sup> (20) is a good example of what Tannen (1989: 111) classes as 'dialogue as instantiation'; in other words, the reported utterance was not necessarily actually spoken, but is offered as an instantiation of a response type.

<sup>15</sup> Another device with similar textual functions can be found in examples (10) and (18) where the phrase *like that* appears to operate textually like an unquote marker.

<sup>16</sup> Note also that example (21) contains the proximal deictic *here* as part of the quotation format. Milroy and Milroy (1977) have also noted the use of quotative sequences incorporating deictic elements such as *here's me*, *here's him*, which alternate with *I says*, *she says*, etc. in vernacular Belfast English.

<sup>17</sup> The existence of *be all* as a quotative in its own right is documented in Rickford *et al.* (2005). No examples of this variant were found in the preadolescent data.

In Singler's view (2001: 275), constructions which involve collocations of *like* with other quotatives (e.g. *say like*, *go like* and *think like*) may constitute intermediate stages in the evolutionary trajectory of *be like* (see also Winter 2002: 9). These forms have also been documented in British English vernaculars (Macaulay 2001: 15; Stenström *et al.* 2002: 109; Buchstaller 2004: 188), although the status of these variants is not yet clear.<sup>18</sup> If we take a construction-based approach to the study of grammaticalisation (see e.g. Bybee 2003), then there may be a number of evolving sequences consisting of lexical collocations with *like* that are used in quotative frames, possibly with varying pragmatic functions.

In the following section, I turn to a quantitative analysis of the major quotative variants in the preadolescent corpus.

## 7.6 DISTRIBUTIONAL ANALYSIS

Table 7.1 below shows the overall distribution of quotative variants in the preadolescent corpus.

The results show that the most frequent form in the corpus is *say*, although *go* is not very far behind. From a proportional perspective, it can be seen that the preadolescents principally employ a bipartite quotative system: *go* and *say* account for 90% of all quotative forms used, with other forms such as *be like* occupying a much more liminal position in the quotative system.

---

<sup>18</sup> Researchers do not deal with *go like* and *say like* in a consistent way. Singler (2001: 266, f.n.10), for instance, combines occurrences of *go like* with quotative *go*, whereas (Macaulay 2001: 11, Table 5) combines distributional results for *be like (that)* and *go like (that)* in his Glaswegian data. In the distributional analysis of the preadolescent data, I separate tokens of *be like* from tokens of *say like/go like*.

TABLE 7.1

OVERALL DISTRIBUTION OF QUOTATIVE VARIANTS IN THE  
PREADOLESCENT CORPUS

QUOTATIVE	%	N
<i>say</i>	48	249
<i>go</i>	42	215
<i>be like</i>	5	24
<i>say like /go like</i>	3	14
<i>zero</i>	1	5
<i>think</i>	1	4
Miscellaneous	0.6	3
TOTAL		514

Both the zero quotative and *think* are negligibly instantiated in the corpus.<sup>19</sup> The infrequency of the zero quotative in the preadolescent data corroborates Hickman's (1993) observation that unframed quotes are rare in this age range. Other studies, however, have reported much higher frequencies of the zero quotative, with young speakers cited as heavy users of this form (see e.g. Buchstaller 2004: 279), although, as Singler (2001: 268) observes, this is an area of vast individual variation (see also Winter 2002: 10 on frequency differences between varieties). The remaining

<sup>19</sup> In Tagliamonte and D'Arcy's (2004: 501) analysis of the quotative system in Canadian English, *think* accounts for only 2% of variant forms. Moreover, in the Canadian data it is restricted to speakers aged between 17 and 19 (see Tagliamonte and D'Arcy 2004: 502, Figure 2). Stenström *et al.* (2002: 110) note that the zero quotative is common in the COLT corpus, but do not give frequency counts for this form. Macaulay (2001: 10-12, Tables 4 and 6) shows that adolescents in Glasgow use the zero variant more frequently than adults.

quotative variants in the corpus categorised under 'miscellaneous' consist of a restricted number of graphic dialogue introducers such as *shout* and *shout out*.

These results contrast with the broader range of forms that are attested in adolescent speech (compare Macaulay 2001: 10), and differ markedly from the results in Tagliamonte and D'Arcy (2004: 501), and D'Arcy (2004: 332), who show that *be like* is the most frequent quotative marker used by the young Canadian speakers that they analysed. Moreover, even if the comparison is restricted to preadolescent speakers, it is apparent from D'Arcy's (2004: 333) data that Canadian preadolescents (aged 8-11) make proportionally far greater use of *be like* than the London preadolescents investigated here.<sup>20</sup>

#### 7.6.1 *Distribution of common quotative variants by age and gender*

Figure 7.2 below gives a more detailed breakdown of the use of quotative *say*, *go* and *be like* by age and gender. It is clear from these data that children as young as 7 make some limited use of the *be like* variant, but in neither age cohort can *be like* be claimed to compete to any major extent with either *say* or *go*. The younger children make use of *be like* more often than the older children, and females use it somewhat more frequently than males in the 7-8 year old cohort. There is a restricted number of tokens of *be like* in the speech of the 10-11 year old females, but 10-11 year old males do not use it at all.<sup>21</sup>

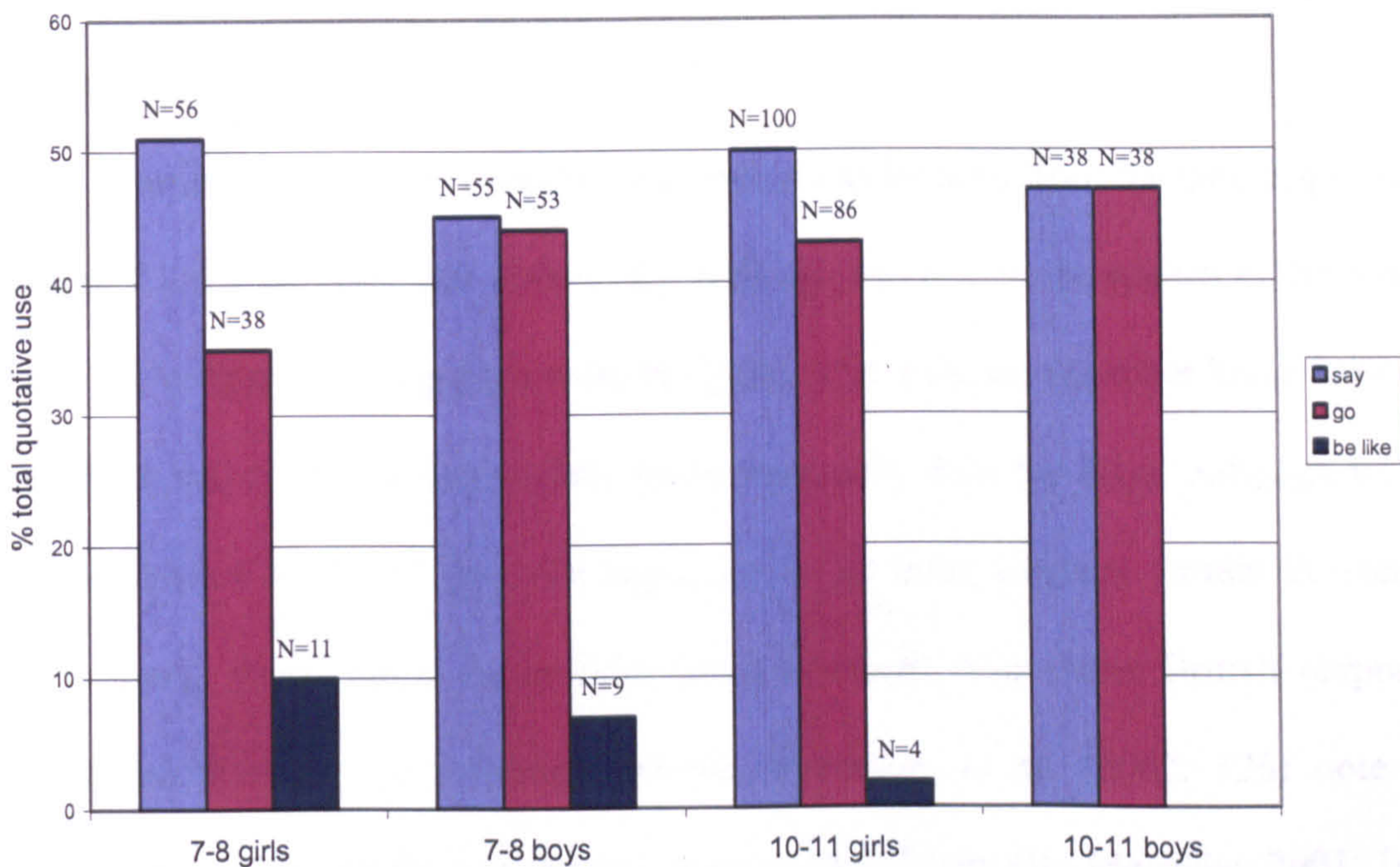
---

<sup>20</sup> This contrast is even more striking if we consider the different sociocultural contexts in which D'Arcy's (2004) Canadian data (recorded in Newfoundland) and my preadolescent data were obtained. D'Arcy (2004: 324) classifies St. John's English as an autonomous regional Canadian variety which is situated within a 'linguistic relic area' (D'Arcy 2005b: 330), and which differs from General Canadian phonologically, lexically and morphologically. By contrast, the London preadolescents were recorded in an area which is located squarely within the influence of mainstream linguistic developments in British English. However, it should be borne in mind that in comparing the Canadian and British data sets, D'Arcy's (2004) results are based on a group of young girls only.

<sup>21</sup> The mixed forms *go like* /*say like*, although only nominally instantiated in the preadolescent data, are more evenly distributed between males and females, as well as across the different age cohorts.

FIGURE 7.2

DISTRIBUTION OF *SAY*, *GO* AND *BE LIKE* BY AGE AND GENDER



Evidently, the paucity of *be like* tokens in the preadolescent data precludes a detailed discussion of how this variant is socially constrained in this particular age group. Furthermore, the absence of substantial data from older speakers impedes speculation about how this variant may be filtering through to younger speakers in the London metropolitan area.<sup>22</sup> Another complicating factor is the matter of social class. Recall that the preadolescents in this study are predominantly from working-class backgrounds. The extent to which social class correlates with the use of *be like* has not been fully investigated in a British context (see Cheshire *et al.* 2005a:143-4), so it is not clear how this parameter impacts on the frequency of *be like* usage. Some preliminary indications that *be like* is sensitive to social class membership can be found, however, in Macaulay's (2005) study of common discourse features in the

<sup>22</sup> The severely restricted occurrence of *be like* in the COLT London adolescent data recorded in 1993 prohibits any firm conclusions being drawn about the internal linguistic or social constraints operating on this variant in the speech of London teenagers.

speech of Glaswegian adolescents. Drawing on a small number of *be like* tokens, Macaulay (2005: 83) found that *be like* occurred more frequently in the speech of middle-class adolescents than in speech of teenagers from working-class backgrounds.

Turning to the other quotative variants, it can be seen from the data displayed in Figure 7.2 that *go* competes extensively with *say*, more so in the speech of the boys in both age cohorts in comparison with the girls. The girls, on the other hand, have a tendency to use quotative *say* slightly more frequently than the boys, although with the exception of the 10-11 year old boys, *say* is the most frequent variant in every group. Again, these results differ from those obtained from other British corpora targeting adolescents. As mentioned above, Stenström *et al.* (2002: 126) note a marked female predominance in the use of *go* in COLT (see also Macaulay 2001: 14; Buchstaller 2004: 273 for similar correlations), although no such correlation obtains in the preadolescent data.<sup>23</sup>

### 7.6.2 *Content of the quote*

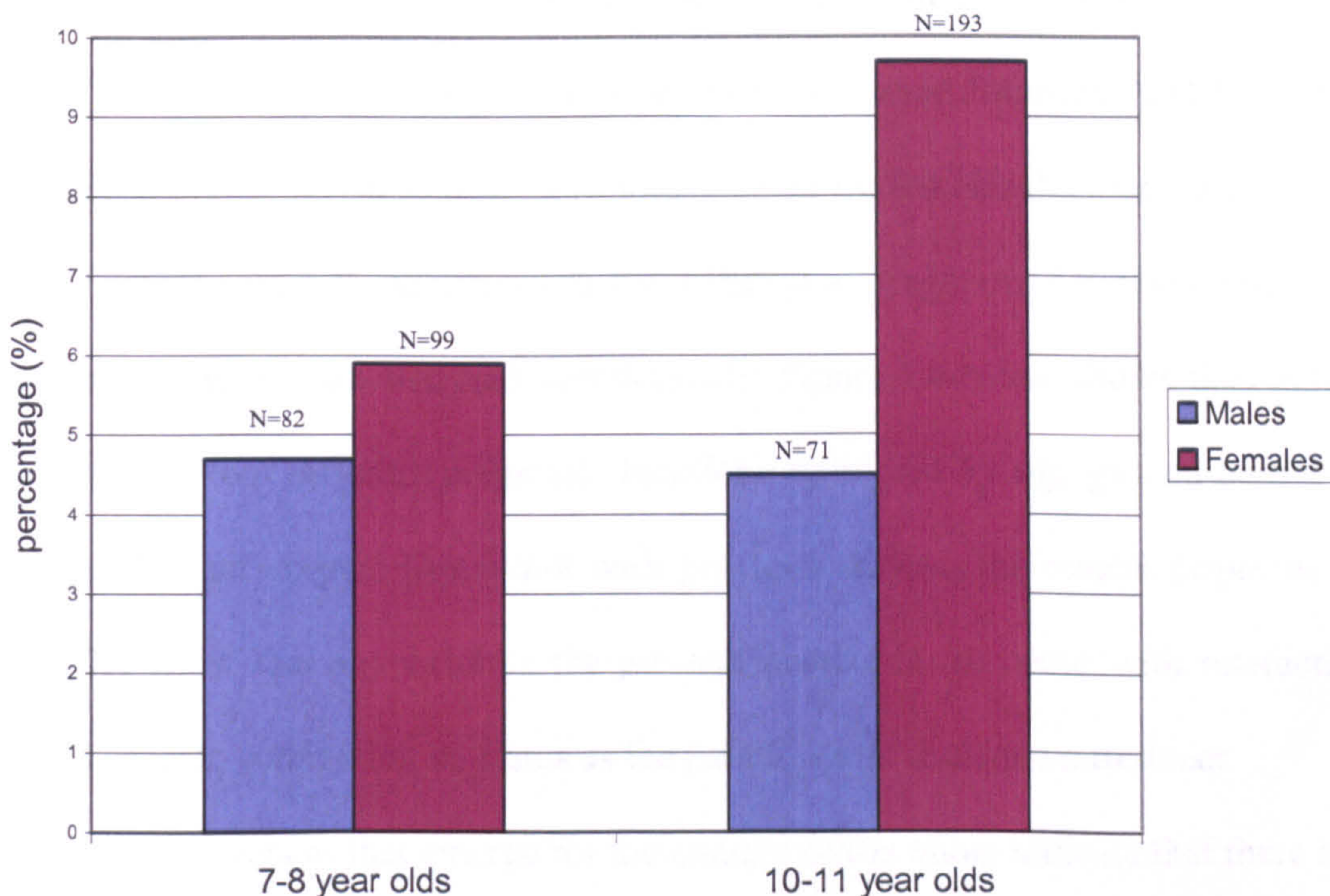
In order to explore how pragmatic differences in the use of quotative variants correlate with age and gender, I now turn to an analysis of the content of the quote. I first consider the most common discursive function of quotation in the corpus: the representation of direct speech. Figure 7.3 below depicts the frequency of direct speech in the corpus according to age and gender.

---

<sup>23</sup> Although boys make slightly greater overall use of *go* than girls, girls use *go* more frequently than boys to introduce direct speech (see Figure 7.5 below)

FIGURE 7.3

FREQUENCY DISTRIBUTION (PER 1000 WORDS) OF DIRECT SPEECH BY AGE AND GENDER



The girls in both age cohorts are ahead of the boys in their use of direct speech. Moreover, while there is a salient increase in the girls' overall use of direct speech with age, there is a slight decrease in the boys' use of direct speech as they get older. These findings exhibit interesting parallels with previous research. Macaulay (2001) found greater use of quoted dialogue in the speech of adolescent females than males. Out of 246 instances of reported dialogue, 76% occurred in the speech of female adolescents (Macaulay 2001: 8; see also Stenström *et al.* 2002: 126). Similarly, Ferrara and Bell (1995: 272) report that 74% of females included direct speech in their narratives in contrast with only 61% of males. Within the broader research literature, there is evidence that women have a greater propensity than men to report dialogue in their narratives. For example, Johnstone (1993: 73) observes that women in midwestern America report speech more often and at greater length in their stories

than men. Reports of speech were also found to occur more frequently in the extrathematic detail of women's narratives than in men's. This tendency also receives some support in the developmental literature, where, as mentioned earlier, it has been noted that females report speech more frequently than males at an early age (see e.g. Ely and McCabe 1993). In the preadolescent data, then, as in Johnstone's (1993) adult data, speech events appear to be a more salient detail for females than for males.

Further contrasts emerge when the differential pragmatic functions embodied by quotative variants are analysed quantitatively. Figure 7.4 below shows the relative proportions of four different pragmatic functions embodied by *say*, *go* and *be like* in the preadolescent corpus. Consistent with previous studies, the results displayed in Figure 7.4 show that *say* exhibits the greatest propensity to occur with interactive reported speech, confirming its status as the paradigmatic dialogue introducer.

The differences that emerge for the content of the quote indicate that there is a functional split between *say* and the other quotatives: whereas *say* is almost categorically restricted to verbally expressed quotations, the other two variants, *go* and *be like*, have a broader range of discursive functions.<sup>24</sup>

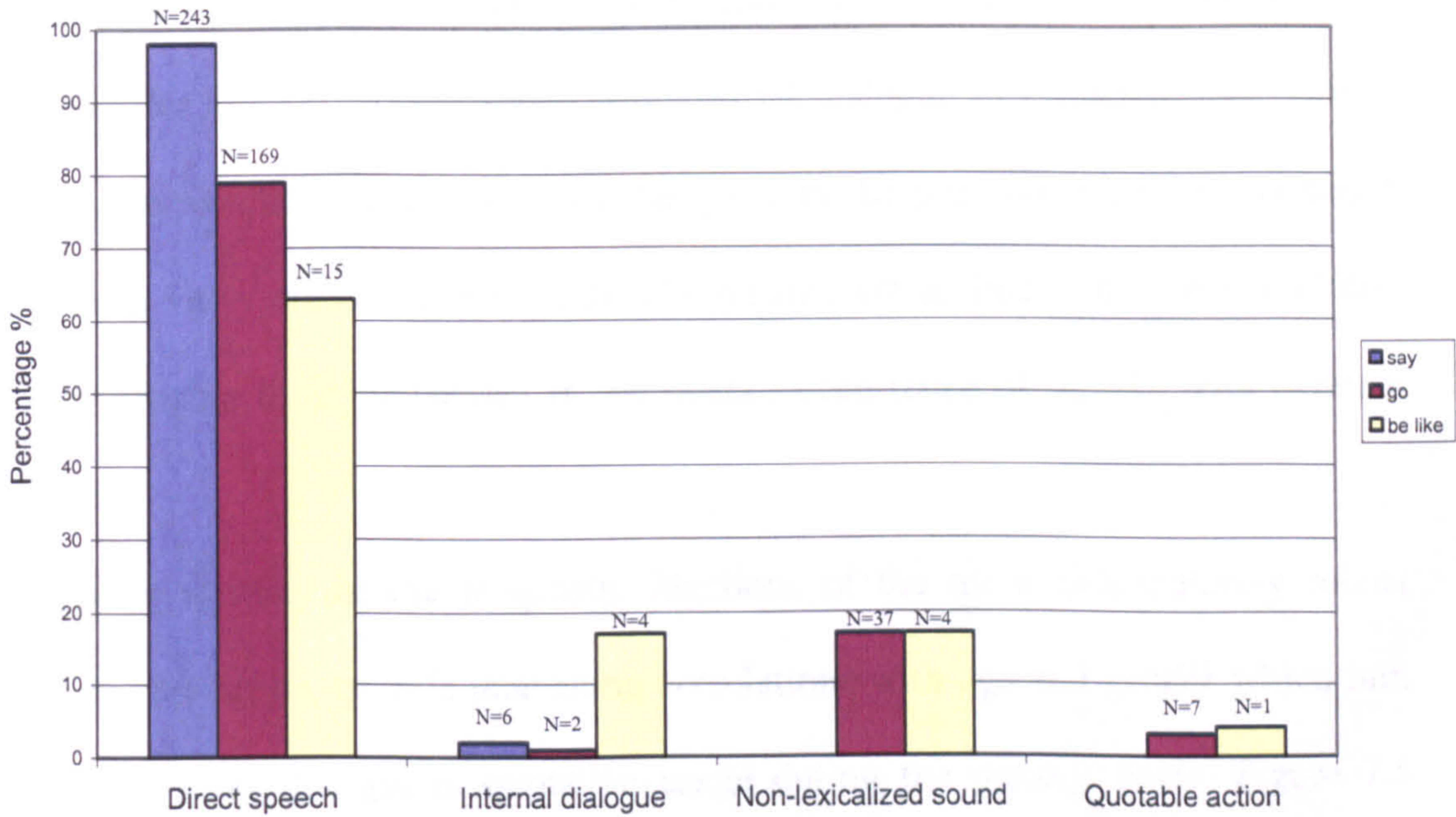
---

<sup>24</sup> Romaine and Lange (1991: 240) claim that quotative *go* contrasts with *be like* in that *go* 'has special connections with the auditory-vocal channel [...] and introduces sounds or onomatopoeic expressions rather than words'. Similarly, Lucy (1993:98) argues that a major pragmatic function of *go* is to 'foreground form - especially of non-speech sounds and non-referential speech.' While it is the case that the usage data depicted in Figure 7.4 illustrate some of the pragmatic functions embodied by *go* documented by Romaine and Lange (1991) and Lucy (1993), it can be seen that direct speech is proportionally the most felicitous context for the use of *go* in the preadolescent data. Barbieri's (2005: 248) large-scale survey of quotative use in American English also revealed that *go* was predominantly used to introduce direct speech.



FIGURE 7.4

PROPORTIONAL COMPARISON OF *SAY*, *GO* AND *BE LIKE* ACCORDING TO THE CONTENT OF THE QUOTE



In contrast with *say*, *go* and *be like* are more frequently used by the preadolescents to foreground mimetic enactments associated with the ‘delivery aspects’ of quoted material (see Clark and Gerrig 1990; Buchstaller 2003). The results show that there is some pragmatic overlap between the use of *be like* and *go*, particularly for quoting non-lexicalised sounds, as well as for reporting gestures. On the other hand, *be like* is differentiated from *go* and *say* in that it is used relatively more often as an introducer of internal dialogue (see Ferrara and Bell 1995: 279).

In spite of its very limited occurrence in the data, the distribution of *be like* across different pragmatic functions shown above offers some endorsement of Buchstaller’s (2004) suggestion that this variant functions as a ‘wild card’ quotative which can be used in a functional range of contexts. What is particularly striking in terms of pragmatic function in the preadolescent data is the use of *be like* to introduce

direct speech. Recall that according to Ferrara and Bell (1995: 279), the use of *be like* with external dialogue is developmental and diagnostic of its ongoing grammaticalisation within the quotative system. In contrast with the results obtained by Tagliamonte and Hudson (1999), who found that *be like* was preferred as an introducer for non-lexicalised sound and internal dialogue in Canadian and British English, the results for the content of the quote in the preadolescent data, although based on a very small number of *be like* tokens, are in line with Carey's (2004) findings relating to an expansion in the overall occurrence of *be like* with external dialogue.

Closer scrutiny of the pragmatic functions of the more quantitatively robust quotative variant, *go*, reveals interesting correlations with age and gender which hint at developmental changes in quotative usage during the preteen years. Figure 7.5 below shows that while the girls slightly increase their use of *go* with direct speech as they mature, this change is more salient for the boys than the girls. Note too that the different pragmatic functions of *go* are not evenly distributed across males and females, with the boys in both age cohorts making comparatively greater use of *go* to introduce non-lexicalised sounds and quotable actions.<sup>25</sup> This is particularly the case for the 7-8 year old boys whose use of quotative *go* to foreground non-lexicalised sounds competes very vigorously with their use of this form to introduce external dialogue. Nevertheless, both males and females decrease their use of *go* with non-lexicalised sounds and quotable gestures as they grow older, concomitant with a rise in the frequency of *go* to report external dialogue.<sup>26</sup>

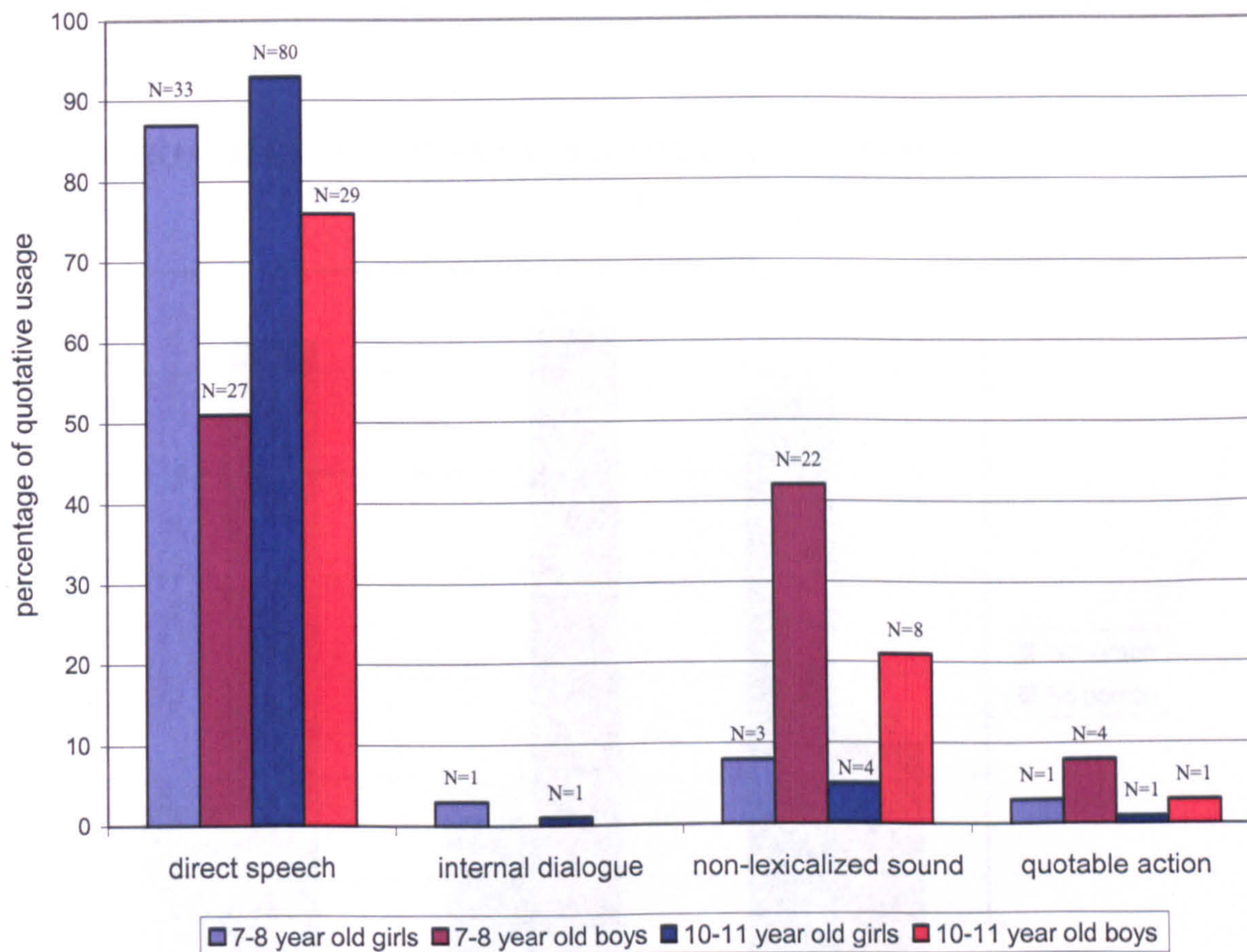
---

<sup>25</sup> The same effect can be discerned in the limited number of *be like* tokens in the youngest preadolescent cohort (7-8 years), with boys using this variant more frequently to report non-lexicalised sounds than the girls, who, in turn, use *be like* more often than the boys to foreground direct speech.

<sup>26</sup> The boys' tendency to use quotative *go* for mimetic performances may be indicative of more general stylistic differences between the preadolescent boys and girls. In this regard, it is interesting to note that Stenström *et al.* (2002: 114) found mimesis to be more a male than a female adolescent phenomenon.

FIGURE 7.5

FUNCTIONAL DISTRIBUTION OF *GO* BY AGE AND GENDER



7.6.3 Grammatical person

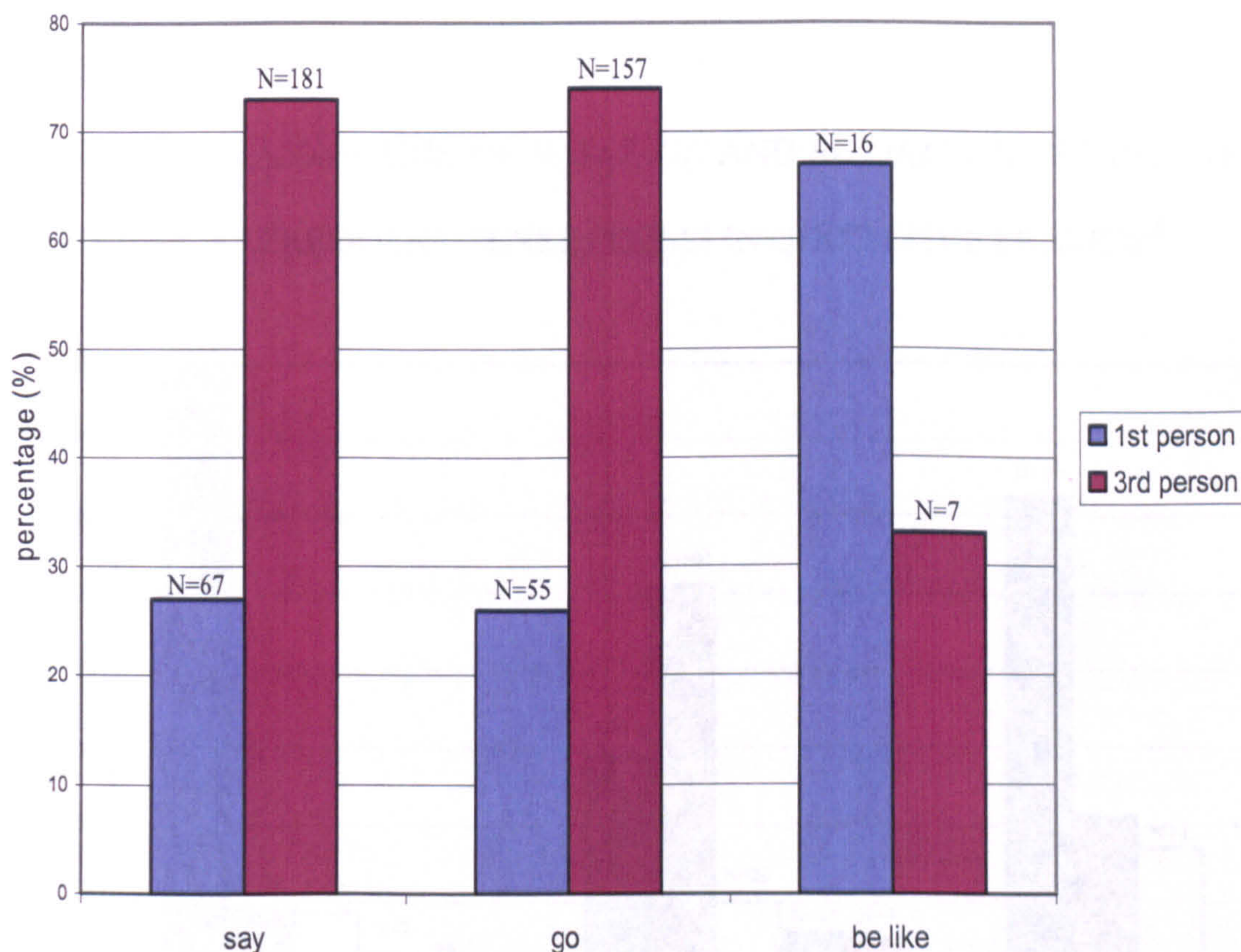
Figure 7.6 below shows that the distribution of *be like* across first and third person subjects is consistent with the findings of previous research (Carey 2004; Tagliamonte and D’Arcy 2004; Dion and Poplack 2005; Baker *et al.* 2006) which has noted the favouring effect of first person subjects on the use of *be like*.<sup>27</sup> Taking into account the restricted occurrence of this variant in the data, this conditioning effect tentatively suggests that the contextual constraints on the occurrence of *be like* in

<sup>27</sup> Although Tagliamonte and Hudson (1999: 169, f.n.10) observe that a highly favoured context for *be like* in British English is with existential *it*, there is only one such example in the preadolescent corpus: *it was like*, ‘There, that’s a big giant snowball’ [14M7/8]. In this particular extract, it is used to report the collective reaction of two boys to an unexpected event (see also Ferrara and Bell 1995: 278 on the use of *it’s like* to report the collective thoughts of a group).

preadolescent discourse, at least as far as grammatical person is concerned, demonstrate affinities with the constraints on *be like* reported for other varieties of British English (see Baker *et al.* 2006) as well as North American varieties of English.

FIGURE 7.6

DISTRIBUTION OF QUOTATIVES BY GRAMMATICAL PERSON<sup>28</sup>



With regard to the other quotative markers in the corpus, the results for *say* and *go* exhibit remarkably parallel profiles as far as the effect of grammatical person is concerned. Both quotatives occur predominantly in third person contexts, which is congruent with D'Arcy's (2004: 327) findings for Canadian English, and Barbieri's (2005: 244) results for American English, but which diverges from Carey's (2004: 8)

<sup>28</sup> A small number of occurrences of quotative variants with the second person have been omitted.

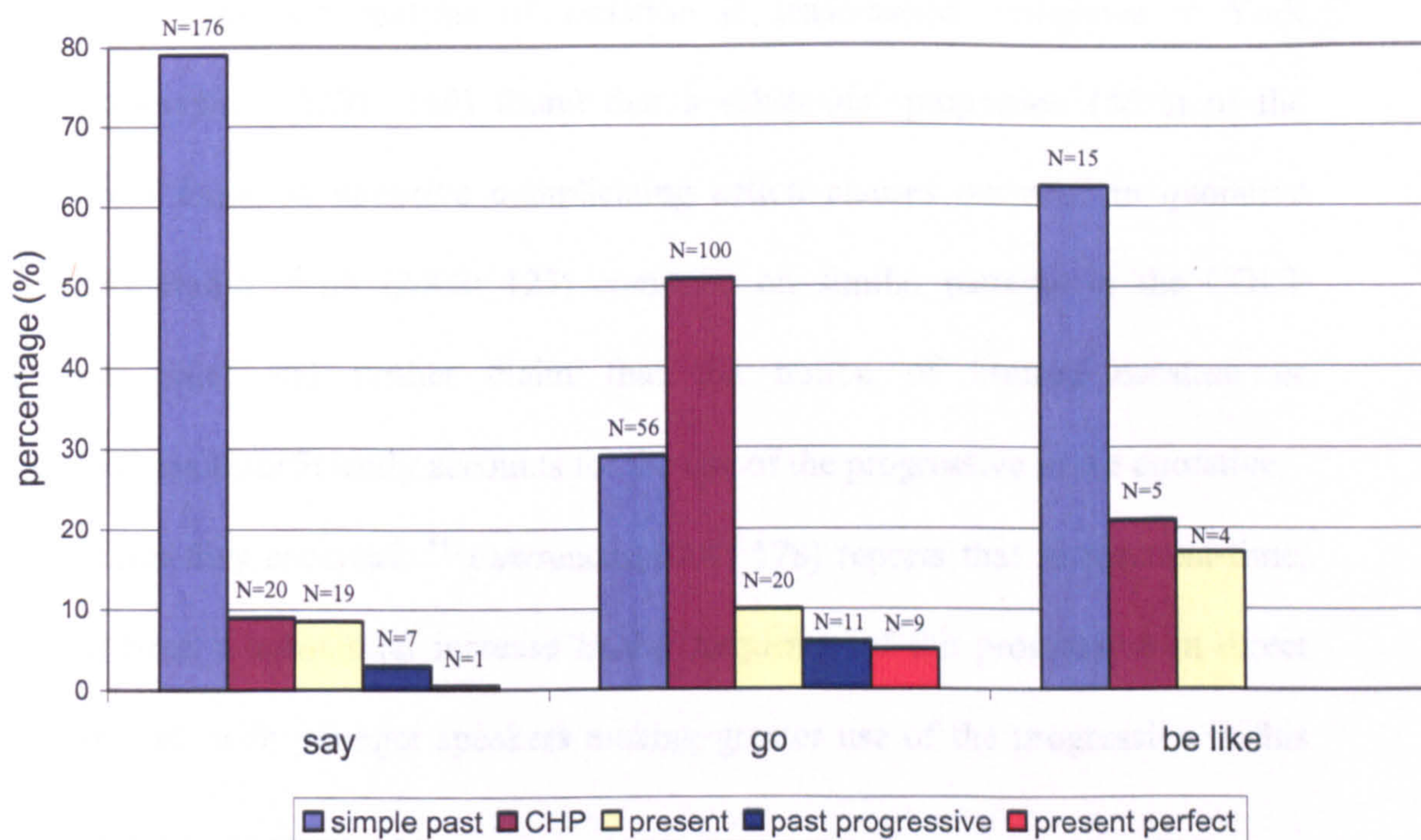
results based on British English, where *go* is reported to occur preferentially with first person subjects.

#### 7.6.4 Tense variation

Turning now to tense variation, Figure 7.7 illustrates that lexical quotatives are temporally marked by a range of forms.

FIGURE 7.7

PROPORTIONAL COMPARISON OF *SAY*, *GO* AND *BE LIKE* ACCORDING TO THE FIVE MOST FREQUENT TENSE FORMS IN QUOTATIVE FRAMES<sup>29</sup>



Contrary to previous research on North American English (see Blyth *et al.* 1990: 218; Singler 2001: 277; D'Arcy 2004: 335; Barbieri 2005: 222) which establishes that *be like* is more frequent in the present tense, Figure 7.7 shows that in the preadolescent data, *be like* occurs most often in the past tense. This finding dovetails with the results

<sup>29</sup> A small number of future tense as well as conditional forms have been omitted from the results displayed in Figure 7.7.

obtained by Macaulay (2001) and Carey (2004), who both report that *be like* occurred most frequently in simple past contexts in their British data.<sup>30</sup> *Say* also occurs most often in the simple past, whereas *go* shows a marked preference for the CHP (see also Chapter 5).

Although the incidence of *say* and *go* with the past progressive and the present perfect is comparatively low, similar patterns have been noted in other varieties of English (see Chapter 5; Winter 2002:11-16 on the use of the progressive and the present perfect with quotative verbs in Australian English; Ferrara and Bell 1995: 267 on the use of the progressive in quotative frames in American English).

In a large-scale analysis of variation in tense-aspect categories in York English, Lawrence (2001: 169) found that a substantial proportion (46%) of the progressives found in narrative complicating action clauses occurred in quotative frames. Stenström *et al.* (2002: 123) comment on similar patterns in the COLT adolescent data, and further claim that the notion of limited duration or incompleteness insufficiently accounts for the use of the progressive in the quotative constructions they analysed.<sup>31</sup> Lawrence (2001: 178) reports that in apparent-time, there has been a substantial increase in the frequency of the progressive in direct speech frames, with younger speakers making greater use of the progressive in this

---

<sup>30</sup> Singler (2001) alludes to stylistic considerations in his discussion of variation in the use of vernacular quotatives with particular tense forms. He argues that vernacular variants such as *go* and *be like* frequently co-occur with the CHP because 'both reflect greater informality' (2001: 271). Note, however, that in Chapter 5, I showed that preadolescents use the CHP at rates which are similar to those reported for adults. On the basis of the putative stylistic connection between the CHP and *be like* usage advanced by Singler (2001), we might expect *be like* to be far more robust in the preadolescent corpus, but no such correlation obtains in the data (see also Tagliamonte and D'Arcy 2007).

<sup>31</sup> Previous attempts to account for the use of the progressive in narrative clauses containing a quotative frame have not been entirely convincing. Fleischman (1995: 536) suggests that such uses may have a ruminative function; that is, the progressive highlights something which the speaker is turning over in her/his mind. Similarly, Bybee (p.c. cited in Fleischman 1995: 536-7) proposes that the continuous meaning encoded by progressives may be used to highlight an ongoing, possibly backgrounded, mental attitude. Neither of these theories exhaustively accounts for all uses of the progressive in quotative contexts in either the preadolescent corpus, or the COLT examples discussed in Stenström *et al.* (2002: 123), where there is often no discernible emphasis on either iterativity or extended mental rumination. In the preadolescent data, progressive forms in quotative frames largely occur in semelfactive contexts.

context than older speakers. This finding is consistent with the general observation that the progressive has become more frequent over the past 30 years and is competing with non-progressive forms (see Mair and Hundt 1995: 111).

### 7.6.5 *Multivariate analysis*

I now turn to a multivariate analysis of the data. Owing to the marginal status of quotative *be like* in the quotative cohort, and its uneven distribution across different contexts, this variant has been excluded from the multivariate analysis. Similarly, other variants such as the mixed forms *say like/ go like*; the zero quotative; and the graphic dialogue introducers *shout/shout out* have been omitted from the multivariate analysis as they account for a negligible proportion of overall quotative usage. In addition, certain factors have been collapsed or excluded on account of either extreme infrequency or poor data distribution in order to allow the variable rule analysis to run (see Tagliamonte 2006a: 210). For example, in the case of the factor group relating to the content of the quote, I have collapsed the use of quotatives to foreground internal dialogue, non-lexicalised sounds, and quotable gestures into one factor, which is contrasted with the use of quotatives to introduce direct speech (see also Barbieri 2005: 235-6 for a similar approach).

Table 7.2 below shows the results of two separate multivariate analyses of factors contributing to the use of *say* and *go* in the preadolescent corpus.

TABLE 7.2

TWO SEPARATE MULTIVARIATE ANALYSES OF FACTORS CONTRIBUTING TO THE SELECTION OF *SAY* AND *GO* IN THE PREADOLESCENT CORPUS<sup>32</sup>

INPUT	SAY		GO		Total N
	FW	%	FW	%	
<b>Gender</b>					
Male	0.49	51	0.51	49	183
Female	0.51	56	0.49	44	276
<b>Age</b>					
7-8	0.57	55	0.43	46	201
10-11	0.44	53	0.56	47	258
Range	13		13		
<b>Grammatical Person</b>					
First	0.47	54.5	0.53	45.5	121
Third	0.51	54	0.49	46	338
<b>Content of the Quote</b>					
Direct Speech	0.58	59	0.42	41	407
Other	0.07	11.5	0.93	88.5	52
Range	51		51		
<b>Tense</b>					
Past	0.74	76	0.26	24	232
CHP	0.13	17	0.87	83	120
Present	0.45	49	0.55	51	39
Past Progressive	0.44	39	0.56	61	18
Range	61		61		

The most significant factor group, as indicated by the range, is tense. The results show that there are notable correlations between quotative variants and particular tense forms, with the vernacular quotative *go* strongly favouring the use of the CHP

<sup>32</sup> Note that the ranges for those factors which are selected as significant are the same for *say* and *go*. As the results are generated by two separate analyses which involve only two variants, with *say* run first as the application value followed by *go*, the variants are essentially 'two sides of the same coin.' I thank David Britain (p.c. March 2007) for discussion of this point.



and, to a lesser extent, the past progressive and the present, whereas *say* only favours the use of the simple past.<sup>33</sup>

The second most significant factor group, the content of the quote, highlights the functional asymmetries of *say* and *go*, with *say* favouring direct speech in contrast with *go*, which is strongly favoured for other quotative functions (i.e. introducing internal dialogue, non-lexicalised sounds and quotable gestures). Thus, in line with previous research, *say* is used to 'present an icon of a speaker's actual words' (Lucy 1993: 96), whereas *go* is used to foreground non-referential speech and gestural effects.<sup>34</sup>

The weakest statistically significant effect is associated with age. Although *go* is used by the younger children in the preadolescent cohort, it is the older children who favour this vernacular variant. The favouring effect for the older children may reflect convergence towards the focused vernacular norms of the peer group which emerge as children negotiate the transition from parent-oriented to peer-oriented networks during the preadolescent years (Kerswill 1996: 196).

The lack of a statistically significant correlation between gender and choice of quotative confirms that neither *say* nor *go* embodies any particular sociosymbolic function for either the girls or the boys in the corpus. In fact, there is evidence that gender-related constraints on quotative use are developmental: in their study of quotative variation in the speech of Canadian youth, Tagliamonte and D'Arcy (2004: 510) report that gender-related differences in quotative usage were highly salient in only the speech of 15-16 year olds.

---

<sup>33</sup> Although the present and the CHP are morphologically identical, I have kept contexts where the present has past temporal reference (i.e. CHP contexts) separate from other uses of the present to facilitate comparison with earlier studies.

<sup>34</sup> A similar functional dichotomy between quotative forms exists in Swedish where *säga* 'say' is a pragmatically neutral quotation marker, whereas the vernacular variant *ba* 'just' conveys particular pragmatic effects associated with the expression of emotional attitude, sound effects, and internal thought (Eriksson 1995).

## 7.7 SUMMARY

The examination of quotative variation in the corpus has highlighted that the quotative system employed by the preadolescents is largely split between *say* and *go*, with the remaining variants occupying a trivial niche within the pool of available forms. Thus, there is little evidence in the preadolescent data examined here that *be like* is expanding at the expense of the other major variants, *say* and *go*. The vernacular quotative, *go*, remains a robust variant, and even appears to be favoured as the children get older. Just how far the patterning of *go* in the preadolescent data is consistent with the wave-hypothesis advanced by Buchstaller (2006:22) according to which *go* sometimes advances and sometimes retreats in apparent-time lies beyond the confines of this study, although judging from the evidence presented here, London preadolescents are peak users of quotative *go*.

In this chapter, I have also shown that when the specific pragmatic functions of *go* are taken into consideration, age- and gender-related differences emerge in the distributional patterns exhibited by this variant. Of particular interest in this chapter is the boys' reduction in the use of *go* to introduce non-lexicalised sounds as they get older, and the concomitant age-related increase in their use of *go* to introduce direct speech.<sup>35</sup>

From a more general perspective, the quantitative analysis of the pragmatic functions of the quotatives employed by the preadolescents revealed that girls were

---

<sup>35</sup> Whether this phenomenon may be linked to the developmental imperative discussed by Eckert (1997a) warrants further investigation. According to Eckert (1997a:158), 'an emphasis on growing up, on being age-appropriate [...] is particularly intense and foregrounded in the early years.' If, as Romaine and Lange (1991: 267) claim, a heavy reliance on paralinguistic and prosodic devices are stylistic features which figure prominently in the early stages of language acquisition, then with increasing maturity, children may reduce their use of *go* with non-lexicalised sounds and other paralinguistic features to avoid sounding childlike. Thus, even within the narrow age range explored in this chapter, there is additional scope for exploring the link between the emergence of age-specific variation in quotative usage and social meaning.

more liable to foreground conversational interaction in their speech than boys, which corroborates similar gender-related patterns reported in previous research.

Although the preadolescents' use of *be like* exhibits interesting parallelisms with other varieties in terms of its expansion into direct speech contexts, its limited productivity in the age range examined here vitiates any claims that can be made about its functional trajectory in London vernacular English. Obviously, its infrequency in the preadolescent data provokes questions of a broader nature about the differential rates of diffusion of this innovation in different varieties of English and the social mechanisms which are constraining its dissemination. The fact that *be like* is amply attested in the vernacular of Canadian preadolescents, as demonstrated by D'Arcy (2004), indicates that age alone cannot exclusively account for its low rate of occurrence in the London data.

The use of *be like*, although very restricted, by children as young as seven in the preadolescent corpus furnishes preliminary evidence that it has percolated to a limited extent down the age spectrum to preteen speakers in London.

### VARIATION IN THE USE OF GENERAL EXTENDERS

#### 8.1 INTRODUCTION

In this chapter, I present a quantitative and qualitative analysis of tags which typically occur clause-finally. Examples of these tags taken from the preadolescent corpus of recordings are given below:

- (1) he was having a wrestling match with me and punching me and kicking  
*me and everything* [19F7/8]
- (2) they go to church *and all that* [13M7/8]
- (3) they found some sheep where there was these dogs chasing them  
*and stuff like that* [8F10/11]
- (4) they teamed up on to him and started hitting him with weapons  
*and things* [4M10/11]

The clause-final tags italicised above have been variously characterised as set-marking tags (Dines 1980), extension particles (Dubois 1992), vague category identifiers (Channell 1994), generalized list completers (Jefferson 1990), and co-ordination tags (Biber *et al.* 1999). Following Overstreet (1999), I refer to them as 'general extenders' in order to avoid imposing a term which implies that these tags are confined to specific functions such as set-marking or list-completion. In fact, as I show below, general extenders are typically multifunctional, making it difficult at times for the analyst to determine which of several possible functions is being foregrounded in a specific extract of discourse (see also Cheshire *in press*).

In this chapter, I investigate the frequency, distribution, and function of general extenders in preadolescent speech, and specifically explore the extent to which patterns of distribution correlate with age and gender. Furthermore, I investigate textual constraints on the use of general extenders, especially with reference to their use in narrative discourse.

I first contextualize the discussion of variation in the use of general extenders in preadolescent speech by drawing on other studies of these structures in other varieties of English. I also briefly consider research on other languages in which there are structural and functional equivalents to general extenders.

## 8.2 GENERAL EXTENDERS IN OTHER VARIETIES OF ENGLISH AND IN OTHER LANGUAGES

### 8.2.1 *Cross-variety perspectives on general extenders*

According to Dines (1980: 18) general extenders in English can be structurally characterised as adhering to the pattern AND/ OR [PRO FORM] (LIKE THAT). They are semi-fixed sequences that contain slots for lexical material to be either inserted or deleted. For example, the comparative element *like that* may not always be present, and even the conjunction may very occasionally be deleted (see below; Overstreet 1999: 11; Cheshire *in press*). Overstreet (1999: 143-144) notes that they 'typically occur in clause-final position, are non-specific in reference and extend otherwise grammatically complete utterances.' They are found in a variety of genres including written and spoken academic discourse (see Ruzaitė 2004), although some variants such as *or so*, *and so on* are more commonly encountered in written registers than in speech (Overstreet 1999: 5). They are most frequently encountered in informal conversation between familiar interlocutors (Overstreet 1999: 6; Dubois 1992: 198).

Overstreet (2005: 1846) classifies general extenders as a subset of discourse markers, and Dubois (1992: 181) also notes that they have discourse-organisational functions which make them comparable in many respects to other discourse markers. However, as can be seen from examples (1)-(4) above, they are characteristically less syntactically mobile than other discourse markers such as *you know* or *like*; they do not occur clause-initially, nor are they necessarily phonologically reduced (see Brinton 1996: 33-5). Nevertheless, as I shall illustrate below, they are optional, and, in common with other discourse markers, they can operate within several discourse components including the ideational, textual and interpersonal levels of discourse (see Schiffrin 1987).

General extenders have been noted in several varieties of English including British English (Aijmer 1985, 2002; Channell 1994); Scots (Macaulay 1985, 1991); American English (Ball and Ariel 1978; Jefferson 1990; Overstreet 1999); New Zealand English (Britain 1992; Stubbe and Holmes 1995); as well as in Australian English (Dines 1980; Winter and Norrby 2000; Norrby and Winter 2002). Even within specific varieties, however, there appears to be a great deal of variation that remains unexplored.

Cheshire *et al.* (1999) note that there may be regionally constrained variation in the form of general extenders. According to Cheshire *et al.* (1999), in Hull, negative concord does not affect tags appended to clauses such as *they do not know what they're doing or anything*, whereas in Reading and Milton Keynes, negative forms such as *or nothing* occur variably in the clause-final tag.<sup>1</sup> Cheshire's (*in press*) recent study of general extenders used by adolescents in Hull, Reading and Milton

---

<sup>1</sup> In terms of the variable effects of negative concord, the London preadolescent data pattern with Milton Keynes and Reading: both *or anything* and *or nothing* occur in general extenders appended to clauses with negative polarity. There are similar comparisons to be drawn with the *Newcastle Electronic Corpus of English* (NECTE) data below; see especially examples (6) and (7).

Keynes highlights differences in the frequency of common variant forms in these geographical locations. For example, the variant *and things* is used less frequently in Hull than in either Reading or Milton Keynes, suggesting, possibly, that there is a north-south division in the distribution of certain types of general extender.

Other forms appear to be dialectal variants that are specific to certain areas and do not have wider geographical currency. Consider the following examples from NECTE:

(5) we're trying to get it through the door you know into the kitchen eeh I

never lifted the carpet *or nowt like that*

[male/aged 51-60/lower-middle class]

(6) do you not get funny looks *or owt like that*

[male/aged 16-20/working class]

(7) she didn't say, 'Keep the change' *or nowt*

[female/aged 16-20/working class]

Examples (5) to (7) show that these forms are used by both males and females from different social classes. However, the extent to which forms such as *or nowt* /*or owt* are grammatically constrained, or sociosymbolically emblematic of local identities, awaits further investigation.

### 8.2.2 *General extenders in other languages*

Cross-linguistic evidence reveals that there are structural equivalents of general extenders in a range of languages, suggesting that these constructions are functional, and play a role in the construction of coherent discourse as well as in the negotiation of social meaning. Cross-linguistic comparisons are, however, somewhat complicated by the fact that general extenders are highly formulaic (Overstreet 2005),

although formal similarities are discernible in typologically affiliated languages. Examples (8) to (12) illustrate clause-final tags in a range of languages. Structures which are comparable to general extenders in English are italicised:

(8) Swedish

vi satt å prata *å så där*

We sat and talked *and that*

(Aijmer 1985: 389; cited in Overstreet 1999: 11)

(9) French

Je jouais au hockey pour le collège, au hockey-salon au: au hockey

intérieur tu sais *des affaires de même*

I played hockey for the college, floor-hockey, inside hockey, you know,

*things like that*

(Dubois 1992: 180; cited in Overstreet 1999: 9)

(10) Brazilian Portuguese

Tava jogando futebol de salão *e tal*

He was playing indoor soccer *and all*

(Roth-Gordon 2001: 130)

(11) Japanese

naraigoto mo hajimerareta shi, *toka*<sup>2</sup>

and I could also start lessons, *and stuff like that*

(Lauwereyns 2002: 244)

---

<sup>2</sup> Elsewhere in Lauwereyns' discussion of *toka*, this particle seems to resemble the pragmatic marker *like* (see Chapter 6) in English. *Toka* can also be used in quotative contexts; e.g. *Okane nai kara toka itte I go like 'cause we have no money'* (Lauwereyns 2002: 254). Romaine and Lange (1991) also note that in some of its uses in English, *like* seems to behave in a similar way to structures which Dines (1980: 22) calls 'terminal tags' (i.e. general extenders).



(12) Arabic

niswān u-`akil u-kulšī

It was women and food *and everything*

(Farghal and Haggan 2005: 401)

(13) Hebrew

tás leshavà, ... xamishá yamim, *máshehu kaze*

[he] flew for a week, five days, something like this

(Maschler 2001: 308)

### 8.3 PREVIOUS RESEARCH ON THE SOCIOLINGUISTIC DISTRIBUTION OF GENERAL EXTENDERS

A number of published studies (e.g. Dines 1980; Dubois 1992; Stubbe and Holmes 1995) indicate that general extenders are socially diagnostic, and are differentially distributed among different social classes. Dines (1980: 18) found that in terms of frequency, general extenders were more commonly used by working-class speakers than by middle-class ones.

According to Dubois (1992: 199), the lexical elements which comprise general extenders may also have different social connotations. Dines (1980) mentions that the form *and that* appears frequently in the speech of working-class speakers in her Australian data, and Britain's (1992) study of general extenders based on data collected in New Zealand reveals that working-class speakers use *and that* far more often than middle-class speakers, who, in turn, use the form *or something* more often than working-class speakers. Cheshire's (*in press*) study of British adolescent speech also revealed social class differences in the use of certain variants, with *and that* favoured by adolescents from working-class backgrounds, and *and stuff* and *and*

*things* preferred by middle-class teenagers. Cheshire (*in press*) adds that the social-class affiliated uses of particular variants may indicate that certain forms function as social indicators for speakers. From an evaluative perspective, Dines (1980: 19-20) reports that both working-class and middle-class speakers do, however, converge in reacting negatively towards speech samples containing general discourse extenders, with adverse reactions based on the perceived association between the use of general extenders and vague or inexplicit speech.

Studies which explore the interaction between general extender usage and social class also reveal that the latter is intertwined in complex ways with other social parameters such as gender and age. Although Stubbe and Holmes (1995: 75) found, as far as overall frequencies were concerned in their New Zealand data, that there was no significant difference in the use of general extenders by young working- and middle-class speakers, they note that there was an interesting interaction between social class and gender, with working-class male speakers and middle-class female speakers producing the highest frequencies of general extenders. Britain (1992) observes that women used general extenders considerably more often than men in the New Zealand corpus he examined.<sup>3</sup> Furthermore, in terms of gender-affiliated preferences for certain forms, Britain's (1992) data also show that young female speakers used *and that* more than males.

Meyerhoff (1992: 9) also considers gender-specific differences in her discussion of pre- and post-nominal hedges in the speech of New Zealanders and

---

<sup>3</sup> Britain's (1992) findings are based on data collected from young, middle-aged and old speakers. Furthermore, speakers differed in terms of ethnicity (i.e. Maori or Pakeha backgrounds). In Britain's data, males receive an index score of 23.19 for general extender usage in contrast with females, who have a score of 43.14. The index score is calculated by dividing the total number of general extenders for each sex by the number of minutes of recording for men and women respectively. I am extremely grateful to David Britain for the use of this unpublished data.

suggests that forms such as *or something* are more significant in women's talk than men's, where, it is claimed, they may function as affiliative or in-group markers.

I return to the issue of gender differences in relation to the use of general extenders in the discussion of the distribution of these forms in the preadolescent corpus below. For the moment, it is important to note that previous research reveals that there is no simple relationship between gender and the use of general extenders. This point is foregrounded in the comparative study of general extender usage in two large corpora of Canadian French undertaken by Dubois (1992). Dubois (1992: 198-9) discovered that women used more general extenders than men in the 1971 corpus she examined, although this pattern was no longer evident in the second corpus compiled in 1984. Similarly, in the *Bergen Corpus of London Teenage Language* (COLT), Stenström *et al.* (2002 :102) report that although female adolescent speakers use general extenders slightly more frequently than male adolescents, the overall gender difference is marginal.

One of the reasons why studies report varying frequencies of general extender usage relates to divergent methodologies in collecting spoken data, as well as varying social and situational factors which impinge on the nature of the data collected. Schiffrin (2001: 66) stresses that if discourse marker usage reflects the underlying cognitive, expressive, textual and social organization of a discourse, then the properties of the discourse itself, which are contingent upon a speaker's goals and social situation, will necessarily affect possibilities for marker usage. Thus, when recorded data differ in terms of topics, social distance between speakers as well as a number of other social parameters, there may be significant disparities between data sets with regard to the frequency of occurrence of discourse-pragmatic features. Such

disparities have a crucial bearing on the extent to which discourse features in different data sets may be reliably compared (see also Cheshire *in press*).

Turning to correlations between age and general extender usage, it has been repeatedly found in previous studies that young people are heavy users of these forms. Dubois (1992: 179) reports that in her study 'the clearest result pertaining to overall rates of [general extender] use is an age-grading one, whereby speakers use fewer particles as they get older'. According to Dubois (1992: 185), the youngest speakers use these particles most frequently, with a downward trend in usage after the teen years. Similar results corroborating high frequencies of usage in youth are reported in other studies: Stubbe and Holmes (1995: 72) observe that young speakers use general extenders approximately twice as often as middle-aged speakers. Further confirmation of this age-related trend is found in Britain's (1992) data, in which young people use general extenders more often than middle-aged or old speakers.<sup>4</sup>

It is interesting too that similar patterns of social distribution have been reported for other languages which have pragmatic markers that are functionally equivalent to general extenders. Lauwereyns (2002: 246) notes that the Japanese particle *toka* (see example 11 above), which is pragmatically comparable to forms such as *or something, and stuff like that*, is approximately 3.5 times more frequent in the speech of younger Japanese speakers than in older speakers. Lauwereyns (2002: 254) accounts for the higher frequency of *toka* in youth speech in terms of its solidarity-enhancing potential for expressing an age-group identity centred on youthfulness. Winter and Norrby (2000) make a similar appeal to the affiliative or solidarity-based functions of general extenders employed by Australian and Swedish-

---

<sup>4</sup> Britain's (1992) index scores for general extender usage according to speaker age are as follows: young = 47.09; middle-aged = 27.01; old = 32.88.

speaking adolescents, who, they claim, use general extenders to enact adolescent youth identities (see also Chapter 6).

Before investigating in more detail how general extenders may be conscripted for social purposes, it is first necessary to review previous research on the function of these forms in discourse.

#### 8.4 THE FUNCTIONS OF GENERAL EXTENDERS

According to Overstreet (1999: 11), previous research has been preoccupied with explaining the referential functions of general extenders (see e.g. Dines 1980; Ward and Birner 1993; Channell 1994). Dines (1980: 22), for example, claims that general extenders 'cue the listener to interpret the preceding element as an illustrative example of some more general case.' This function is illustrated in example (14) below taken from the preadolescent corpus, where the general extender *and all that* refers anaphorically to a set containing the preceding nominal, *a duck*, and instantiates an illustrative example of an inferable broader category:

(14) there's like other animals lives there... a duck *and all that* [10F10/11]

By appending a general extender to a particular item, the speaker assumes the listener can identify the category of items that the speaker intended to refer to (Overstreet and Yule 1997: 251). As Dubois (1992: 181) observes, general extenders can serve as intermediaries between a set of elements and the concept that the speaker has just expressed.

The category-implicative functions of general extenders have given rise to their being characterised as set-marking tags (Dines 1980). This characterisation also features prominently in Jefferson's (1990) discussion of the role of general extenders in list construction. Jefferson (1990: 66) claims that in compiling a list, speakers tend

to construct normative sequences consisting of three parts as in *she bought apples, oranges and stuff like that*, where the general extender functions as a 'generalised list-completer' which implicates the existence of further instances of an enumerable set.

Although general extenders can undoubtedly have category-implicative functions in discourse, one of the main criticisms levelled at this interpretation of their use is that this function is not necessarily prominent in all instances. Channell (1994: 140-1) discusses the problematic issues raised by the following example she recorded in relation to news about the birth of a baby:

(15) it's a boy or a girl *or something*

As Channell (1994) points out, both members of the set of possible human offspring are mentioned prior to the addition of the general extender, which means that there are, in principle, no further possible members of the set which can be implicated by the use of the general extender (see also Macaulay 1985: 113-4; Macaulay 1991: 61; and Aijmer 2002: 215 for additional problems with category implication associated with general extenders). In the example cited above, the addition of the general extender *or something* does not have obvious category-implicative functions, but appears instead to be operating as marker of negative politeness enabling the speaker to foreground her or his uncertainty about the sex of the baby (see Overstreet 1999: 92).

There are similar problems with viewing general extenders as list-completers. Corpus-based studies do not corroborate Jefferson's (1990) hypothesis that general extenders are typically integrated into three-part lists. For example, Overstreet (1999:

25) notes that only 21% of the general extenders in her corpus were used to complete three-part lists, with 74% of general extenders occurring after only one item.<sup>5</sup>

The inference of additional members of a category implied by the use of a general extender also crucially depends in many cases on common cultural scenarios and shared social knowledge between interlocutors (Dubois 1992: 182). A particular case in point is where general extenders are appended to items in discourse to create ad hoc or non-lexicalised categories (Barsalou 1983: 211). Overstreet (1999:43), for example, found that the vast majority of general extenders in her data were used to implicate non-lexicalised categories rather than lexicalised ones. Given that categorization can vary across individuals, subgroups and cultures (Schiffrin 1994: 307), the reference of general extenders appended to non-lexicalised categories can frequently be subjective and highly context-dependent.

The knowledge which hearers need to draw on in order to interpret what may be implicated by the use of a general extender can range along a continuum which encompasses both general cultural and social knowledge shared by many, as well as knowledge that is restricted to a small number of participants (Overstreet 1999: 69; Roth-Gordon 2001). However, as Overstreet and Yule (2002: 787) point out, even though discourse participants may not necessarily share an identical range of reference when using a general extender, the lack of compatibility of reference does not necessarily impede successful communication. As Östman (1981: 19) notes in relation to the discourse marker *you know*, which can also be used to indicate knowledge shared between speaker and hearer (Schiffrin 1987: 268), 'it is the pretence of shared knowledge on the part of the speaker that achieves intimacy and facilitates verbal interaction.' General extenders can similarly function in the capacity

---

<sup>5</sup> It is interesting that similar results are reported for other languages. According to Shinmura (1998: 1901), the Japanese particle *toka* (see example 11) is typically used to provide one example while omitting the rest.

of what Schiffrin (1987: 284) refers to as markers 'of information state transition' by signalling that the speaker has meta-knowledge of knowledge which is assumed to be shared with the addressee (Schiffrin 1987: 269). This is a function that is common to a number of other discourse markers (see e.g. Meyerhoff 1994 on the tag particle *eh* in New Zealand English). From a social perspective, then, what is popularly construed as inexplicit or vague speech can fulfil important interactional functions (see Jucker *et al.* 2003). Tannen (1982:18) argues that non-autonomous language, which is characteristically associated with the properties of the real-time dynamics of spoken language, deliberately builds on interpersonal involvement and demands maximum contribution from the audience in supplying socio-cultural and contextual knowledge. Lack of explicitness in informal discourse can therefore be interactionally advantageous: being understood without explicitly stating what one means contributes to a sense of involvement through mutual participation in sense-making (Tannen 1989: 23).

To illustrate some of these points more clearly, consider the following example from the preadolescent corpus where a ten year old girl is commenting on her perception of some of the differences between boys and girls:

(16) well normally like boys think that football's a boy game cos like

David Beckham *and all that* are boys [20F10/11]

In this case, the phrase *David Beckham and all that* is used to supply an illustrative example of a broader category (i.e. football players). Furthermore, in order to communicate her meaning successfully, the speaker in the above example is assuming that the addressee has specific shared knowledge of who David Beckham is (i.e. a famous English footballer). Overstreet and Yule (1997: 250) argue that a central function of general extenders is to serve as markers of intersubjectivity through which



speakers can index assumptions of conjoint knowledge, shared experience and social connection. The example above serves to reiterate a point made earlier, namely, general extenders, like other pragmatic particles, can fulfil a number of different discourse functions, sometimes simultaneously.

Note too in the above example the lack of grammatical agreement between *and all that* and the proper noun, *David Beckham*, to which the general extender anaphorically refers. Pawley and Syder (1983: 560) observe a similar use of the general extender *and them* in New Zealand English, which is apparently employed as 'a marker of unspecified associates' after proper nouns and other definite NPs denoting a person. Pawley and Syder (1983: 561) add that another marker, *and that*, is used in New Zealand English with reference to the unspecified associates of an inanimate definite NP (e.g. *I put it with the cutlery and that*). What is interesting about example (16), however, is that, in contrast with the examples from New Zealand English, the general extender *and all that* is used to refer to unspecified associates of an animate proper noun rather than an inanimate NP, possibly indicating that this variant has undergone, or is in the process of undergoing, decategorialization. It is not difficult to find other examples in the preadolescent corpus where the head noun in a general extender either fails to show syntactic or semantic characteristics which match those of a preceding nominal, as well as cases where the head noun in the general extender refers anaphorically to either a non-nominal referent, or even a preceding predicate:

(17) it is worrying me to go round with all the older children... the fifteen-year-olds *and things* [16F10/11]

(18) because we've been to MacDonald's on our own, haven't we ? [15F10/11]

<unclear> with my sister *and everything* [16F10/11]

(19) the police has found him *and everything* [19M7/8]

(20) they're really silly *and everything* [7F6]

(21) like he was crying *and stuff* [17F10/11]

In (17), *and things* occurs with an animate plural noun; and in (18), *and everything* is appended to an animate singular noun. In (20), *and everything* may refer to the preceding adjective, or it may have wider scope over the entire preceding predicate. This seems to be the case too in (21), where instead of appearing with a preceding mass noun (see Dines 1980), *and stuff* is appended to a verb phrase. These collocational patterns suggest that the lexical constituents of certain general extender variants appear to have lost- or are in the process of losing- semantic and syntactic properties, resulting in their becoming part of desemanticised phrases which are grammaticalising as single processing units (see Bybee 2003: 603; Aijmer 2002: 223; and Cheshire *in press*). According to Traugott (1995), structural decategorialization can result in increased bonding within a grammaticalising construction as well as increased pragmatic force and semantic abstraction. These processes may in turn lead to greater syntactic freedom. Additional evidence that certain general extenders instantiate some of the structural and pragmatic properties correlated with the characteristics of grammaticalisation can be found in Overstreet (1999) and Cheshire (*in press*). Overstreet (1999: 104) discusses instances in her database where the general extender *and stuff* is not attached to propositional information and appears to have migrated from clause-final position to other syntactic locations. Cheshire's (*in press*) analysis of British adolescents' use of general extenders revealed that *and that* and *and everything* appear to be the most grammaticalised variants. Other variants such as *and stuff* and *and things* were found to be less advanced in terms of

grammaticalisation processes. Different variants may therefore be situated at different points along the cline of grammaticalisation.

In the remainder of this chapter, I look at some of the structural properties and distributional features of general extenders in the preadolescent corpus. Before I discuss the distribution of general extenders in the data, I first outline the methodology employed to delimit the envelope of variation.

## 8.5 METHODOLOGY

One of the major methodological problems encountered in quantitative analyses of discourse variation, as I have already highlighted in Chapter 6 with regard to discourse *like*, concerns the delimitation of the envelope of variation and the related issue of the principle of accountability. The rigorous methodological procedures outlined in Labov (1982: 30) according to which a closed set of variants of a given variable is first defined, and all the non-occurrences of the variants in relevant contexts are taken into account raises a number of dilemmas as far as general extenders are concerned.

Firstly, as Dines (1980: 29) notes, there is no complete taxonomy of general extenders which can be used to define a closed class of variants. General extenders, like other formulaic expressions, are characterised by their semi-fixedness: in other words, they typically contain a number of slots where optional lexical material can potentially be inserted (see Wray 2002: 57). This means that many novel creations can be formed by exploiting the slots that are available for open class material (see Overstreet 1999: 4; and Wray 2002: 45). One possible solution to this problem that I have drawn upon in the analysis of discourse *like* is to confine quantitative work to uncovering the relative frequency of occurrence of a variable in a defined section of

speech (see Labov 1982: 87; Macaulay 2005: 6). In this chapter, raw numbers of general extenders in each age and gender cohort are converted to normalised frequencies of occurrence per 10,000 words to enable consistent comparisons to be drawn across age and gender groups within the corpus, as well as with the findings of other studies (see Macaulay 2002: 285; Cheshire *in press*).

From a theoretical perspective, another major difficulty in defining the envelope of variation hinges on the extent to which variant types of general extender may be viewed as equivalent to each other. This is of course a problem which bedevils the extension of quantitative methodology to non-phonological levels of variation (see Milroy and Gordon 2003: 169). In this chapter, I follow the practice of earlier studies in proposing that a general extender variable can be postulated on the basis of a common discourse function (Dines 1980: 15).

As far as the identification of relevant structures are concerned, each instance of a general extender was tagged in the initial compilation of the corpus to facilitate the extraction of tokens at a later stage. Identification of general extenders was based on the structural characteristics outlined in Dines (1980), Dubois (1992), Channell (1994), Overstreet (1999), and Aijmer (2002). By and large, general extenders in the preadolescent corpus conform structurally to the characterisation outlined in Dines (1980: 18): AND/ OR [PRO FORM] (LIKE THAT). A small number of deviations from this characteristic arrangement of constituents were also tagged as general extenders, as illustrated in example (22) below, where there is no conjunction present before the pro form, and in example (23) which atypically consists of a single item: *things*. This form was interpreted as a reduced variant of the tag *and/or things (like that)*:

- (22) it's on at 9 o'clock I think *something like that* [10M10/11]

(23) and then they'd say, 'Don't say ain't or yeah or no' ...*things*...

'Say it properly' [7F10/11]

Cases where a general extender was appended to material that was inaudible or unclear were eliminated from the quantitative analysis; for example:

(24) I was with my dad looking for ... for starfishes <UNCLEAR>

*things like that* [18M7/11]

In total, 148 tokens were retained for analysis and coded for a number of structural features and social factors discussed in previous studies of general extenders (e.g. Dines 1980; Aijmer 1985; Dubois 1992; Overstreet 1999). I briefly review the coding decisions that were made below.

Each form, or type, was given a unique code to facilitate type/token analysis. Thus, the form *and stuff* was given a different code from the lexically related variant *and stuff like that*, which, in turn, was coded separately from the form *and all that kind of stuff*.

Secondly, following Overstreet (1999: 4), each token was coded as to whether it was an adjunctive (i.e. beginning with *and*) or a disjunctive general extender (i.e. beginning with *or*), as in (25) and (26) below:

(25) they've ate all their insides *and everything* [22F7/8]

(26) these girls called Jane and Sandy they were always teasing me if I didn't have the right hairstyle *or something* or the right clothes [16F10/11]

Previous research has highlighted that the bipartite distinction between adjunctive and disjunctive types is significant.<sup>6</sup> In Overstreet's (1999: 8) database, disjunctive forms (N=89 (57%)) outnumber adjunctive ones.

With regard to the length of forms, Aijmer (1985: 373) notes that short forms (e.g. *and things*) are generally preferred to expanded forms (e.g. *and things like that*). Each token was therefore additionally coded for the number of individual word forms constituting the general extender.

Following Channell (1994: 132), who reports that general extenders most commonly follow noun phrases, verb phrases and embedded sentences, the type of grammatical constituent immediately preceding the appended general extender was also incorporated into the coding protocol.

In order to ascertain whether general extenders have a list-completing function in the corpus, I also coded each token as to whether it was a part of a list-like sequence (Jefferson 1990). In order to do this, I coded each token according to whether it was used to complete either a two-part (27), three-part (28), or longer list-like sequence (29) to investigate the existence of any normative arrangement of components in an enumeration containing a general extender.

(27) my dad phoned a ambulance *and all that* [9F10/11]

(28) it was like a shop that sells lots of sellotape and marker  
pens *and all that* [17M7/8]

(29) it was really fat and all purple and like... green *and everything* [19F10/11]

---

<sup>6</sup> Although Overstreet (2005) reports a similar split between adjunctive and disjunctive forms in German, it is not clear to what extent the distinction between adjunctive and disjunctive forms is operative in other languages where there are discourse equivalents to general extenders. Roth-Gordon (2001: 121) found no disjunctive examples in Brazilian Portuguese, but reports instead that vernacular speakers use adjunctive forms (e.g. *e o caralho* 'and shit'), or forms which lacked a connector (e.g. *não sei que* 'I don't know what'). Similarly, Dubois (1992: 192) reports that 42% (N=1141) of the general extenders she analysed in two corpora of Canadian French appeared without a preceding connector.

Furthermore, in order to investigate whether general extenders co-occur with other discourse-pragmatic features (see Stubbe and Holmes 1995), the presence of any additional clause-internal discourse markers was coded for in order to uncover any salient collocational patterns in the data.

Given the sensitivity of discourse-pragmatic features to contextual factors such as discourse type (see Stubbe and Holmes 1995: 66; Vincent and Sankoff 1992), I also examined the distribution of general extenders in different text types within the preadolescent data. General extenders were coded on the basis of whether they occurred in opinions, explanations, descriptions or narratives. Taking into account previous research on discourse markers which has shown that they may be differentially distributed throughout narrative macrostructure (see e.g. Schiffrin 1987: 274 on the function of *you know* in narrative backgrounded clauses), a more detailed coding protocol was used to investigate the distribution of general extenders in different narrative sub-components. In order to conduct a systematic analysis of general extenders in narrative, I draw on the framework for analysis outlined in Labov and Waletzky (1967), and Labov (1972), in order to explore any specific pragmatic roles they might have in narrative discourse.

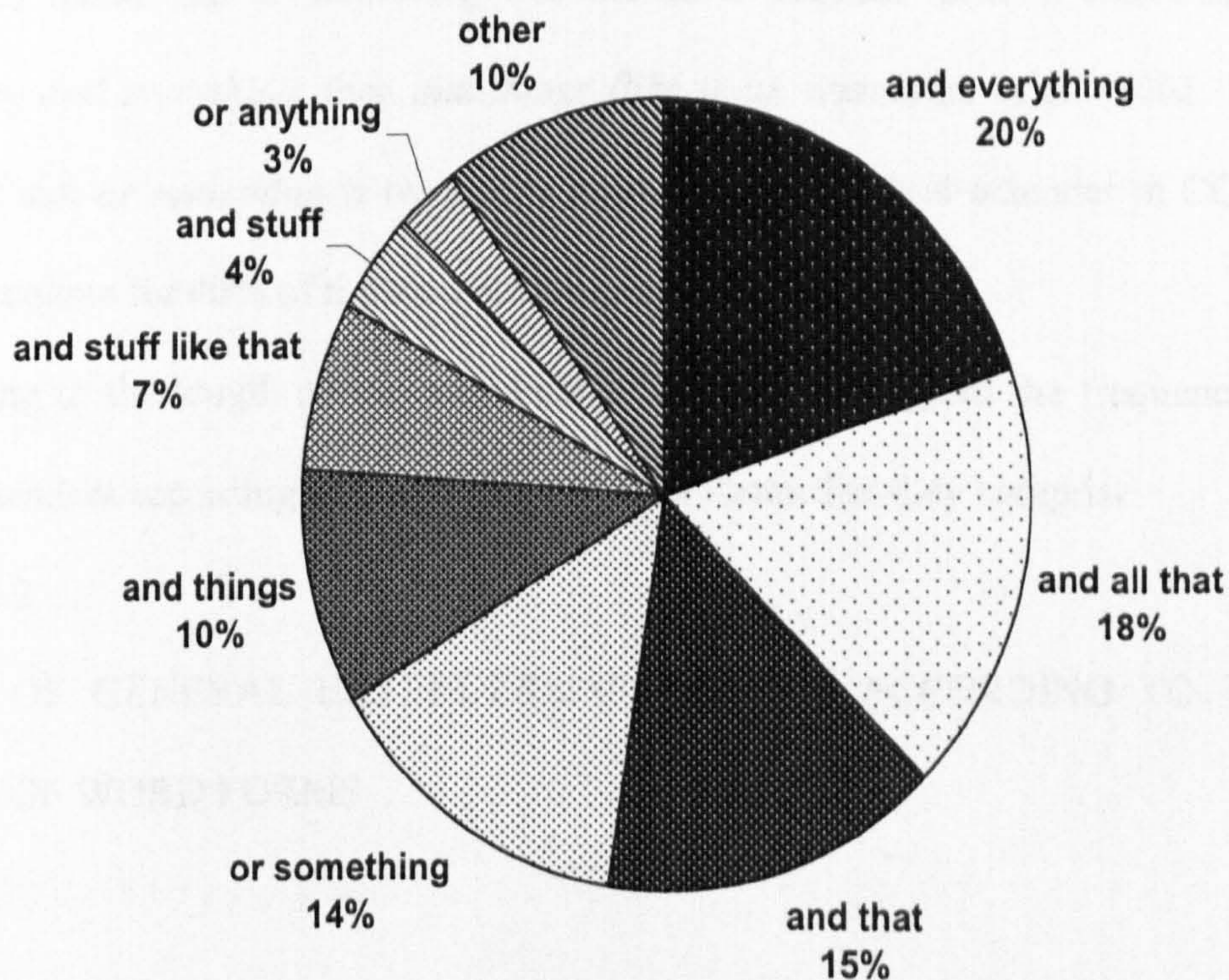
Finally, each token was coded according to the gender and age of individual speakers.

## 8.6 DISTRIBUTIONAL ANALYSIS

Figure 8.1 below gives an overview of the frequency of general extender types in the corpus.

FIGURE 8.1

FREQUENCY OF GENERAL EXTENDER VARIANTS IN THE  
PREADOLESCENT CORPUS



The form *and everything* is proportionally the most dominant, followed closely by *and all that* and the lexically related form *and that*. The OTHER category in Figure 8.1 consists almost entirely of nonce forms (see appendix). Examples (30) to (32) illustrate some of the forms that appear only once in the corpus.

(30) well they talk like, 'Wha::t's that?' *and like that* [9F10/11]

(31) he says, 'Say it properly or get your ... the right sounds' *or whatever* [10F10/11]

(32) the girl, she had to sit on a broken chair and like mum and auntie *and everyone* was getting really annoyed [27F7/8]

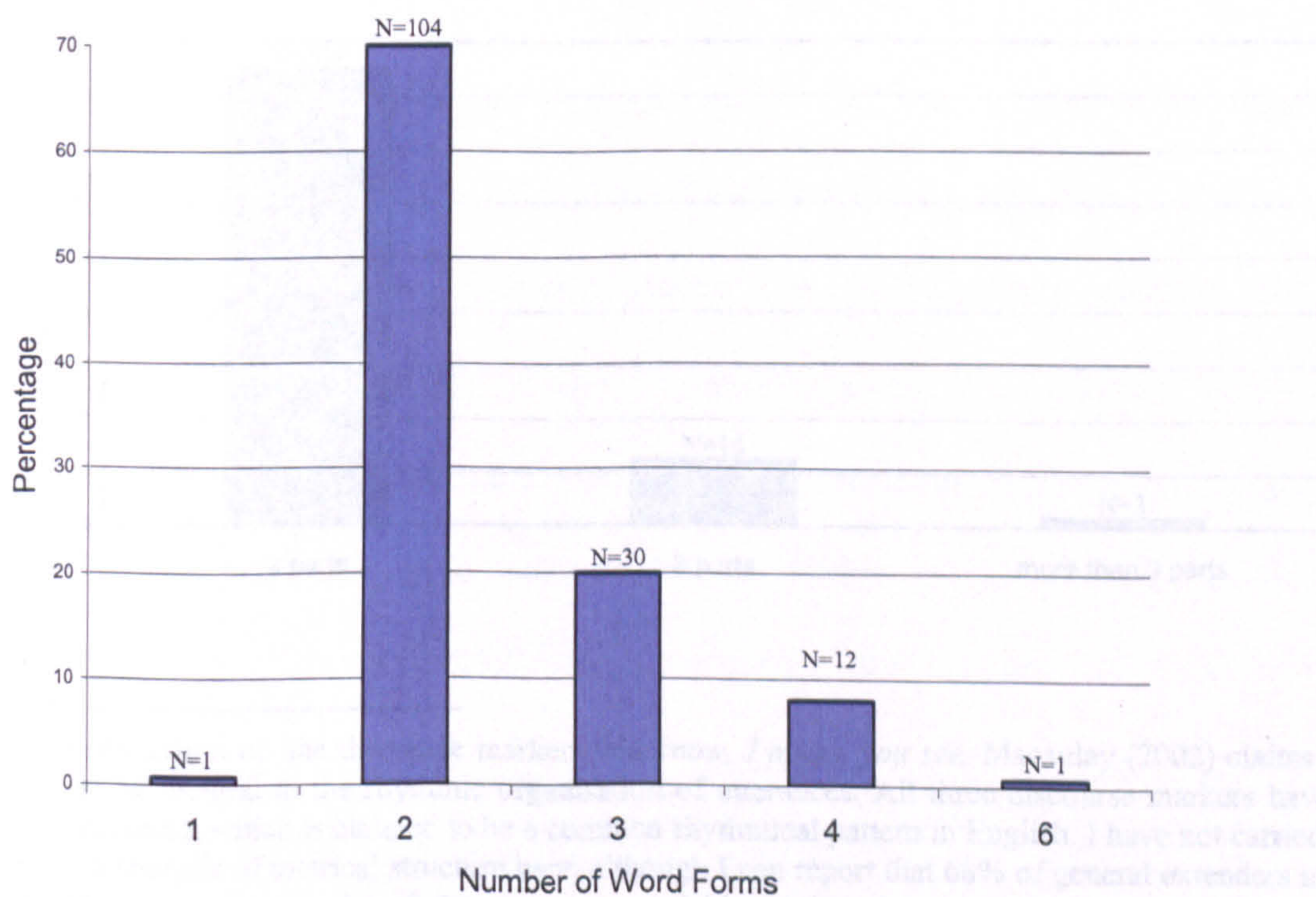


*Or something* accounts for 14% of the total number of general extenders, and is the only relatively frequent disjunctive form used by the preadolescents. Other studies have found that *or something* accounts for a much greater proportion of general extender usage than is the case here. Drawing on a far larger corpus, Biber *et al.* (1999: 116) found that *or something* was the most frequent form in conversation followed by *and everything*, then *and things (like that)*. Stenström *et al.* (2002: 102) also report that *or something* is by far the most frequent general extender in COLT, where it accounts for 40% of the total occurrence of forms.

Turning to the length of variants, Figure 8.2 below compares the frequency of general extenders according to the number of word forms that they comprise.

FIGURE 8.2

LENGTH OF GENERAL EXTENDERS COMPARED ACCORDING TO THE NUMBER OF WORD FORMS



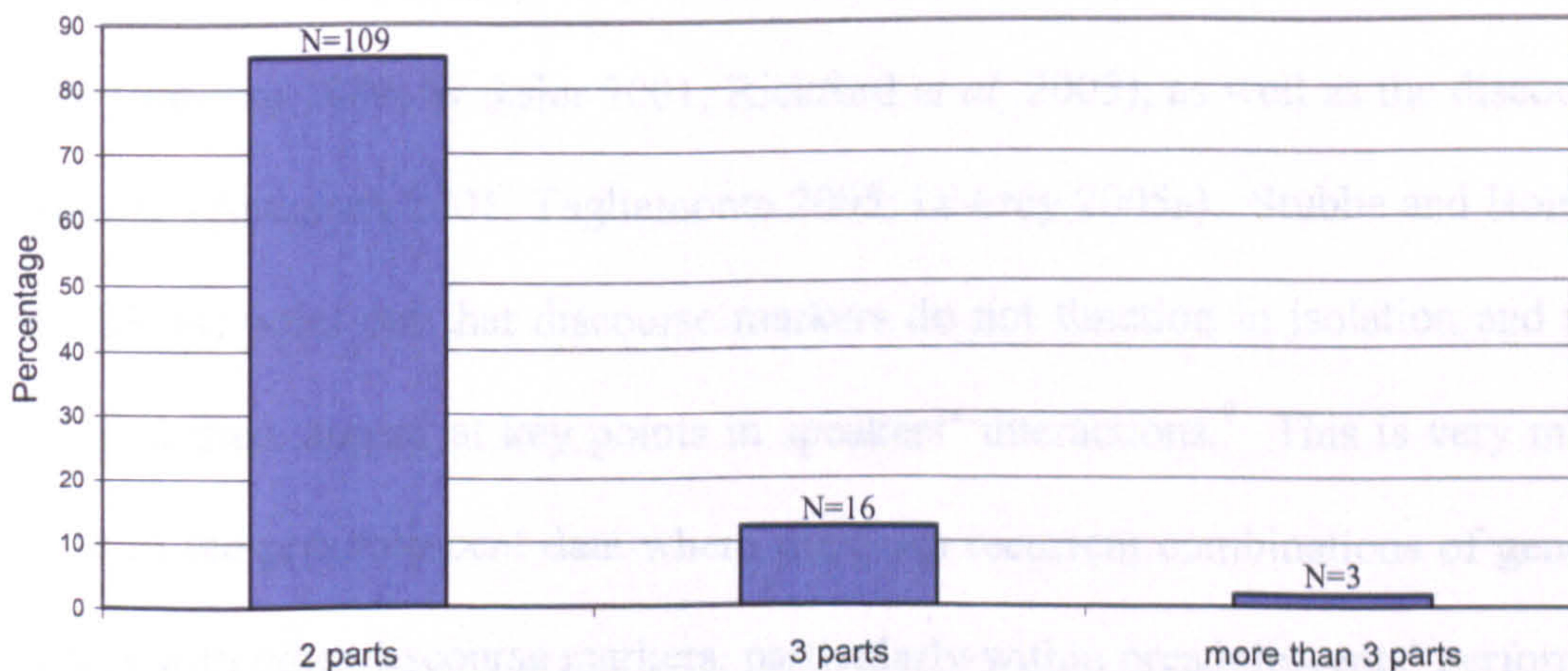
The distribution displayed above is consistent with the findings discussed in Aijmer's (1985:373) study of general extenders, in which short forms were generally found to be preferred to longer ones.<sup>7</sup>

In terms of the grammatical constituents which immediately precede general extenders in the corpus, 50% are nouns (N=74); 22% are verbs (N=32); and 13% are adjectives (N=19). A further 13.5% (N=20) are appended to direct quotations (discussed below); and 2% (N=3) are appended to numerals. Norrby (2002) also reports that general extenders are most frequently appended to nouns in her Swedish data. However, the high frequency of general extenders in post-nominal position may simply result from the greater number of noun phrases in the data relative to other grammatical categories.

Figure 8.3 below examines the frequency of list-like sequences in the data.

FIGURE 8.3

FREQUENCY OF LIST-LIKE SEQUENCES IN THE CORPUS<sup>8</sup>



<sup>7</sup> In his discussion of the discourse markers *you know*, *I mean*, *you see*, Macaulay (2002) claims that they may be integral to the rhythmic organisation of utterances. All three discourse markers have an iambic structure, which is claimed to be a common rhythmical pattern in English. I have not carried out a detailed analysis of metrical structure here, although I can report that 68% of general extenders in the preadolescent corpus consist of three or more syllables rather than two associated with an iambic structure.

<sup>8</sup> General extenders which followed reported speech (N=20) were not counted in the distributional analysis of list-like sequences.

In line with Overstreet's (1999:26) findings based on adult usage of general extenders, there is no evidence in the preadolescent data that general extenders are typically incorporated into three-part lists (i.e. 2 items + general extender). 85% of general extenders in the preadolescent data follow one item only, with only 12.5 % participating in a three-part arrangement. In infrequent cases where general extenders are appended to more than 2 items, the cumulative effect of a number of components in an enumeration appears to be evaluative (see Schiffrin 1994: 308; and Schiffrin 2006: 171-2). In example (29) above, repeated for convenience as example (33) below, the general extender *and everything* is appended to a concatenation of adjectives which iconically help to convey the seriousness of an eye injury that a child described:

(33) it was *really* fat and *all* purple and *like...* green *and everything*

[19F10/11]

The example above also graphically illustrates that general extenders collocate with other discourse particles such as the intensifier *really*; the adverbial intensifier *all* (see Cheshire 1989; Waksler 2001; Rickford *et al.* 2005); as well as the discourse marker *like* (Andersen 2001; Tagliamonte 2005; D'Arcy 2005a). Stubbe and Holmes (1995: 83-84) point out that discourse markers do not function in isolation and that complex clusters appear at key points in speakers' interactions.<sup>9</sup> This is very much the case in the preadolescent data where there are recurrent combinations of general extenders with other discourse markers, particularly within preadolescents' performed narratives. 26% of general extenders in the corpus occur in conjunction with other

---

<sup>9</sup> This observation highlights the importance of taking into account the surrounding context when studying discourse variables. The need to adopt a microscopic approach when dealing with discourse features points to methodological differences between the analysis of discursive variables and the study of variables from other levels of the grammar where careful attention to the discourse context may be less integral to the analysis of a specific feature. In the case of co-occurring discourse particles, which often embody a range of interacting functions, it is also important to note, as Stubbe and Holmes (1995: 84) point out, that 'the whole is somewhat greater than the sum of its parts.'

discourse markers. The most frequent of these are *all* and *like* which occur in various combinations with general extenders, as shown in examples (34) to (37) below:<sup>10</sup>

(34) there was *like* loads... *like* dinosaur bones *and that* [15M7/8]

(35) she was calling me *all* names *and everything* [16F10/11]

(36) they kept eating *like all these* goats *and things* [27F7/8]

(37) they found *all like* dinosaurs *and all that* [13M7/8]

In the examples above, all taken from narrative recounts, the general extender is appended to discourse-new information. As one of the characteristic pragmatic functions of general extenders is to indicate knowledge shared by the speaker and hearer (Overstreet 1999), their use in discourse presupposes that the hearer is familiar with what the speaker is talking about. Brinton (1990: 54) has argued that the presentation of new information as if it were old or familiar information can be exploited by speakers to improve hearer reception of discourse-new entities.

The cluster of discourse signals which helps to draw the hearer's attention to discourse-new information not only includes *like* and *all* used in conjunction with general extenders, but also the plural demonstrative *these* as in (36), which is one of a number of devices that can reduce the distance between the taleworld and the storytelling situation, and by extension, between the teller, tale and audience (see Georgakopoulou and Goutsos 1997: 140; Toolan 1988: 158). In (36) above, we can see how discourse-pragmatic markers in this particular context operate synergistically, or are mutually reinforcing (see also Macaulay 2005: 9).<sup>11</sup>

---

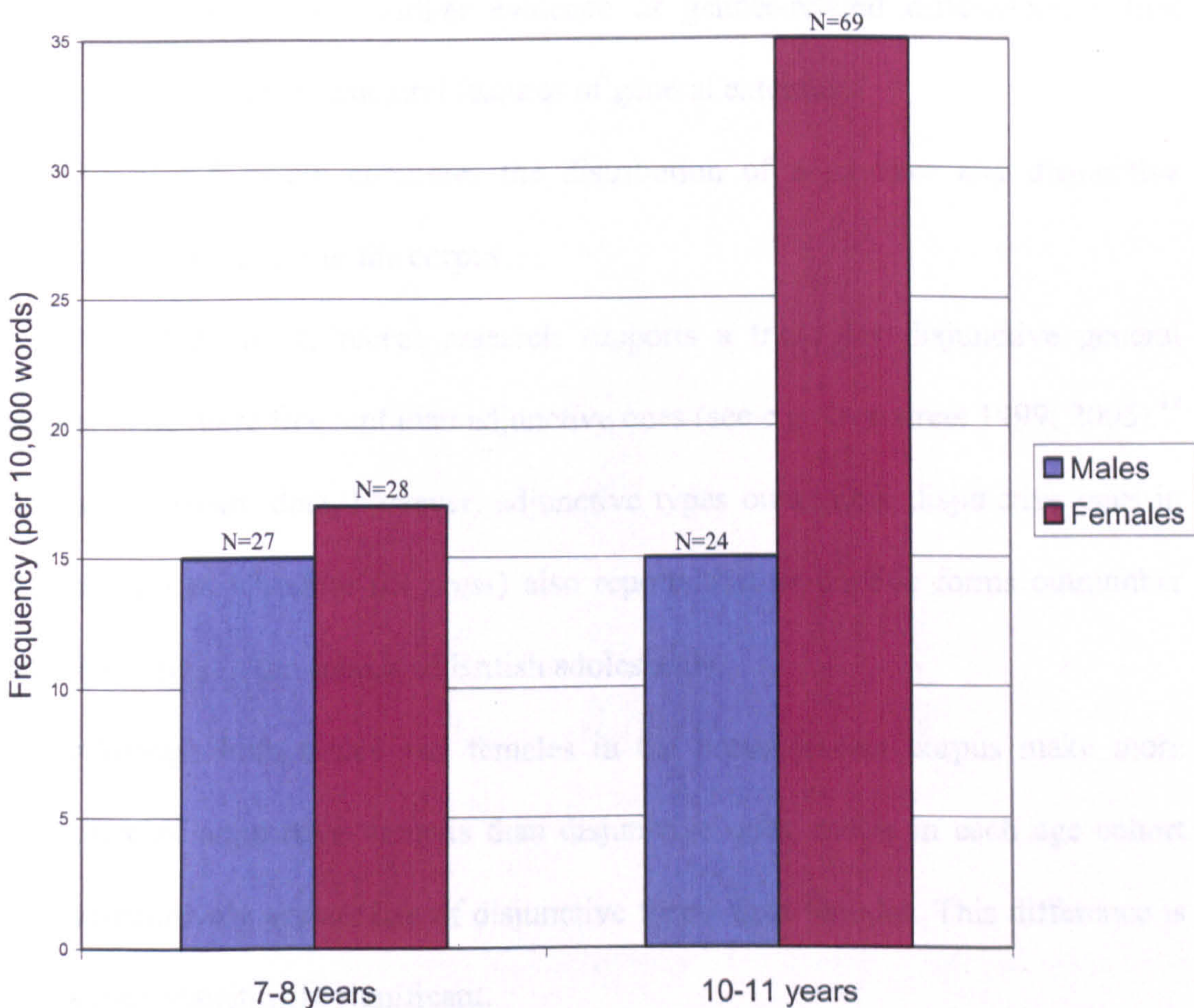
<sup>10</sup> The absence of *you know* in collocations with general extenders in the preadolescent data is conspicuous: Overstreet (1999: 74-6) notes that *you know* often co-occurs with general extenders in the American English data that she analysed. The pragmatic functions associated with *you know* may be (at least partially) fulfilled in the preadolescent corpus by other discourse markers.

<sup>11</sup> Fleischman (1999) cites Hansen's (1998) observation that pragmatic markers often functionally overlap. Fleischman (1999) argues that even when discourse markers do co-occur, each marker is carrying out a different pragmatic function. However, it is not clear how the distinct functions of different discourse markers occurring within a cluster can be systematically teased apart.

Turning next to social factors, Figure 8.4 below compares the frequency of general extenders in the preadolescent corpus by age and gender.

FIGURE 8.4

FREQUENCY OF GENERAL EXTENDERS ACCORDING TO THE AGE AND GENDER OF SPEAKERS



(chi square value = 8.297, df = 1, p<0.01)

In both age cohorts, females make greater use of general extenders than males. However, although the difference between males and females is relatively small in the 7-8 age cohort, it is much more salient in the case of the older children where females show a noticeable increase in their use of general extenders. In brief, girls show a marked upturn in general extender usage as they get older, whereas the frequency of

usage for boys remains relatively constant with age. This finding suggests that fluctuations in the frequency of general extenders in preadolescent vernacular usage cannot be accounted for in terms of a simple female/ male dichotomy; instead, the data suggest, as Holmes (1997: 203) contends, that gender is a complex social construct that interacts with other social dimensions such as age.

In order to explore whether female and male preadolescent speakers' use of general extenders exhibits further evidence of gender-related differences, I first consider variation in the structural features of general extenders.

Figure 8.5 below compares the distribution of adjunctive and disjunctive forms by age and gender in the corpus.

As noted above, recent research supports a trend for disjunctive general extenders to be more frequent than adjunctive ones (see e.g. Overstreet 1999; 2005).<sup>12</sup> In the preadolescent data, however, adjunctive types outnumber disjunctive ones in both age cohorts. Cheshire (*in press*) also reports that adjunctive forms outnumber disjunctive forms in the speech of British adolescents.

Although both males and females in the preadolescent corpus make more frequent use of adjunctive variants than disjunctive ones, males in each age cohort make comparatively greater use of disjunctive forms than females. This difference is not, however, statistically significant.

---

<sup>12</sup> Overstreet (1999; 2005) found that disjunctive general extenders outnumbered adjunctive forms in her English and German corpora respectively. Norrby (2002), on the other hand, found that adjunctive forms were proportionally dominant in her Swedish data.

FIGURE 8.5

COMPARISON OF THE FREQUENCY OF ADJUNCTIVE AND DISJUNCTIVE GENERAL EXTENDERS BY AGE AND GENDER<sup>13</sup>

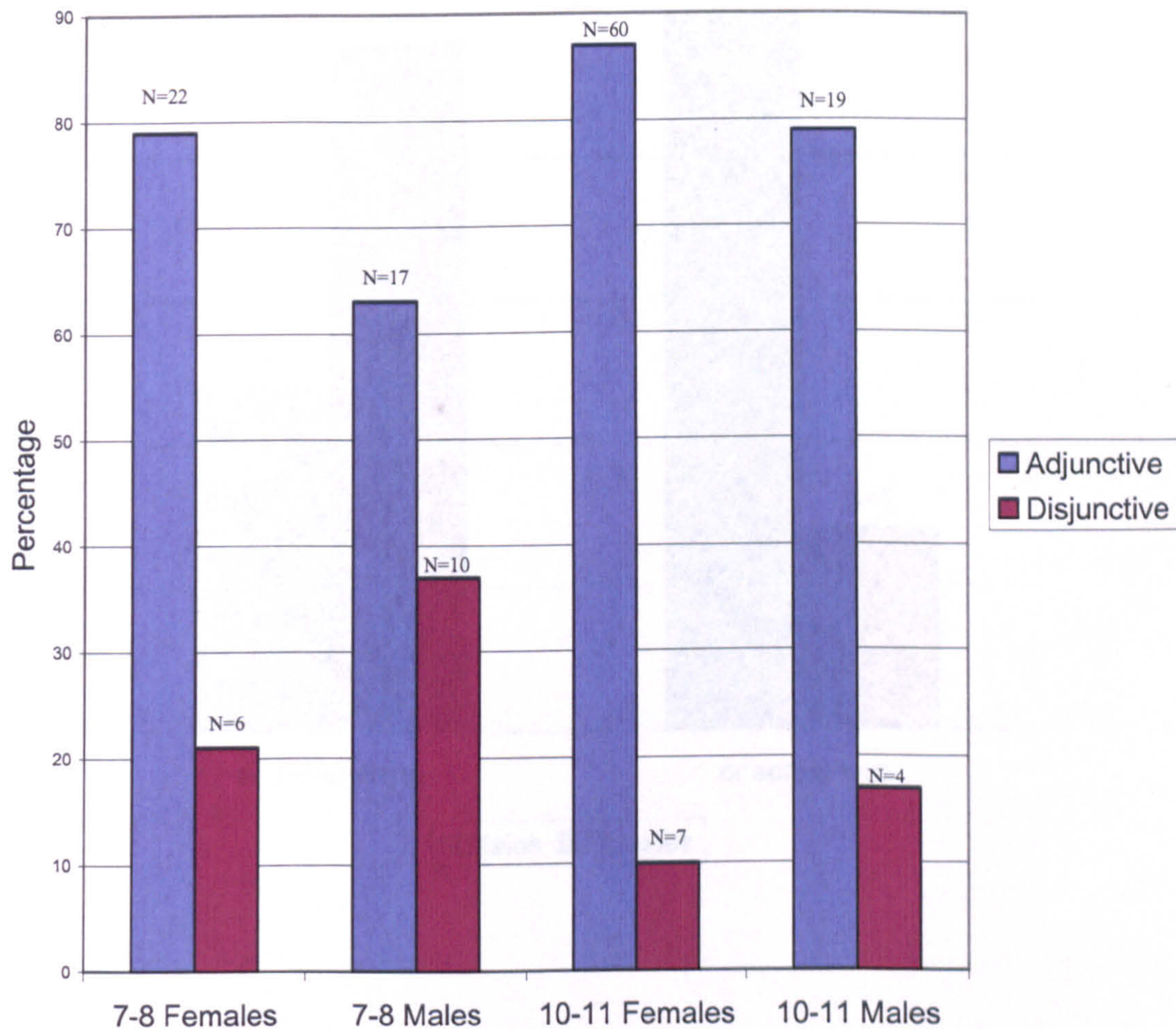
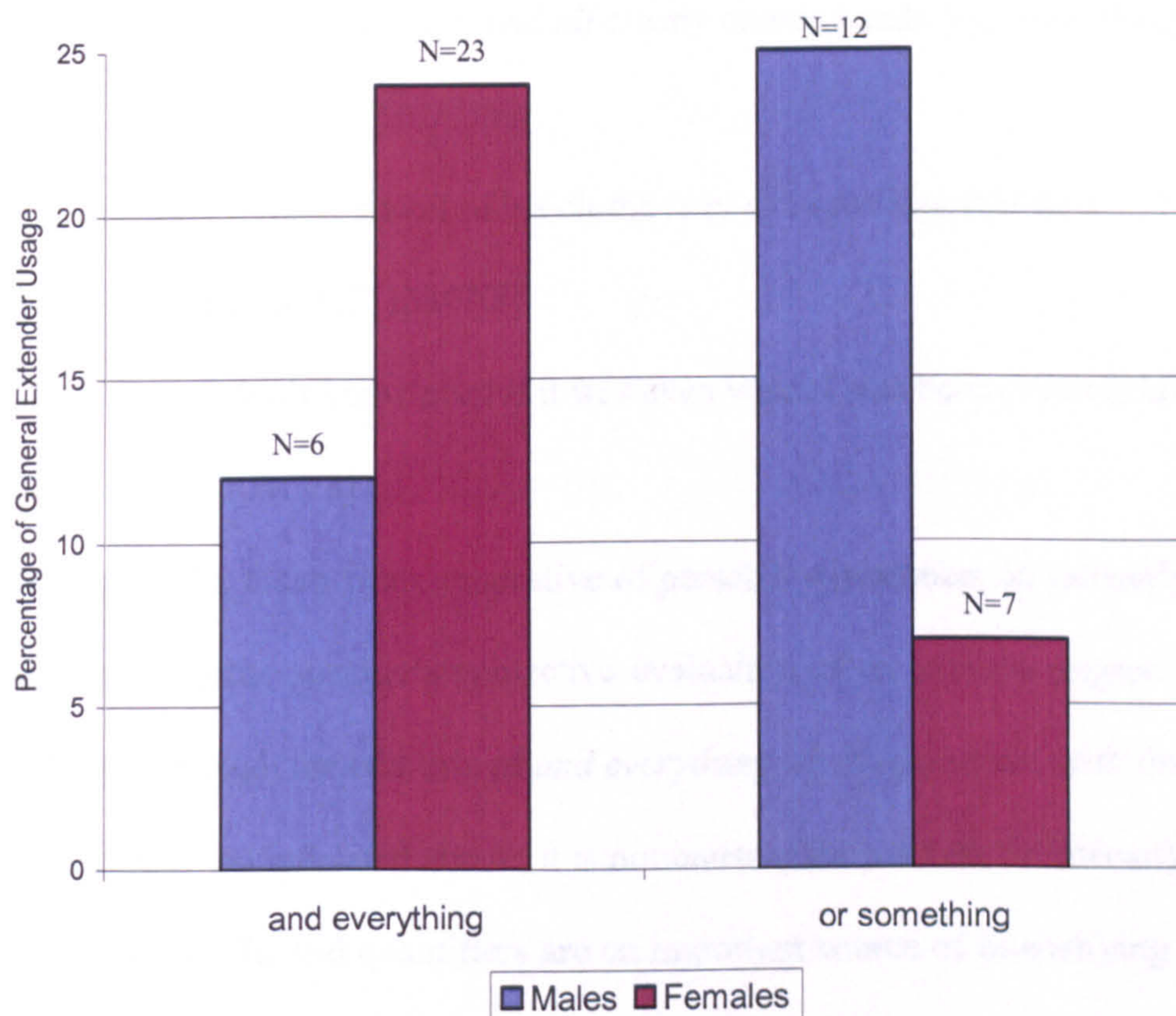


Figure 8.6 below compares the frequency of the most common adjunctive form, *and everything*, with the most common disjunctive form, *or something*, in order to explore whether these are differentially distributed in the speech of females and males in the corpus.

<sup>13</sup> 3 tokens were not preceded by a conjunction.

FIGURE 8.6

COMPARISON OF THE FREQUENCY OF *AND EVERYTHING* AND *OR SOMETHING* BY GENDER



Although multifunctionality is an integral feature of general extenders, Overstreet (1999: 80; 104-5; 146-7) claims that frequently occurring variants such as *and everything* as well as *or something* can have pragmatically different functional orientations in discourse. Overstreet and Yule (2002: 788) claim that *and everything* has an intensifying function in discourse and evokes a scale, with information preceding *and everything* being at a high or extreme point on this scale (see also Ward and Birner 1993; Aijmer 2002: 242). *Or something*, on the other hand, is claimed to be used to reduce threats to negative face (see Brown and Levinson 1987: 65-66) and can function as a strategy of negative politeness or tentativeness



(Overstreet 1999: 105). These pragmatic orientations can be discerned in preadolescents' use of *and everything* and *or something*, as illustrated below:

(38) I knew that they were going to do it, but I never knew what time and I never knew they had *all* creepy crawly hands *and everything*

[15F10/11]

(39) I think it's a bit of his clothe (*sic*) *or something* that he's ripped off [13M7/8]

(40) I don't know even if it was even when I was born *or something* [20M7/8]

In example (38), taken from a narrative of personal experience, an intensifying effect is conveyed by the speaker's subjective evaluation of an extreme degree of 'creepy crawliness' through the addition of *and everything* which co-occurs with the adverbial intensifier *all*. As indicated above, it is not unusual for markers of intensity to cluster (Labov 1984b: 57), and quantifiers are an important source of intensifying devices in narrative.

In example, (39) *or something* is used to tentatively express one of a possible number of options, with the speaker's use of *I think* additionally indicating that *or something* is functioning as a hedge or marker of uncertainty in this context. Similarly, in example (40), the tentativeness implicit in the boy's utterance is not only marked by *or something*, but is also indicated by *I don't know* in the matrix clause, which marks epistemic uncertainty (see also Overstreet 1999: 115 on collocations of *or something* with *I think* and *I don't know*).

If we concede that *and everything* and *or something* may have different pragmatic orientations in discourse, then there is scope for arguing that their differential frequencies in the speech of preadolescent boys and girls may reflect

different discourse styles, or different orientations towards the construction of discourse (see also Chapter 6 and Chapter 7). However, this suggestion is necessarily provisional, not only because of the low number of tokens involved, but also because it is not clear to what extent other discourse-pragmatic features in the surrounding context contribute to particular pragmatic inferences.<sup>14</sup>

More compelling evidence of gender-related differences in the use of general extenders emerges from a type/token analysis of variant forms. Table 8.1 below breaks down the number of types and tokens of general extender by age and gender. The most salient finding to emerge from the tabulation of the data is that the older girls make use of a much more diverse range of variants than any other group. This finding has interesting parallels with Vincent and Sankoff's (1992) study of 'punctors' or pragmatic particles in Canadian French (e.g. *là* 'there', *vois-tu* 'do you see', and *n'est-ce pas* 'isn't it so'). Vincent and Sankoff (1992: 207) note that as the number of 'punctor' occurrences increases, so does the number of different forms used. This appears to be the case too for the preadolescent girls in the oldest cohort, who are not only the most prolific users of general extenders, but are also the speakers with the most extensive repertoire of variant forms.

---

<sup>14</sup> Although my interpretation of the qualitative differences in the boys' and girls' use of *and everything* and *or something* may seem overly cautious in the light of the gender pattern displayed in Figure 8.6, it is important to note that results obtained by other researchers (e.g. Cheshire *in press*) do not corroborate Overstreet's (1999: 97) claim that adjunctive general extenders can function as strategies of positive politeness, whereas disjunctive general extenders can function as strategies of negative politeness. Cheshire (*in press*) found that both adjunctive and disjunctive general extenders could have negative politeness functions. Furthermore, in my own data, it was sometimes difficult to establish interpretations of *or something* as a marker of negative politeness objectively. Consider the following example from a seven-year-old boy who explained to me how to play a game I was unfamiliar with: *and once you've beat like five yeah... then you get your place in history in the game or something yeah* [20M3]. In this example, there are no other co-occurring discourse features which support an interpretation of tentativeness on the part of the speaker. Furthermore, the speaker was clearly more familiar with the game than I was, so there was little risk of his eliciting a face-threatening contradiction about the nature of his assertions from myself. I draw on this example to highlight the fact that the assignment of a primary pragmatic function to general extender types is not necessarily straightforward.

TABLE 8.1

## TYPE/TOKEN ANALYSIS OF GENERAL EXTENDERS ACCORDING TO AGE AND GENDER

Speaker group	Number of types	Number of tokens
7-8 girls	7	28
7-8 boys	5	27
10-11 girls	19	69
10-11 boys	7	24

When we consider developmental aspects of general extender usage in the preadolescent corpus, there is evidence that some of these forms perform text-organising functions that are more characteristic of older speakers than younger ones; in other words, the functions of general extenders seem to diversify as children mature. Consider examples (41) to (44) where a general extender follows direct speech:<sup>15</sup>

(41) they go, 'No you !' *and stuff like that* [8F10/11]

(42) boys say, 'Come here man' *and everything like that* [15F610/11]

(43) they talk like, 'Wha::t's tha::t?' *and like that* [9F10/11]

(44) he missed and then... 6R was saying, 'Oh youse are going to lose' *and all that* and then Mark went over to him and they had a fight [9M10/11]

13.5% (N=20) of general extenders in the corpus follow quoted speech. Overstreet (1999: 118) notes that the use of general extenders following reported dialogue is

<sup>15</sup> Prosodic changes (e.g. pitch configuration, loudness and timing) were used as cues to determine whether general extenders were employed as boundary markers for reported speech. In a few cases, the prosodic format of reported utterances suggested that a general extender occurred *within* the boundaries of reported speech. Such cases are excluded from the discussion above which only focuses on general extenders that are set off prosodically from reported utterances.

apparent in her data too. In the preadolescent corpus, it is only the older children who make use of general extenders in this environment, with only one token found following quoted speech in the 7-8 age cohort. It is certainly not the case that the younger children do not use reported speech; on the contrary, their narratives are typically polyphonic (Bakhtin 1981a) and make extensive use of direct speech, as I have shown in Chapter 7. As I have already noted, the metalinguistic skills required for integrating reported speech into narrative appear to be developmentally constrained (see Hickman 1993:64). What we may be seeing in the older preadolescent age group, therefore, is a tendency to make the boundaries of reported speech more explicit by using a general extender as a text-organising marker which signals the end of quoted talk (see Bolden 2004).

It is also the case that the general extenders in examples (41) to (43) above perform an illustrative function by marking the preceding quoted material as an approximation in form and content to what could have been said in a specific instance. In (42) and (43) above, the appended general extender is used by the young female speakers to indicate that the preceding quote is an illustrative remark attributed to the generic voice of the boys in the speaker's class. In (43), the speaker additionally makes use of expressive phonology resulting in a 'polyphonic layering' of voices (see Günther 1999). By framing a reported utterance in this way, the female speaker concerned also displays aspects of her own gendered identity by disaffiliating herself from the perceived speech style of boys. Thus, in addition to marking boundaries between chunks of discourse, (e.g. between quoted speech and the resumption of narrative discourse, as in (44) above), general extenders, in conjunction with other discourse-pragmatic features, can function at the interpersonal level to help signal speakers' social alignments or positionings relative to the content of their discourse

(see also Norrby and Winter 2002). The ability to evaluate discourse in this way appears to emerge as children mature (see e.g. Kernan 1977: 101, who claims that in narrative, there is an expansion in the use of evaluative devices with increasing age).

### 8.6.1 *Distribution in narrative*

I now turn to a consideration of the distribution of general extenders in different text types, focusing specifically on narrative. Table 8.2 below gives a breakdown of the frequency of general extenders in four different text types in the preadolescent corpus.

TABLE 8.2

#### DISTRIBUTION OF GENERAL EXTENDERS ACROSS FOUR DIFFERENT TEXT TYPES

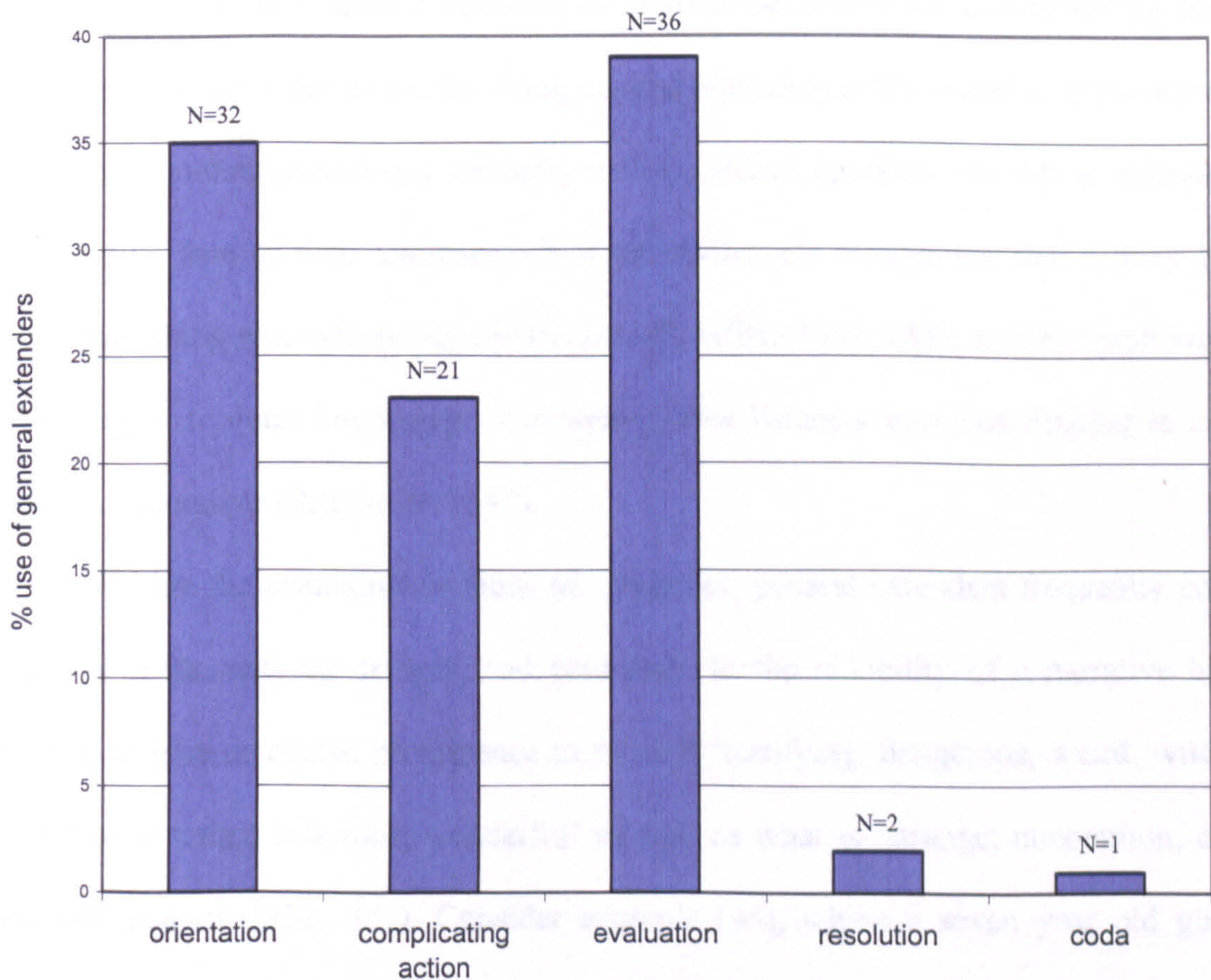
Text Type	Number of tokens	% of total general extenders
Narrative	92	62
Description	26	18
Opinion	25	17
Explanation	5	3

Table 8.2 shows that general extenders occur in a number of different text types in the data, with the largest proportion found in narrative discourse. The concentration of general extenders in this particular genre can most probably be attributed to the fact that narratives are the predominant text type in the corpus, reflecting one of the original aims of the study which centred on eliciting narratives of personal experience from preadolescent speakers.

If we consider the distribution of general extenders within the various sub-components of narrative macrostructure, we can see that their distribution is unevenly correlated with different structural features, which provides some insight into their possible functional role in narrative.

FIGURE 8.7

DISTRIBUTION OF GENERAL EXTENDERS WITHIN NARRATIVE MACROSTRUCTURE



It can be seen from the results displayed in Figure 8.7 above that general extenders tend to occur most frequently in the evaluation and orientation sub-sections of

narratives.<sup>16</sup> Previous studies that have focused on discourse markers have uncovered similar correlations with specific narrative sub-components. According to Schiffrin (2001: 66), *you know* often introduces a story preface, and Brinton (1990: 54) argues that *you know* has both a grounding and evaluative function in narrative. Given that orientations generally furnish descriptive information about the spatiotemporal setting of a story (Georgakopoulou 1997: 58), as well as additional details about characters and their activities (Labov 1972: 364), it is not surprising that *you know* occurs within orientation sections where there is both a need to introduce background material as well as to draw the listener's attention to material necessary for understanding the following discourse. Similarly, by using general extenders at the outset of a narrative to secure common ground and solidarity with a listener, speakers are able to address the positive face of their audience; elicit the audience's recognition that a story is about something with which they are familiar (Schiffrin 1987: 284); and, by implicitly appealing to conjoint knowledge, vicariously draw listeners into participating in an unfolding narrative (Britain 1998: 47).

Within the evaluative sections of narratives, general extenders frequently co-occur with intensifying devices, and contribute to the tellability of a narrative by helping to give discourse prominence to what is 'terrifying, dangerous, weird, wild, crazy; or amusing, hilarious, wonderful' as well as what is 'strange, uncommon, or unusual' (Labov 1972: 371). Consider example (45), where a seven year old girl recounts a fight with her brother who had taken a book from her:

---

<sup>16</sup> Very few children used abstracts in recounting narratives of personal experience, hence their absence in Figure 8.7. As Johnstone (2001: 639) points out, it has been observed repeatedly that not all stories have abstracts or codas. It may also be the case, as Romaine (1985: 92) suggests, that since these narratives were elicited in an interview-like situation, children counted the questions, which elicited narratives on particular topics, as abstracts.

(45) he wouldn't give it back  
so I went and told mum  
and she went, 'Oh just stay away from him'  
and I was like, 'OK'  
so I stayed away from him  
and he kept coming up to me  
and <unclear> he kept swearing at me *and everything*  
and kicking me *and everything*  
so then I picked him up  
cos I'm quite...  
I can pick him up  
cos I'm quite strong  
and I chucked him ... on to ... against the wall  
and he broke his leg

This recount contains many of the features of what Wolfson (1982: 25) terms a performed narrative, including the use of direct speech, asides, as well as repetition (see Chapter 5). In this narrative, the general extender *and everything* appears twice, and is used in conjunction with the semi-auxiliary *keep*, which Labov (1972: 376, f.n.10) identifies as having a possible intensifying function in narrative. The use of parallel constructions (e.g. *he kept coming up to me... he kept swearing at me and everything*) also functions evaluatively (Tannen 1982: 14), and draws the listener's attention to material which is interesting, surprising or unusual in the story. The speaker's use of an appended participle followed by *and everything* (i.e. *he kept swearing at me and everything and kicking me and everything*) is yet another



evaluative element (Labov 1972: 387) which places additional emphasis on the simultaneity of distinct actions (Toolan 1988: 156), and helps to heighten the dramatic tension in the narrative before allowing the resolution to be delivered with greater impact (see Labov 1972: 374). We can also see that the repeated use of the general extender *and everything* in the above example is embedded in what Georgakopoulou and Goutsos (1997: 122) refer to as a tripartite action pattern which can be schematised as follows:

- a. and he kept coming up to me
- b. and <unclear> he kept swearing at me *and everything*
- c. and kicking me *and everything*

This tripartite pattern, which Georgakopoulou and Goutsos (1997: 122) claim to be commonly encountered in narratives of many cultures, helps to focalise the actions of the antagonist before the transition to a new event-schema signalled by *so then I picked him up*, which moves the narrative action forward. Thus, in example (45) above, the general extender *and everything* participates in larger organisational structures which operate evaluatively and are strategically deployed to enhance plot advancement.

A possible pragmatic effect of *and everything* in the narrative above, where it is uttered with contrastive pitch accent (see Overstreet and Yule 2002), is to foreground that there is something remarkable in the information to which it is appended (see Aijmer 1985: 385). As in (38) above, it appears to 'scale up' the accompanying information as being at an extreme point (Overstreet and Yule 2002: 788). Not only does *and everything* mark the speaker's affective involvement in her narrative, it also invites listeners to draw on their common knowledge in interpreting its inferential scope. As far as story structure is concerned, *and everything*, in concert

with other constructions, helps to draw the listener's attention to material in the story which is important for understanding the speaker's subsequent actions reported towards the end of the narrative.

Additional evidence that *and everything* has an intensifying function in preadolescent narratives can be adduced from the fact that it constitutes 42% of all general extenders in narrative evaluations, whereas in the corpus as a whole, it accounts for only 20% of all forms. Since narrative evaluations are a prime locus for intensity markers (Labov 1972), it seems plausible that the proportionally greater use of *and everything* in evaluation sections is attributable to its intensity-marking role. Further indications that general extenders comprising universal quantifiers can evolve intensity-marking functions can be found by considering cross-linguistic evidence: Farghal and Haggan (2005: 410) argue that although the general extender *u-kulši* (*and everything*) in Arabic can have general extender functions, it may also exclusively operate as an intensifier.

## 8.7 SUMMARY

In the preceding sections, I have presented a quantitative and qualitative analysis of general extenders in preadolescent speech. I have shown that they are typically multifunctional and have referential, textual and interpersonal uses which are exploited by preadolescent speakers. Crucially, these forms have a functional orientation based on the co-operative exploitation of shared knowledge and experience which can help to sustain a sense of rapport and reinforce solidarity between interlocutors (Overstreet 1999: 18).

It was also shown that general extenders occur in other varieties of English, as well as in a number of typologically diverse languages. Although general extenders

appear to be highly formulaic constructions cross-linguistically, there is some evidence that they are used for comparable discourse-pragmatic functions, although further research is needed to substantiate superficial cross-linguistic parallels.

General extenders were also found to occur in a number of different text types. Evidence from preadolescent narratives revealed that in terms of frequency, they correlate with particular aspects of narrative structure. Higher frequencies of usage in orientation and evaluation sections of narratives suggest that, like other addressee-oriented features of discourse, general extenders occur in precisely those sub-components of narrative that appeal to the listener's empathy and foreground common knowledge (see Britain 1998: 50-51). Thus, their concentration in evaluation sections may be accounted for in terms of the need to elicit the interactive support of the listener in appreciating the intrinsic interest-value of a story. Likewise, their use in orientation sections draws the listener's attention to important background information necessary for understanding the unfolding of events in a story. Evidently, general extenders do not act alone in these functions, and they are part of a network of discourse devices which are recruited for similar purposes (see e.g. Britain 1998 on *High Rising Terminal* contours; and Schiffrin 1987, and Brinton 1996 on *you know*).

One of the most salient findings to emerge from the distributional analysis of the data is that general extender usage appears to be sensitive to the age and gender of the speakers in the preadolescent cohort examined here. It was found that not only did the females in the older age cohort show a marked upturn in their frequency of use of general extenders, but also that the inventory of variant forms they use is considerably broader than that of other speaker groups in the corpus.

## CHAPTER 9

### CONCLUDING REMARKS

#### 9.1 THE SOCIOLINGUISTIC COMPETENCE OF PREADOLESCENTS

This study has aimed to contribute to empirical characterizations of the structured nature of variation in preadolescence. Using a variationist methodology to explore differences in a tightly compressed age range (7-11), one of the overarching aims of the study has been to explore the existence of internally systematic patterns of grammatical variation in this age group, and to investigate the intersection of internal linguistic constraints with the social parameters of age and gender. From a broad perspective, the examination of six key variables drawn from different levels of the grammar confirms Roberts' (1994: 3) observation that in acquiring the dialect of their community, children acquire variable structures along with categorical features of language.

Nevertheless, although children acquire patterns of inherent variation at an early age, the results of the distributional and multivariate analyses in the preceding chapters corroborate Kerswill's (1996:199) point that not all variables are acquired at the same rate, and that differential rates of acquisition of variable structures are dependent on the linguistic level to which they belong. Furthermore, the lack of congruity in the operation of social constraints on variables from different grammatical sub-components indicates, as Smith *et al.* (2007: 65) have observed, that the impact of external and internal constraints on variation is not isomorphic across all variables.

### 9.1.1 Summary of analyses

To explore these issues further, it is instructive to consider Table 9.1, which presents a summary of the major findings emerging from the multivariate analyses in Chapters 3, 4, 5 and 7. This summary is supplemented by further details provided below relating to the findings emerging from the distributional analyses of discourse features in Chapters 6 and 8.

TABLE 9.1

#### SUMMARY OF FINDINGS EMERGING FROM MULTIVARIATE ANALYSES

Variable	Significant internal linguistic constraints	Significant age constraints	Significant gender constraints
<i>was/were</i>	✓	X	✓
relativization	✓	X	X
tense variation	✓	✓	X
quotatives	✓	✓	X

In line with previous studies, the tabulated summary of findings indicates that internal linguistic factors have primacy as far as contextual constraints on variation are concerned (see Preston 1991: 33). Moreover, several of these variables reveal quite complex internal conditioning, as evidenced by the configuration of factors constraining the occurrence of variant forms. Indeed, one of the advantages of the analytical method employed in the study has been the use of multivariate analyses to probe deeper constraints operating on dialect-specific structures (see e.g. *was/were* variation and relativization) which might not otherwise have been evident from the

superficial inspection of frequency distributions of variant forms.<sup>1</sup> The fact that these variable structures are systematically constrained in the speech of preadolescents bolsters recent research (Smith *et al.* 2007) demonstrating that complex internal conditioning of variables is an intrinsic feature of language acquisition in childhood.

Turning to external variables, age attains statistical levels of significance in two of the multivariate analyses, whereas gender is only significant in *was/were* variation. However, the magnitude of significant effects is generally greatest for internal linguistic constraints, with age exerting a much weaker effect on linguistic variation. The only exception to the greater magnitude of statistically significant internal constraints vis-à-vis external ones is found in *was/were* variation, where the range for gender is slightly larger (a point spread of 30) than the range for grammatical person (a point spread of 26; see Table 3.9). In this particular case, the direction of the gender effect is in line with the recurrent tendency for males to use a higher frequency of non-standard forms than females (Labov 1989: 205). This finding indicates that patterns of gender differentiation which have been documented in adult populations in connection with the use of non-standard morphosyntactic features are also evident in the language of preteen speakers.

---

<sup>1</sup> One of the principal advantages of using Goldvarb to carry out multivariate analyses of the data is that it can reveal the simultaneous impact of factor effects, and indicate the relative magnitude of individual effects, in addition to showing which effects attain levels of statistical significance (Tagliamonte 2006a: 215). However, Goldvarb is not without its drawbacks. From the practical perspective of using the program, Bayley (2002: 130) notes that variable rule programs have largely been restricted to research in sociolinguistics and second language acquisition, with the result that there is little general awareness of these programs in research establishments, thereby making it difficult for users to access advice and assistance when troubleshooting areas of difficulty. Secondly, as noted in 2.3.3, Goldvarb does not overtly check for interactions between factors (see Tagliamonte 2006a: 151). Although Tagliamonte (2006a: 151) does not view this as a defect in variable rule analysis, the onus is on the researcher to test for the presence of interaction effects systematically (Sigley 2003: 250). Of course, other programs and statistical measures could be conceivably used to analyse the data contained in this study. Other computer assisted techniques include Principal Components Analysis (see e.g. Horvath and Sankoff 1987). One of the strengths of this approach is that it first groups speakers in terms of their linguistic behaviour. The social interpretation of the data in terms of the social characteristics of the speakers then proceeds on the basis of the grouping of linguistic data which serves as the initial input to the program (Milroy and Gordon 2003: 166).

With other variables, there are maturational differences in the use of vernacular variants, as documented in Chapter 5 and 7 which show that there are significant shifts in children's use of the conversational historical present and the vernacular quotative *go* as they get older.<sup>2</sup> Structured variation in the use of both these variants is evident in the speech of the youngest age cohort (7-8 years), but age-related differences highlighted in the respective multivariate analyses suggest that the interval between the youngest and oldest speakers in the corpus is a period in which sociolinguistic competence in the use of these forms is consolidated. As mentioned at several points in this study, the age range investigated here coincides with significant developments in preadolescent social organization which have important linguistic correlates. It is during the later preadolescent years that children begin to take the speech of their peers as a model, 'resulting in convergence towards relatively focused norms' (Kerswill 1996: 191). It is also a period which witnesses an expansion in children's stylistic repertoires (Romaine 1984a). Thus, ability to alternate tense forms for stylistic purposes may be a concomitant of a more general expansion in narrative skills with age, as suggested in Chapter 5. According to Berman and Nir-Sagiv (2004: 369), the use of tense shifting as a rhetorical device is developmental in childhood: in the speech of younger children, tense shifts are locally motivated by temporal semantic relations between adjacent clauses, but with increasing maturity, tense shifting takes on more discourse-motivated functions associated with encoding differences between background and foreground in narrative. This perhaps points to an important difference between the acquisition of variables which encode discourse-

---

<sup>2</sup> The fact that age is statistically significant for two of the variables suggests that the linguistic differences between children examined within a very small age interval may well be greater than the differences which emerge from the examination of an equivalent age interval in adult populations, as hypothesised by Cameron (2005a: 54, f.n. 9). Linguistic differentiation in later preadolescence is consistent with current research which 'reveals marked developments in linguistic proficiency across the school years, interfacing with sociocognitive developments and increasing communicative competence' (Ravid and Berman 2006: 120).

pragmatic functions and variables from other levels of the grammar. Whereas Kerswill (1996: 192) argues that phonology is more or less fully developed by the age of 6 or 7, the acquisition of discourse-pragmatic variables may be acquired over a more protracted period. In fact, such a scenario is suggested by shifting frequencies in the use of discourse *like* across the two preadolescent age divisions examined in Chapter 6. Age differentiations in the use of this discourse particle point to specifically localised patterns of variation that emerge in later preadolescence. Although 7-8 year old children were found to use *like* in a variety of syntactic positions, *like* principally occurred in their speech with clause-internal constituents. On the other hand, in the speech of the older children, *like* had scope over larger discourse units. Furthermore, in both age cohorts quantitative variation in the syntactic position of *like* intersected in interesting ways with gender, exhibiting provocative parallels with similar patterns located in adult populations (D'Arcy 2005a). Methodologically, these findings emphasise the need to conduct more fine-grained analyses of younger age cohorts in order to detect trends that might otherwise be obscured in the study of aggregated data from broader age ranges (Llamas 2007:73; Eckert 2000).

Other important interactions between age and gender were uncovered in the distributional analysis of preadolescents' use of general extenders. The interplay between age and gender in patterns of general extender variation is all the more interesting given that this is the least frequent of the six variables targeted for analysis. Previous research suggests that the frequency of a variable in discourse has a bearing on its potential recruitment to fulfil sociosymbolic functions. Thus, as mentioned in Chapter 4, infrequent syntactic variables such as relative markers appear less readily available for social evaluation in comparison with high frequency



phonological variables (see Cheshire *et al.* 2005a: 139).<sup>3</sup> In the case of general extenders, however, their relative infrequency in the corpus does not appear to detract from the fact that females in the older preadolescent age cohort make markedly greater use of these forms than the boys.

The frequency of general extenders in the speech of the older girls, which is sharply differentiated from that of the older boys, and contrasts in turn with the usage of the younger preadolescents, may possibly be a discursive manifestation of different gendered orientations towards the construction of talk which emerge in childhood. According to Cameron (2005a), such gender-differentiated patterns may be explicated with reference to varying degrees of gender segregation which obtain at different points in the lifespan. Cameron (2005a: 24) argues that children's preference for same-gender affiliation is developmental in preteen populations, and emerges around the ages of 3 to 4, with the degree of separation at its height in middle childhood or early adolescence. Increasingly intensive interaction in single-sex peer groups offers a crucial context for gender-differentiation in speech and the crystallization of distinct styles of speaking (see Romaine 1999: 196).

Further evidence of different gendered interactional styles in preadolescence can be found in girls' consistently greater use of direct speech in both age cohorts, with, yet again, the widest degree of gender differentiation found in the oldest age group. That the widening of the degree of differentiation between the boys and the girls is not simply coincidental, or an artefact of the data set, is bolstered by previous research on a comparable age continuum. Particularly relevant is Cameron's (2005b) study of word initial and internal (dh) used by Chicago elementary school children in the fifth grade (10-11 year olds) and in the second grade (7-8 year olds), which

---

<sup>3</sup> The issue of frequency and saliency is a complex one, as Hoffmann (2004: 189) has demonstrated in his discussion of how low-frequency complex prepositions may be cognitively salient as well as engaged in processes of change.

revealed that the point spread separating the girls from the boys in relation to the use of this variable was larger in the fifth grade than in the second grade. Although the sociolinguistic variables in the Chicago study and my own investigation are drawn from different components of the grammar, both studies converge in showing that the frequency usage of certain variants becomes more sharply contrastive as children reach the 10-11 age range. However, as Cameron (2005a: 30) acknowledges, interactions between age and gender in the expression of sociolinguistic variation very much depend on the type of variable available for social assessment, with, for example, stable variables behaving differently from changes in progress. I examine this issue in more detail below where I review how the analyses presented in the previous chapters illuminate the extent to which preadolescents participate in patterns of language variation and change.

## 9.2 PARTICIPATION IN PATTERNS OF VARIATION AND CHANGE

What can we surmise from the multivariate and distributional analyses of the preadolescent data about the participation of preteen speakers in patterns of variation and change ?

Firstly, it is clear from the data presented in Chapter 6 that, contrary to previous reports (Miller and Weinert 1995), preadolescents actively engage in the use of discourse marker *like* which is known to be rapidly innovating throughout contemporary varieties of English (Cheshire *et al.* 2005a: 154). Furthermore, as pointed out earlier, preadolescents appear to be converging with adolescents in their use of discourse *like* in certain syntactic positions resulting in the loss of the more traditional clause-final variant. This process of convergence may well be a specific instantiation of the more general phenomenon of dialect levelling whereby locally or

socially marked variants are eradicated in favour of forms which have wider geographical currency (Milroy and Gordon 2003: 130). Further evidence of the eradication of recessive localised variants in the southeast can be inferred from the distributional analysis of relativization, which shows that the relative marker *as*, attested in the traditional dialects of the southeast (Edwards 1993: 228-9; Anderwald 2004: 190), is not encountered in the preadolescent corpus at all.

Additional indications of the participation of preteen speakers in patterns of variation and change can be adduced from the analysis of *was/were* variation. Recall that this is an area of grammar that has undergone longitudinal re-organization (Nevalainen 2006), and may in fact still be involved in ongoing structural reconfiguration in the London metropolitan area, as tentatively suggested by Anderwald (2001). However, in this case, the findings are complex as contrastive norms in inner and outer London suggest, contrary to claims that the non-standard grammar of the southeast appears 'remarkably homogenous' (Edwards 1993: 219), that there are nuanced grammatical differences within this area. Of particular interest in the preadolescent data is the reversal of the NP/PRO constraint attested since the Middle English period in northern British dialects, as well as in some southern ones.<sup>4</sup> That the absence of this constraint is not simply a feature of the incomplete acquisition of internal linguistic constraints by preadolescents is strengthened by the existence of similar patterns in other dialects in the southeast and East Anglia (Britain 2002a; Britain *et al.* 2005). The detection of such subtle grammatical conditioning graphically illustrates that even though features such as *was/were* variation may have

---

<sup>4</sup> Recall too that data from other corpora based on the vernacular of young Londoners suggest that where the NP/PRO constraint is not reversed, the magnitude of this effect is vitiated.

a widespread geographical distribution, 'quantitative differences may be hiding behind qualitative similarity' (Anderwald 2004: 193).<sup>5</sup>

Yet another interesting dimension of *was/were* variation is the distributional configuration of *was* and *were* to mark polarity distinctions. Whether this apparent case of remorphologization is an active change in progress in London requires access to robust longitudinal data spanning a range of age cohorts that is not yet available, but the fact that similar patterns have been documented in the southeast (and elsewhere) suggests that preadolescents are actively engaged in acquiring a vernacular pattern that differs from the putatively more basilectal pattern based on levelling to *was* and *wasn't* in all persons and numbers (Chambers 2004: 132).<sup>6</sup>

Although I mentioned at the outset of this study that the research literature suggests that preteen children can participate in language change (Roberts 2002), the nature of any claims about the engagement of preteen children in changes in progress must necessarily be moderated by the fact that it is not always entirely possible to rule out the existence of age-grading effects. As Guy *et al.* (1986: 30-33) observe, composite sources of evidence drawn from age, class and sex, as well as real-time historical evidence, are required to corroborate a bona fide change in progress. In the present study, the current non-availability of systematically collected data relating to the vernacular norms of older speakers in London often hampers any firm conclusions that can be drawn about children's participation in changes that may be percolating

---

<sup>5</sup> See also Poplack and Tagliamonte (1991: 316), who make the strong claim that in resolving the status of a variable feature in a given variety, 'it is not its current *existence* that is decisive, nor even its rates of occurrence, but its *distribution* in the language, as determined by the hierarchy of constraints conditioning its appearance.' In the light of this claim, it is noteworthy that one of the reasons why the southeast of England is believed not to be grammatically distinctive is because few sociolinguistic studies have dealt with this region in any depth (Anderwald 2004: 175).

<sup>6</sup> There are other viable theoretical frameworks for investigating non-standard concord patterns in preadolescence which I have not explored in this study (see e.g. Adger and Smith 2005). Cornips and Corrigan (2005b: 4) note that the exploration of grammatical variability among preadolescent language users opens up interesting avenues for studies which integrate variationist linguistics with more theoretical approaches.

through the speech community. A further complication arises in cases where genuine innovation is not easily disentangled from variation that may represent historical residue.

Several of these problems came to the fore in the study of tense variation in preadolescent narratives. While patterns of tense shifting between the simple past and the conversational historical present in preadolescent narratives are reminiscent of similar patterns documented in adult narratives (Schiffrin 1981), alternation between these tense forms and the present perfect is much less extensively described. Although such variation in the preadolescent corpus appears to have an essentially pragmatic orientation associated with plot advancement, segmentation of narrative events, and the evaluation of narrative episodes, it is unclear whether uses of the present perfect in narrative complicating action clauses suggest a possible grammatical innovation or whether, viewed against the diachronic backdrop of extensive variation between the preterite and the present perfect in earlier varieties of English (Denison 1993), they reflect the retention of earlier patterns that have persisted in non-standard dialects. However, taking the broader perspective adopted in Chapter 5, the fact that similar incursions of the present perfect into narrative contexts typically associated with the preterite have been recently documented in Antipodean varieties of English (see Engel and Ritz 2000; Cox 2005) indicates a possible cross-varietal 'drift' (see Sapir 1921; Trudgill 2004). Although speculative, the existence of such parallel developments would be typologically consistent with the tendency towards increasing use of periphrastic expressions in tense/aspect categories in modern spoken English.<sup>7</sup>

At the very least, the patterns which emerge from the investigation of preadolescents' use of the present perfect call into question Kortmann's (2006: 608)

---

<sup>7</sup> See e.g. the progressive, which has 'gently increased in the past few centuries' (Nesselhauf 2007:191).

recent assertion that ‘the dominant pattern across varieties of English [...] clearly is the simple past (increasingly) doing service for all the major uses of the present perfect apart from the continuative perfect.’ It may be the case that the use of the present perfect by preadolescents highlighted in this study represents a potential site of variation that may be modified as children get older (Roberts 2005: 162), and which may, ultimately, impact on the dialect patterns of the future. Clearly, the data in Chapter 5 point to the need for more substantive surveys of the use of the present perfect in contemporary varieties of English, particularly in view of Tagliamonte’s (1996: 389) remarks that ‘little, if anything, is known about the linguistic and extra-linguistic conditioning of variability in this area of the grammar.’

In other analyses of variation in the data, it is the restricted occurrence of variants in the corpus that are otherwise claimed to be rapidly innovating in international varieties of English which is remarkable, and which points to possible areas of discontinuity between preadolescent and adolescent patterns of variation in Britain that have not been extensively investigated (see also Eckert 1988: 185 on the difficulty in documenting such discontinuities). Particularly striking is the nominal occurrence of *be like* in the preadolescent data. Contrary to expectation, given the correlations between *be like* and young speakers which are pervasive in the research literature (Barbieri 2007: 27), this variant is marginalized in the preadolescents’ quotative cohort, which is largely circumscribed to just *say* and *go*. As pointed out in Chapter 7, this focused pattern of variation based primarily on *say* and *go* is not a defining feature of preadolescent usage reported in other varieties of English: preadolescents in North America clearly engage in the frequent use of *be like* (see D’Arcy 2004; Tagliamonte and D’Arcy 2004). It is not clear why preadolescents in North America should make robust use of *be like*, whereas preadolescents living in

proximity to the largest urban centre in Britain should not.<sup>8</sup> Although *be like* appears to have made inroads into the quotative cohort in British English at a later date than in North American varieties (see Buchstaller 2006a), recent studies (e.g. Baker *et al.* 2006) indicate that it appears to be increasing in British English, and displacing other quotative forms. However, as Barbieri (2007: 25) notes, most previous studies of quotative variation have been based on surveys conducted in one single location, 'with most of the current data [...] over-subscribed to a specific sector of the population, namely speakers between 18 and 24' (Tagliamonte and D'Arcy 2004: 494).<sup>9</sup>

In order to contextualize quotative variation in preadolescent usage in London vis-à-vis the usage of older speakers, intergenerational data are required. Without qualifying evidence from older speakers relating to the social distribution of quotative variation in London and the southeast, it is difficult to say how *be like* may be diffusing to successive generations, but, if *be like* is an incoming feature associated with the adolescent age cohort, then it may be the case, as Tagliamonte and D'Arcy (2004: 509) suggest, that preadolescents do not have extensive exposure to this feature until they are fully integrated into teenage peer networks.<sup>10</sup> However, what is intriguing in the London preadolescents' very restricted use of *be like* is that their

---

<sup>8</sup> Tagliamonte and D'Arcy (2004) do not consider speakers under the age of 10. Nevertheless, it is clear that *be like* accounts for approximately 50% of the quotative usage of 10-12 year olds in their data (2004: 502). If we compare this figure with the distribution of *be like* in the speech of the 10-11 year olds in the London preadolescent corpus, where this variant accounts for only 1% of this cohort's total quotative usage, then the difference is striking.

<sup>9</sup> This is apparently the case for British English more so than for North American varieties. The vast majority of studies on *be like* in British English (Tagliamonte and Hudson 1999; Carey 2004; Baker *et al.* 2006) target college-aged populations. Furthermore, the aforementioned studies were undertaken in the same geographical locale: York. One exception is Macaulay (2001), based on the speech of Glaswegian adults and adolescents.

<sup>10</sup> When preadolescent quotative usage is broadly contrasted with that of adolescents, there does seem to be evidence of an expansion in the range of variants used as children enter the teenage years, including greater use of semantically richer quotative forms such as *shout*, *yell*, etc. (see D'Arcy 2004: 332). Since many quotative variants are essentially lexical, it is not surprising that speakers restructure the quotative system with age as vocabulary continues to be acquired throughout the lifespan (see also Aitchison 2000 on possible psycholinguistic reasons related to lexical expansion in adolescence).

deployment of this quotative with direct speech furnishes provisional evidence of a functional expansion which matches the inferred direction of change documented for *be like* in other varieties of English (Tagliamonte and D'Arcy 2004). Recall too that unlike the quotatives *say* and *go* in the preadolescent corpus, *be like* is more frequent with the first person, which is also consonant with its functional profile in other varieties of English. Thus, even though the preadolescents make very limited use of this quotative overall, they nevertheless appear to be sensitive to some of the constraints which operate on this variant in older populations.

Whatever the precise reasons underlying the infrequency of *be like* in the preadolescents' quotative system, the paucity of this variant in the preadolescent data is in keeping with previous claims that preadolescents are able to adopt new features, while 'adolescents may be the most influential transmitters of change' (Kerswill 1996:177; see also Eckert 2000: 16).

### 9.3 EDUCATIONAL IMPLICATIONS OF THE RESEARCH

According to Watt and Smith (2005: 103), 'a recurring popular perception of innovations in the speech of young speakers of any variety, is that of 'sloppiness' or 'laziness'. This study, however, has highlighted the structured nature of grammatical variation in preadolescents' speech, many facets of which would nonetheless be considered non-standard in educational settings.

In the light of the findings presented in this study, it is a matter of importance that professionals who are concerned with assessing children's language should have 'a realistic concept of its variable and patterned nature' (Milroy 1987: 200; Watt and Smith 2005). This is particularly the case in the school setting where the evaluation of children's language, both spoken and written, is a recurrent focus of attention. In



practical terms, this means that teachers should be made aware of norms of spoken English as well as aspects of spoken language which reflect normal social and regional variation (Romaine 1984a). Failure to do so can result in what Milroy and Milroy (1991: 163) term 'unrealistic remediation strategies'. One such case that I witnessed in the school where the fieldwork was undertaken concerns children's use of discourse *like*. On one occasion, a teacher engaged in somewhat Sisyphean attempts to get children with whom she was talking to edit out discourse *like* from their speech on the grounds that it was 'unnecessary'. In this case, the teacher seemed to be operating on the assumption that discourse markers such as *like* are redundant features of language, rather than normal features of spoken discourse. Such assumptions on the part of professionals are often grounded in the norms of the codified written language (Milroy and Milroy 1991: 89) which are inappropriately used as a baseline for assessing spontaneous spoken discourse.

In order to counter such perceptions, detailed information about the normal patterns of variation in children's speech, including innovations that children may be engaging in, is of relevance to teachers and speech therapists, not only to avoid misdiagnosis of children's language abilities, but also to make professionals aware of potential conflict sites which children speaking non-standard dialects may encounter when acquiring the norms of standard English (see further Watt and Smith 2005).

Several of the findings arising from the present study could be usefully incorporated into teaching training programmes as part of language awareness modules in order to counter some of the pervasive negative evaluations of children's spoken language, and to encourage teachers to acknowledge and draw on children's speech as an educational resource for incorporating language work into the

classroom.<sup>11</sup> For example, the findings discussed in relation to children's use of discourse variation (e.g. *like* and *general extenders*) could be used to illustrate how aspects of children's spoken language reflect its interactive and collaborative nature. The findings for these particular features could also serve as a basis for discussing how the dynamics of spoken language affects the kinds of structures that speakers use, and how language varies from one context to another. Many of these discourse features are representative of the differences between informal spoken English and formal written English, and could be used to exemplify some of the typical differences between speech and writing.

Other findings relating to *was/were* variation and quotative variation could be used to illustrate how language changes over time, and could be more broadly situated within discussions of how English varies locally and globally. Stigmatised features such as non-standard *was*, and non-standard *what* used as a relative marker, could also be used in teacher training programmes to explore reactions to linguistic diversity and regional variation, and to explore ways in which beliefs about language are disseminated.

By making some of the findings discussed in the present study accessible to teachers in training, as well as to qualified teachers in educational settings, the expectation is that not only will teachers be more aware of the forms and functions of spoken and written English, as well as the use of linguistic variation to index social differences, they will be better equipped to broaden children's own knowledge about language and to encourage children to 'share their own experiences of linguistic diversity [...], empowering them to face the adult world' (Cheshire and Edwards 1993: 50).

---

<sup>11</sup> Several of the ideas explored in this section are also discussed in the document entitled *Introducing the Grammar of Talk* published by the Qualifications and Curriculum Authority in 2004, and circulated to schools.

#### 9.4 DIRECTIONS FOR FUTURE RESEARCH

The findings presented in the preceding chapters have been generated within the confines of a small-scale study. The interpretation of the data here will depend on future work using larger scale corpora which target a similar age range. In particular, it will be important to investigate whether similar patterns highlighted in the present study are found in the vernacular usage of other preteen speakers in the southeast as well as elsewhere. It will also be necessary to extend such investigations to younger school-aged populations than the one targeted in the present study which has furnished insights into variation in later preadolescence, but which has not explored the 'developmental genesis' of these patterns in early childhood (see Romaine 1984a: 82; Roberts 1994: 26).

Future work will also benefit from more fine-grained analyses of variation in preadolescence in order to probe the nexus between specific configurations of linguistic and stylistic variation in this age group. As Eckert (2000: 213) comments, 'while the individual variables in a dialect may correlate with various aspects of social membership and practice, most of them take on interpretable social meaning only in the context of the broader linguistic styles to which they contribute.' Young speakers' recruitment of linguistic variables in the construction of distinct styles of speaking may well operate in tandem with their engagement in the manipulation of other semiotic resources such as clothing, hair style and recreational activities (see Eckert 2001: 124), the investigation of which necessitates a larger ethnographic perspective than has been possible in the present study.

Yet another dimension that future work needs to consider is the changing social status of preadolescence. Although researchers often resort to chronological demarcations in their definition of preadolescence, it is important that studies are

sensitive to the fact that preadolescence itself is a dynamically situated social construct. Llamas (2007) observes that social and economic changes may mean that speakers experience life stages quite differently from previous generations. Economic factors such as the rise of new consumer-oriented markets targeting preadolescent 'tweenies' can have an impact on the life-style characteristics and norms of behaviour of this particular age group (Llamas 2007:72). The linguistic correlates of such changes open up interesting avenues for further research projects targeting this life stage.

To conclude, I hope to have provided in this study further empirical support for both the existence and evolution of structured heterogeneity in the language of preadolescents. The findings presented here corroborate the position taken by Roberts (1994: 177), who argues that 'just as the inclusion of variation is necessary to form a complete picture of child language, so is the inclusion of children in variational research important in formulating a complete picture of variation in the speech community.'

## APPENDIX

### SPEAKER INFORMATION DATABASE <sup>1</sup>

Pseudonym	Gender	Age Range	Father's Occupation	Mother's Occupation
Danny	male	10-11 years	unemployed	unknown
Patrick	male	10-11 years	factory worker	housewife
Gareth	male	10-11 years	own business	driver
Mikey	male	10-11 years	British Telecom Engineer	cook
Aaron	male	10-11 years	self-employed	learning support assistant
Darren	male	10-11 years	window cleaner	housewife
James	male	10-11 years	carpet fitter	shop assistant
Carl	male	10-11 years	surveyor	housewife
Harry	male	10-11 years	history teacher	nursery teacher
Jason	male	10-11 years	own retail business	veterinary assistant
David	male	10-11 years	painter and decorator	school secretary
William	male	10-11 years	unemployed	childminder
Alison	female	10-11 years	surveyor	housewife/ childminder
Charley	female	10-11 years	engineer	learning support assistant
Elizabeth	female	10-11 years	furniture retail	bank cashier
Ruth	female	10-11 years	retail	retail
Diana	female	10-11 years	furniture installation	housewife
Helen	female	10-11 years	postman	school lunchtime supervisor

<sup>1</sup> In cases where information about parents' occupations was not readily obtainable either from the children themselves, or existing records in the school, this gap is indicated in the table above as 'unknown'.

<b>Pseudonym</b>	<b>Gender</b>	<b>Age Range</b>	<b>Father's Occupation</b>	<b>Mother's Occupation</b>
Laura	female	10-11 years	car factory worker	personnel officer
Rosie <sup>2</sup>	female	10-11 years	car factory worker	personnel officer
Annie	female	10-11 years	printer	housewife
Chloe	female	10-11 years	builder	works in a bar
Abbie	female	10-11 years	builder	bank cashier
Tracey	female	10-11 years	carpet fitter	administrator
Casey	female	10-11 years	tax inspector	office worker
Sasha	female	10-11 years	bank cashier	housewife
Dawn	female	10-11 years	unknown	housewife
Carla	female	10-11 years	unknown	housewife
John	male	7-8 years	accountant	cleaner
Brian	male	7-8 years	history teacher	nursery teacher
Ryan	male	7-8 years	roofer	hairdresser
Liam	male	7-8 years	self-employed	teaching assistant
Doug	male	7-8 years	unknown	housewife
Graham	male	7-8 years	unknown	housewife
Alfie	male	7-8 years	repairs shoes	works in supermarket
Tony	male	7-8 years	delivers furniture	works in building society
Freddie	male	7-8 years	heating engineer	tailor
Sally	female	7-8 years	furniture retail	works in travel agents
Bella	female	7-8 years	works in nut factory	housewife

<sup>2</sup> Rosie is Laura's twin sister. The database includes a small number of children who are siblings.

<b>Pseudonym</b>	<b>Gender</b>	<b>Age Range</b>	<b>Father's Occupation</b>	<b>Mother's Occupation</b>
Denise	female	7-8 years	car dealer	learning support assistant
Lisa	female	7-8 years	office worker	learning support assistant
Danielle	female	7-8 years	office worker	machinist
Suzanne	female	7-8 years	works in a bank	receptionist
Jaimie	female	7-8 years	cleaner	housewife
Maria	female	7-8 years	unknown	housewife
Lola	female	7-8 years	unknown	cashier
Jackie	female	7-8 years	unknown	housewife
Alicia	female	7-8 years	delivery driver	housewife

GENERAL EXTENDER TYPES IN THE PREADOLESCENT CORPUS

General Extender	Number
and everything	29
and all that	26
and that	22
or something	19
and things	15
and stuff like that	10
and stuff	6
or anything	5
and things like that	2
and all	2
and everything like that	1
and all that kind of stuff	1
and everyone	1
and like that	1
and everywhere	1
and all them	1
or whatever	1
or nothing	1
or things	1
things	1
sort of thing	1
something like that	1
<b>TOTAL</b>	<b>148</b>



## REFERENCES

- Adger, David and Jennifer Smith (2005). Variation and the minimalist program.  
In Leonie Cornips and Karen Corrigan (eds.), pp. 149-78.
- Aijmer, Karin (1985). What happens at the end of our utterances ? The use of  
utterance-final tags introduced by *and* and *or*. *Papers from the 8<sup>th</sup>  
Scandinavian Conference of Linguistics*. Copenhagen: Institut for  
Philologie, Copenhagen University, pp. 366-89.
- Aijmer, Karin (2002). *English Discourse Particles: Evidence from a Corpus*.  
Amsterdam: John Benjamins.
- Aitchison, Jean (1983). Review of Suzanne Romaine *Sociohistorical Linguistics: Its  
Status and Methodology*. *Linguistics* 21: 629-44.
- Aitchison, Jean (1992). Relative clauses in Tok Pisin: Is there a natural pathway?  
In Marinel Gerritsen and Dieter Stein (eds.) *Internal and External Factors in  
Syntactic Change*. Berlin: Mouton de Gruyter, pp. 295-316.
- Aitchison, Jean (2000). Shuddering halt or sudden spurt? The linguistic development  
of (pre-)adolescents. In H.W.Kam and A. Pakir (eds.), *Recent Developments and  
Issues in Applied Linguistics: A Collection of SAAL (Singapore Association for  
Applied Linguistics) Lectures*. Singapore: EPB.
- Altendorf, Ulrike and Dominic Watt (2004). The dialects in the south of England:  
Phonology. In Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend  
Mesthrie, and Clive Upton (eds.) *Handbook of Varieties of English*. Vol.1:  
*Phonology*. Berlin: Mouton de Gruyter, pp. 178-203.
- Andersen, Elaine (1990). *Speaking with Style: The Sociolinguistic Skills of Children*.  
London: Routledge.

- Andersen, Gisle (1997). 'They gave us these yeah, and they like wanna see like how we talk and all that.' The use of *like* and other pragmatic markers in London teenage speech. In Ulla-Britt Kotsinas, Anna-Brita Stenström, and Anna-Malin Karlsson, (eds.) *Ungdomsspråk i Norden. MINS 43* Stockholm: Stockholmsuniversitet, Institutionen för nordiska språk, pp. 82-95.
- Andersen, Gisle (2000). The role of pragmatic marker *like* in utterance interpretation. In Gisle Andersen and Torsten Fretheim (eds.) *Pragmatic Markers and Propositional Attitude*. Amsterdam : John Benjamins, pp. 17-38.
- Andersen, Gisle (2001). *Pragmatic Markers and Sociolinguistic Variation: A Relevance-Theoretic Approach to the Language of Adolescents*. Amsterdam: John Benjamins.
- Anderwald, Lieselotte (2001). *Was-were* variation in non-standard British English today. *English World-Wide* 22: 1-21.
- Anderwald, Lieselotte (2002). *Negation in Non-standard British English: Gaps Regularizations and Asymmetries*. London: Routledge.
- Anderwald, Lieselotte (2004). The varieties of English spoken in the southeast of England: Morphology and syntax. In Bernd Kortmann, Edgar Schneider, Kate Burridge, Rajend Mesthrie, and Clive Upton (eds.), pp. 175-95.
- Ariel, Mira (1994). Pragmatic operators. In Richard Asher and John Simpson (eds.) *The Encyclopedia of Language and Linguistics*, Vol.6. Oxford: Pergamon, pp. 3250-53.
- Androutsopoulos, Jannis and Alexandra Georgakopoulou (2003). Discourse constructions of youth identities: Introduction. In Jannis Androutsopoulos and Alexandra Georgakopoulou (eds.) *Discourse Constructions of Youth Identities*. Amsterdam: John Benjamins, pp. 1-25.

- Auer, Peter, Frans Hinskens, and Paul Kerswill (eds.) (2005). *Dialect Change: Convergence and Divergence in European Languages*. Cambridge: Cambridge University Press.
- Bailey, Guy (2002). Real and apparent time. In Jack Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), pp. 312-32.
- Bailey, Guy, Natalie Maynor, and Patricia Cukor-Avila (1989). Variation in subject-verb concord in Early Modern English. *Language Variation and Change* 1: 285-300.
- Baird, Sarah (2001). How *to be like* a Kiwi: Verbs of quotation in New Zealand English. *New Zealand English Journal* 15: 6-19.
- Baker, Zipporah, David Cockeram, Esther Danks, Mercedes Durham, Bill Haddican, and Louise Tyler (2006). The expansion of *be like*: Real time evidence from England. Paper given at NWAV 35, University of Ohio, Columbus, Ohio, 11<sup>th</sup> November, 2006.
- Bakhtin, Mikhail (1981a). *The Dialogic Imagination*. Austin: University of Texas Press.
- Bakhtin, Mikhail (1981b). Discourse in the novel. In M. Holmquist (ed.) *Dialogic Imagination*. Austin: University of Texas Press, pp. 259-422.
- Ball, Catherine (1996). A diachronic study of relative markers in spoken and written English. *Language Variation and Change* 8: 227-58.
- Ball, Catherine and Mira Ariel (1978). Or something, etc. *Penn Review of Linguistics* 3: 35-45.
- Ball, Martin (ed.) (2005). *Clinical Sociolinguistics*. Oxford: Blackwell.
- Bamberg, Michael (1990) The German Perfekt: Form and function of tense alternations. *Studies in Language* 14: 253-90.

- Bamberg, Michael and Victoria Marchman (1990) What holds a narrative together ?  
The linguistic encoding of episode boundaries. *Pragmatics* 4: 58-121.
- Barbieri, Frederica (2005). Quotative use in American English: A corpus-based, cross-register comparison. *Journal of English Linguistics* 33: 222-56.
- Barbieri, Frederica (2007). Older men and younger women: A corpus-based study of quotative use in American English. *English World-Wide* 28: 23-45.
- Barsalou, Lawrence (1983). Ad hoc categories. *Memory and Cognition* 11: 211-27.
- Bauer, Laurie (1994). *Watching English Change: An Introduction to the Study of Linguistic Change in Standard Englishes in the Twentieth Century*. London: Longman.
- Bayley, Robert (2002). The quantitative paradigm. In Jack Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), pp. 117- 41.
- Beal, Joan and Karen Corrigan (2002). Relatives in Tyneside and Northumbrian English. In Patricia Poussa (ed.), pp. 125-34.
- Beal, Joan and Karen Corrigan (2005). A tale of two dialects: Relativization in Newcastle and Sheffield. In Markku Filppula, Juhani Klemola, Marjatta Palander and Esa Penttilä (eds.) *Dialects Across Borders*. Amsterdam: John Benjamins, pp. 87-117.
- Berman, Ruth and Dan Slobin (1994). Becoming a proficient speaker. In Ruth Berman and Dan Slobin (eds.) *Relating Events in Narrative: A Cross-Linguistic Developmental Study*. Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 597-610.

- Berman, Ruth and Bracha Nir-Sagiv (2004). Linguistic indicators of inter-genre differentiation in later language development. *Journal of Child Language* 31: 339-80.
- Bever, Terence and Terence Langendoen (1972). A dynamic model of evolution of language. *Linguistic Inquiry* 2: 433-65.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.
- Blyth, Carl, Sigrid Recktenwald and Jenny Wang (1990). I'm like, say what?!: A new quotative in American oral narrative. *American Speech* 65 : 215-27.
- Bolden, Galina (2004). The quote and beyond: Defining boundaries of reported speech in conversational Russian. *Journal of Pragmatics* 36: 1071-1118.
- Breivik, Leiv Egil (1990). *Existential there: A Synchronic and Diachronic Study*. Oslo: Novus.
- Breivik, Leiv Egil (1999). On the pragmatic function of relative clauses and locative expressions in existential sentences in the LOB Corpus. In Hilde Hasselgård and Signe Oksefjell (eds.) *Out of Corpora: Studies in honour of Stig Johansson*. (Language and Computers: Studies in Practical Linguistics, 26). Amsterdam and Atlanta, GA: Rodopi, pp. 121-35.
- Brinton, Laurel (1990). The development of discourse markers in English. In Jacek Fisiak (ed.) *Historical Linguistics and Philology*. Berlin: Mouton de Gruyter, pp. 45-71.
- Brinton, Laurel (1992). The historical present in Charlotte Brontë's novels: Some discourse functions. *Style* 2: 221-44.

- Brinton, Laurel. (1996). *Pragmatic Markers in English: Grammaticalisation and Discourse Functions*. Berlin: Mouton de Gruyter.
- Brinton, Laurel (2001). Historical discourse analysis. In Deborah Schiffrin, Deborah Tannen and Heidi Hamilton (eds.), pp. 138-60.
- Brinton, Laurel (2006). Pathways in the development of pragmatic markers in English. In Ans van Kemenade and Bettelou Los (eds.), pp. 307-334.
- Britain, David (1992). Things *and that* in Porirua: An analysis of set marking tags. Notes from departmental seminar. Wellington Linguistics Seminar Series.
- Britain, David (1998). High rising terminals in New Zealand English: Who uses them, when and why? *Essex Research Reports in Linguistics* 21: 33-58.
- Britain, David (2001). Review of Paul Foulkes and Gerard Docherty (eds.) *Urban Voices: Accent Studies in the British Isles*. *English World Wide* 22: 121-128.
- Britain, David (2002a). Diffusion, levelling, simplification and reallocation in past tense *BE* in the English Fens. *Journal of Sociolinguistics* 6: 16-43.
- Britain, David (2002b). Space and spatial diffusion. In Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), pp. 603-37.
- Britain, David (2005). Innovation, diffusion, Estuary English and local dialect differentiation: The survival of Fenland Englishes. *Linguistics* 43: 995-1022.
- Britain, David (*in press*). Grammatical variation in England. In David Britain (ed.) *Language in the British Isles*. Cambridge: Cambridge University Press.
- Britain, David and Andrea Sudbury (2000). There's sheep and there's penguins: Drift and the use of singular verb forms of *BE* in plural existential clauses in New Zealand and Falkland Island English. *Essex Research Reports in Linguistics* 28: 1-32.

- Britain, David, Laura Rupp, Susan Fox, Michelle Bray, Susan Baker and Juliette Spurling (2005). The Northern and East Anglian subject rules: Two patterns, one rule ? Invited talk, Department of Linguistics, Queen Mary, London University, 20<sup>th</sup> October 2005.
- Brown, Penelope and Stephen Levinson (1987). *Politeness*. Cambridge: Cambridge University Press.
- Buchstaller, Isabelle (2001). An alternative view of *like*: Its grammaticalisation in conversational American English and beyond. *Edinburgh Working Papers in Applied Linguistics* 11: 21-41.
- Buchstaller, Isabelle (2003). The co-occurrence of quotatives with mimetic performances. *Edinburgh Working Papers in Applied Linguistics* 12: 1-9.
- Buchstaller, Isabelle (2004). The sociolinguistic constraints on the quotative system- British English and US English compared. PhD thesis. University of Edinburgh.
- Buchstaller, Isabelle (2006a). Diagnostics of age-graded linguistic behaviour: The case of the quotative system. *Journal of Sociolinguistics* 10: 3-30.
- Buchstaller, Isabelle (2006b). Social stereotypes, personality traits and regional perception displaced: Attitudes towards the new quotatives in the U.K. *Journal of Sociolinguistics* 10: 362-81.
- Burnley, David (1983). *A Guide to Chaucer's Language*. London: MacMillan.
- Butler, Judith (1990). *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.
- Butters, Ronald (1980). Narrative *go say*. *American Speech* 55: 304-7.
- Bybee, Joan (1988). The diachronic dimension in explanation. In John Hawkins (ed.) *Explaining Language Universals*. Oxford: Blackwell, pp. 350-79.

- Bybee, Joan (2003). Mechanisms of change in grammaticization: The role of frequency. In Brian D. Joseph and Richard D. Janda (eds.) *The Handbook of Historical Linguistics*. Oxford: Blackwell, pp. 602-23.
- Bybee, Joan and Carol Moder (1983). Morphological classes as natural classes. *Language* 59: 251-70.
- Bybee, Joan, Revere Perkins, and William Pagliuca (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Bybee, Joan and Paul Hopper (2001). Introduction to frequency and the emergence of linguistic structure. In Joan Bybee and Paul Hopper (eds.) *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, pp. 1-24.
- Bynon, Theodora (1977). *Historical Linguistics*. Cambridge: Cambridge University Press.
- Cameron, Deborah (2000). *Good to Talk*. London: Sage publications.
- Cameron, Deborah (2005). Language, gender and sexuality: Current issues and new directions. *Applied Linguistics* 26: 482-502.
- Cameron, Deborah, Fiona McAlinden and Kathy O'Leary (1989). Lakoff in context: The social and linguistic functions of tag questions. In Jennifer Coates and Deborah Cameron (eds.) *Women in their Speech Communities*. London: Longman, pp. 74 – 93.
- Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, Ben Rampton and Kay Richardson (1993). Ethics, advocacy and empowerment: Issues of method in researching language. *Language and Communication* 13: 81-94.
- Cameron, Richard (1998). A variable syntax of speech, gesture, and sound effect: Direct quotations in Spanish. *Language Variation and Change* 10: 43-83.



- Cameron, Richard (2000). Language change or changing selves? Direct quotation strategies in the Spanish of San Juan, Puerto Rico. *Diachronica* 17: 249-92.
- Cameron, Richard (2005a). Aging and gendering. *Language in Society* 34: 23-61.
- Cameron, Richard (2005b). Growing up and apart: Gender divergences in two Chicagoland elementary schools. Paper presented at NWAV 34, New York University, New York City, 21<sup>st</sup> October, 2005.
- Campbell, Lyle (1980). Towards new perspectives in American Finnish. In Osmo Ihola (ed.) *Congressus Quintus Internationalis Fenno-Ugristarum*. Pars III Turku: Suomen Kielen Seura, pp. 341-51.
- Carey, Lucy (2004). I was like, 'Oh God.' The changing quotative system in British English. *Durham Working Papers in Linguistics* 10: 1-13.
- Cedergren, Henrietta and David Sankoff (1974). Variable rules: Performance as a statistical reflection of competence. *Language* 50: 333-55.
- Chambers, Jack (2003). *Sociolinguistic Theory* 2<sup>nd</sup> edn. Oxford: Blackwell.
- Chambers, Jack (2004). Dynamic typology and vernacular universals. In Bernd Kortmann (ed.) *Dialectology meets Typology*. Berlin: Mouton de Gruyter, pp.127-45.
- Chambers, Jack, Peter Trudgill, and Natalie Schilling-Estes (eds.) (2002). *The Handbook of Language Variation and Change*. Oxford: Blackwell.
- Channell, Joanna (1994). *Vague Language*. Oxford: Oxford University Press.
- Cheshire, Jenny (1982a). *Variation in an English Dialect: A Sociolinguistic Study*. Cambridge: Cambridge University Press.
- Cheshire, Jenny (1982b). Linguistic variation and social function. In Suzanne Romaine (ed.) *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold, pp. 153-66.

- Cheshire, Jenny (1987). Syntactic variation, the linguistic variable and sociolinguistic theory. *Linguistics* 25: 257-82.
- Cheshire, Jenny (1989). Addressee-oriented features in spoken discourse. *York Papers in Linguistics* 13: 49-63.
- Cheshire, Jenny (ed.) (1991). *English around the World: Sociolinguistic Perspectives*. Cambridge: Cambridge University Press.
- Cheshire, Jenny (1994). Standardization and the English irregular verb. In Dieter Stein and Ingrid Tieken-Boon van Ostade (eds.) *Towards a Standard English, 1600-1800*. Topics in English Linguistics 12. Berlin: Mouton de Gruyter, pp. 115-33.
- Cheshire, Jenny (1996). Syntactic variation and the concept of prominence. In Juhani, Klemola, Merja, Kytö and Matti Rissanen (eds.) *Speech Past and Present: Studies in English Dialectology in Memory of Ossi Ihalainen*. Frankfurt : Peter Lang, pp. 1-17.
- Cheshire, Jenny (1999a). Spoken standard English. In Tony Bex and Richard Watts (eds.) *Standard English: The Widening Debate*. London: Routledge, pp. 129-48.
- Cheshire, Jenny (1999b). Taming the vernacular: Some repercussions for the study of syntactic variation and spoken grammar. *Cuadernos de Filología Inglesa* 8: 59-80.
- Cheshire, Jenny (2002a). Sex and gender in variationist research. In Jack Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), pp. 423-43.
- Cheshire, Jenny (2002b). Who we are and where we're going: Language and identities in the new Europe. In Mike Holt and Paul Gubbins (eds.) *Beyond Boundaries: Language and Identity in Contemporary Europe* Clevedon: Multilingual Matters, pp. 19-34.

- Cheshire, Jenny (2005a). Syntactic variation and beyond: Gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics* 9: 479-508.
- Cheshire, Jenny (2005b). *The Queen Mary Narrative Corpus*. Queen Mary, University of London.
- Cheshire, Jenny (*in press*). Discourse variation, grammaticalisation and stuff like that. To appear in the *Journal of Sociolinguistics*.
- Cheshire, Jenny and Viv Edwards (1993). Sociolinguistics in the classroom: Exploring linguistic diversity. In James Milroy and Lesley Milroy (eds.), pp. 34-52.
- Cheshire, Jenny, Viv Edwards and Pamela Whittle (1993). Non-standard English and dialect levelling. In James Milroy and Lesley Milroy (eds.) pp. 53- 96.
- Cheshire, Jenny, Paul Kerswill and Ann Williams (1999). The role of adolescents in dialect levelling. Final Report to the Economic and Social Research Council. (Project Number R000236180).
- Cheshire, Jenny and Ann Williams (2002). Information structure in male and female adolescent talk. *Journal of English Linguistics* 30: 217-38.
- Cheshire, Jenny, Paul Kerswill and Ann Williams (2005a). Phonology, grammar and discourse in dialect convergence. In Peter Auer, Frans Hinskens and Paul Kerswill (eds.), pp. 135-67.
- Cheshire, Jenny, Susan Fox, Paul Kerswill and Eivind Torgersen (2005b). Reversing drift: changes in the London diphthong system. Paper presented at UKLVC5, University of Aberdeen, 12<sup>th</sup> September 2005.

- Cheshire, Jenny and Susan Fox (2006). New perspectives on *was/were* variation in London. Paper presented at NWAV 35, Ohio State University, 11<sup>th</sup> November 2006.
- Clark, Herbert and Richard Gerrig (1990). Quotations as demonstrations. *Language* 66: 764-805.
- Cofer, Thomas (1975). Performance constraints on relative pronoun deletion. *Linguistics* 15: 13-32.
- Comrie, Bernard (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press
- Comrie, Bernard (1985). *Tense*. Cambridge: Cambridge University Press.
- Cornips, Leonie and Karen Corrigan (2005a). Convergence and divergence in grammar. In Peter Auer, Frans Hinskens and Paul Kerswill (eds.), pp. 96-134.
- Cornips, Leonie and Karen Corrigan (2005b). Toward an integrated approach to syntactic variation: A retrospective and prospective synopsis. In Leonie Cornips and Karen Corrigan (eds.), pp. 1-27.
- Cornips, Leonie and Karen Corrigan (eds.) (2005c). *Syntax and Variation: Reconciling the Biological and the Social*. Amsterdam: John Benjamins.
- Corrigan, Karen (2000). What bees to be maun be: Aspects of deontic and epistemic modality in a northern dialect of Irish English. *English World-Wide* 21: 25-62.
- Coulmas, Florian (ed.) (1997). *The Handbook of Sociolinguistics*. Oxford: Blackwell.
- Couper-Kuhlen, Elizabeth (1998). Coherent voicing. On prosody in conversational reported speech. *Interaction and Linguistic Structures (InLiSt) No.1* Konstanz: University of Konstanz.
- Coupland, Nik (1988). *Dialect in Use: Sociolinguistic Variation in Cardiff English*. Cardiff: University of Wales Press.

- Cox, Rachel (2005). Preterite uses of the present perfect in New Zealand English narratives: A case study. M.A. dissertation, Department of Linguistics, University of Canterbury, New Zealand.
- Croft, William (1990). *Typology and Universals*. Cambridge: Cambridge University Press.
- Crystal, David and Derek Davy (1975). *Investigating English Style*. London: Longman.
- Cukor-Avila, Patricia (2002). *She say, she go, she be like*: Verbs of quotation over time in African American Vernacular English. *American Speech* 77: 3-30.
- Cukor-Avila, Patricia, and Guy Bailey (1995). Grammaticalization in AAVE. In Jocelyn Ahlers, Leela Bilmes, Joshua Guenter, Barbara Kaiser, and Ju Namkung (eds.) *Proceedings of the Twenty-First Annual Meeting of the Berkeley Linguistics Society* (BLS 21). Berkeley: Department of Linguistics, University of California, pp. 401-13.
- D'Arcy, Alex (2004). Contextualizing St. John's youth English within the Canadian quotative system. *Journal of English Linguistics* 32: 323-45.
- D'Arcy, Alex (2005a). *LIKE*: Syntax and development. PhD Dissertation, University of Toronto.
- D'Arcy, Alex (2005b). The development of linguistic constraints: Phonological innovations in St. John's English. *Language Variation and Change* 17: 327-55.
- Dailey-O'Cain, Jennifer (2000). The sociolinguistic distribution and attitudes towards focuser *like* and quotative *like*. *Journal of Sociolinguistics* 4: 60-80.

- Dasinger, Lisa and Cecile Toupin (1994). The development of relative clause functions in narrative. In Ruth Berman and Dan Slobin, (eds.) *Relating Events in Narrative: A Cross-linguistic Developmental Study*. Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 457- 514.
- De Mauro, Tullio, Frederico Mancini, Maximo Vedovello, and Miriam Voghera (1993). *Lessico di Frequenza dell'Italiano Parlato*. Milan: Etaslibri.
- Denison, David (1993). *English Historical Syntax: Verbal Constructions*. London: Longman.
- Denison, David (1998). Syntax. In Suzanne Romaine (ed.) *Cambridge History of the English Language*. Vol IV. Cambridge: Cambridge University Press, pp. 92-326.
- De Villiers, Jill, Helen Tager Flusberg, Kenji Hakuta, and Michael Cohen (1979). Children's comprehension of relative clauses. *Journal of Psycholinguistic Research* 8: 499-518.
- Diessel, Holger and Michael Tomasello (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics* 11: 131-51.
- Diessel, Holger and Michael Tomasello (2005). A new look at the acquisition of relative clauses. *Language* 81: 882-906.
- Dines, Elizabeth (1980). Variation in discourse – and stuff like that. *Language in Society* 9: 13-31.
- Dion, Nathalie and Shana Poplack (2005). 'Im like: I know, I do the same thing': Can minority speakers catch up with the Joneses? Paper presented at NWAV34, New York University, 21<sup>st</sup> October 2005.
- Dubois, Sylvie (1992). Extension particles, etc. *Language Variation and Change* 4: 179-203.

- Dubois, Sylvie and David Sankoff (2001). The variationist approach toward discourse structural effects and socio-interactional dynamics. In Deborah Schiffrin, Deborah Tannen, and Heidi Hamilton, (eds.), pp. 282-303.
- Eckert, Penelope (1988). Adolescent social structure and the spread of linguistic change. *Language in Society* 17: 183-207.
- Eckert, Penelope (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1: 245-67.
- Eckert, Penelope (1997a). Age as a sociolinguistic variable. In Florian Coulmas (ed.), pp. 151-67.
- Eckert, Penelope (1997b). Gender, race and class in the preadolescent marketplace of identities. Paper presented at the Annual Meeting of the American Anthropological Association, Washington, DC.
- Eckert, Penelope (2000). *Linguistic Variation as Social Practice*. Oxford: Blackwell.
- Eckert, Penelope (2001). Style and social meaning. In Penelope Eckert and John R. Rickford (eds.) *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press, pp. 119-26.
- Eckert, Penelope (2003). Language and gender in adolescence. In Janet Holmes and Miriam Meyerhoff (eds.) *The Handbook of Language and Gender*. Oxford: Blackwell, pp. 381-400.
- Eckert, Penelope (2004). Adolescent language. In Edward Finegan and John R. Rickford (eds.) *Language in the USA*. Cambridge: Cambridge University Press, pp. 361-74.
- Edwards, Viv (1993). The grammar of southern British English. In James Milroy and Lesley Milroy (eds.), pp. 214-38.

- Eisikovits, Edina (1991). Variation in subject-verb agreement in Inner Sydney English. In Jenny Cheshire (ed.), pp. 235-56.
- Elsness, Johan (1997). *The Perfect and the Preterite in Contemporary and Earlier English*. Berlin: Mouton de Gruyter.
- Ely, Richard and Allyssa McCabe (1993). Remembered voices. *Journal of Child Language* 20: 671-96.
- Ely, Richard, Jean Berko Gleason and Allysa McCabe (1996). Why didn't you talk to your mommy, honey? Parents' and children's talk about talk. *Research on Language and Social Interaction* 29: 7-25.
- Engel, Dulcie and Marie-Eve Ritz (2000). The use of the present perfect in Australian English. *Australian Journal of Linguistics* 20: 119-40.
- Eriksson, Mats (1995). A case of grammaticalization in modern Swedish: The use of *ba* in adolescent speech. *Language Sciences* 17: 19-48.
- Erman, Britt (1992). Female and male usage of pragmatic expressions in same-gender and mixed-gender interaction. *Language Variation and Change* 4: 217-34.
- Erman, Britt (2001). Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk. *Journal of Pragmatics* 33: 1337-59.
- Escure, Geneviève (1991). Gender roles and linguistic variation in the Belizean Creole community. In Jenny Cheshire (ed.), pp. 595-606.
- Farghal, Mohammed and Madeline Haggan (2005). The tag *and everything* revisited: The case of *u-kulši* in Arabic. *Multilingua* 24: 399-412.
- Feagin, Crawford (1979). *Variation and Change in Alabama English: A Sociolinguistic Study of the White Community*. Washington, DC: Center for Applied Linguistics.



- Ferrara, Kathleen (1997). Form and function of the discourse marker *anyway*: Implications for discourse analysis. *Linguistics* 35: 343-78.
- Ferrara, Kathleen and Barbara Bell (1995). Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of *be + like*. *American Speech* 70: 265-89.
- Fischer, Kerstin (2000). *From Cognitive Semantics to Lexical Pragmatics: The Functional Polysemy of Discourse Particles*. Berlin: Mouton de Gruyter.
- Fischer, Olga (1992). Syntax. In Norman Blake (ed.) *The Cambridge History of the English Language*. Vol II. Cambridge: Cambridge University Press, pp. 207-408.
- Fischer, Olga, Ans van Kemenade, Willem Koopman and Wim van der Wurff (eds.) (2000). *The Syntax of Early English*. Cambridge: Cambridge University Press.
- Fleischman, Suzanne (1985). Discourse functions of tense-aspect oppositions : Towards a theory of grounding. *Linguistics* 23: 851-82.
- Fleischman, Suzanne (1986). Evaluation in narrative: The present tense in medieval performed stories. *Yale French Studies* 70: 199-251.
- Fleischman, Suzanne (1990). *Tense and Narrativity: From Medieval Performance to Modern Fiction*. Austin: University of Texas Press.
- Fleischman, Suzanne (1995). Imperfective and irrealis. In Joan Bybee and Suzanne Fleischman (eds.) *Modality in Grammar and Discourse*. Amsterdam: John Benjamins, pp. 519-51.
- Fleischman, Suzanne (1999). Pragmatic markers in comparative and historical perspective: Theoretical implications of a case study. Paper presented at the 14<sup>th</sup> International Conference on Historical Linguistics, Vancouver, Canada, August 1999.

- Fleischman, Suzanne and Marina Yaguello (2004). Discourse markers across languages. In Carol Moder and Aida Martinovic-Zic (eds.) *Discourse Across Languages and Cultures*. Amsterdam: John Benjamins, pp. 129-47.
- Fludernik, Monika (1991). The historical present tense yet again: Tense switching and narrative dynamics in oral and quasi-oral storytelling. *Text* 11: 365-97.
- Flynn, Suzanne and Barbara Lust (1980). Acquisition of relative clauses: Developmental changes in their heads. *Cornell Working Papers in Linguistics* 1: 33-45.
- Forby, Robert (1830/ 1970). *The Vocabulary of East Anglia*. New York: A.M. Kelley.
- Forrström, Gosta (1948). *The Verb 'TO BE' in Middle English: A Survey of the Forms*. Lund: C.W.K. Gleerup.
- Foulkes, Paul and Gerard Docherty (eds.) (1999). *Urban Voices: Accent Studies in the British Isles*. London: Arnold.
- Foulkes, Paul, Gerard Docherty, and Dominic Watt (1999). Tracking the emergence of structured variation. *Leeds Working Papers in Linguistics and Phonetics* 7: 1-25.
- Foulkes, Paul, Gerard Docherty and Dominic Watt (2005). Phonological variation in child-directed speech. *Language* 81: 177-206.
- Fox, Barbara and Sandra Thompson (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language* 66: 297-316.
- Fraser, Bruce (1999). What are discourse markers? *Journal of Pragmatics* 31: 931-52.
- Fuller, Janet (2003). Use of the discourse marker *like* in interviews. *Journal of Sociolinguistics* 7: 365-77.

- Geisler, Christine (2002). Relativization in Ulster English. In Patricia Poussa (ed.), pp. 135-46.
- Georgakopoulou, Alexandra (1994). Contextualizing modern Greek storytelling: The case of the historical present and constructed dialogue. *Journal of Modern Greek Studies* 12: 203-19.
- Georgakopoulou, Alexandra (1997). *Narrative Performances: A Study of Modern Greek Storytelling*. Amsterdam: John Benjamins.
- Georgakopoulou, Alexandra and Dionysis Goutsos (1997). *Discourse Analysis*. Edinburgh: Edinburgh University Press.
- Givón, Talmy (1983). *Topic Continuity in Discourse*. Amsterdam: Benjamins.
- Godfrey, Elizabeth and Sali Tagliamonte (1999). Another piece for the verbal –s story: Evidence from Devon in southwest England. *Language Variation and Change* 11: 87-121.
- Goffman, Erving (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Golato, Andrea (2000). An innovative German quotative for reporting on embodied actions: *Und ich so/ und er so* and *I'm like/and he's like*. *Journal of Pragmatics* 32: 29-54.
- Goodwin, Marjorie Harness (1990). *He-said-she-said: Talk as Social Organization among Black Children*. Bloomington, Indiana: Indiana University Press.
- Greenberg, Joseph (1966). *Language Universals*. The Hague, Paris: Mouton.
- Guasti, Maria Teresa and Anna Cardinaletti (2003). Relative clause formation in Romance child's production. *Probus* 15: 47-89.

- Günther, Susanne (1999). Polyphony and the layering of voices in reported dialogues: An analysis of prosodic devices in everyday reported speech. *Journal of Pragmatics* 31: 685-708.
- Guy, Gregory (1990). The sociolinguistic types of language change. *Diachronica* 7: 47-67.
- Guy, Gregory, Barbara Horvath, Julia Vonwiller, Elaine Daisley, and Inge Rogers (1986). An intonational change in progress in Australian English. *Language in Society* 15: 23-52.
- Habbick, Timothy (1991). Burnouts versus rednecks: Effects of group membership on the phonemic system. In Penelope Eckert (ed.) *New Ways of Analysing Sound Change*. San Diego, CA: Academic Press, pp. 185-212.
- Haiman, John (1988). Incorporation, parallelism and focus. In Michael Hammond, Edith Moravcsik and John R. Wirth (eds.) *Studies in Syntactic Typology*. Amsterdam: John Benjamins, pp. 303-20.
- Hansen, Maj-Britt Mosegaard (1997). *Alors and donc* in spoken French: A reanalysis. *Journal of Pragmatics* 28: 153-87.
- Hansen, Maj-Britt Mosegaard (1998). *The Function of Discourse Particles: A Study with Reference to Spoken French*. Amsterdam: John Benjamins.
- Harris, John (1984). Syntactic variation and dialect divergence. *Journal of Linguistics* 20: 303-27.
- Hay, Jennifer and Daniel Schreier (2004). Reversing the trajectory of language change: Subject-verb agreement with *BE* in New Zealand English. *Language Variation and Change* 16: 209-35.
- Hazen, Kirk (1998). The birth of a variant: Evidence for a tripartite negative past *BE* paradigm. *Language Variation and Change* 10: 221-44.

- Heine, Bernd and Tania Kuteva (2002). *World Lexicon of Grammaticalisation*.  
Cambridge: Cambridge University Press.
- Henry, Alison (1995). *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: Oxford University Press.
- Henry, Alison (2002). Variation and syntactic theory. In Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), pp. 267-82.
- Henry, Alison (2005). Idiolectal variation and syntactic theory. In Leonie Cornips and Karen Corrigan (eds.), pp. 109-22.
- Hernández-Campoy, Juan Manuel (1999). Geolinguistic models of analysis of the spatial diffusion of sociolinguistic innovations. *Studia Anglica Posnaniensia* 34: 7-42.
- Herrmann, Tanja (2005). Relative clauses in English dialects of the British Isles. In Bernd Kortmann, Tanja Herrmann, Lukas Pietsch and Susanne Wagner (eds.), pp. 21-105.
- Hickmann, Maya (1993). The boundaries of reported speech in narrative discourse: Some developmental aspects. In John Lucy (ed.), pp. 63-90.
- Hickmann, Maya (1995). Discourse, organization and the development of reference to person, space, and time. In Paul Fletcher and Brian MacWhinney (eds.) *The Handbook of Child Language*. Oxford: Blackwell, pp. 194-218.
- Hoffmann, Sebastian (2004). Are low-frequency complex prepositions grammaticalized? In Christian Mair (ed.) *Corpus Approaches to Grammaticalization in English*. Amsterdam: John Benjamins, pp. 171-210.
- Hogg, Richard (1988). *Snuck*: The development of irregular preterit forms. In Graham Nixon and John Honey (eds.) *An Historic Tongue: Studies in English Linguistics in Memory of Barbara Strang*. London: Routledge, pp. 31-40.

- Holmes, Janet (1997). Women, language and identity. *Journal of Sociolinguistics* 1: 195-223.
- Hopper, Paul (1979). Aspect and foregrounding in discourse. In Talmy Givón (ed.) *Syntax and Semantics* Vol. 12: *Discourse and Syntax*. New York: Academic Press, pp. 213-41.
- Hopper, Paul (1991). On some principles of grammaticalisation. In Elizabeth Traugott and Bernd Heine (eds.) *Approaches to Grammaticalisation* Vol. 1. Amsterdam: Benjamins, pp. 17-35.
- Hopper, Paul and Elizabeth Traugott (1993). *Grammaticalisation*. Cambridge: Cambridge University Press.
- Horvath Barbara and David Sankoff (1987). Delimiting the Sydney speech community. *Language in Society* 16: 179-204.
- Horvath, Barbara and Ronald Horvath (2003). A closer look at the constraint hierarchy: Order, contrast and geographical scale. *Language Variation and Change* 15: 143-70.
- Howe, Chad and Scott Schwenter (2003). Present perfect for preterite across Spanish dialects. *University of Pennsylvania Working Papers in Linguistics* 9: 61-75.
- Huddleston, Rodney (1971). *The Sentence in Written English*. Cambridge: Cambridge University Press.
- Hudson, Richard (1985). The limits of subcategorization. *Linguistic Analysis* 15: 233-55.
- Hudson, Richard (1990). *English Word Grammar*. Oxford: Blackwell.
- Hughes, Arthur and Peter Trudgill (1979). *English Accents and Dialects*. London: Edward Arnold.

- Ihalainen, Ossi (1980). Relative clauses in the dialect of Somerset. *Neuphilologische Mitteilungen* 81: 187-96.
- Ihalainen, Ossi (1994). The dialects of England since 1776. In Robert Burchfield (ed.) *The Cambridge History of the English Language. Vol. 5 English in Britain and Overseas: Origin and Development*. Cambridge: Cambridge University Press, pp. 197-270.
- Jefferson, Gail (1990). List-construction as a task and resource. In George Psathas (ed.) *Interaction Competence*. Lanham, Md.: University Press of America, pp. 63-92.
- Johnstone, Barbara (2001). Discourse analysis and narrative. In Deborah Schiffrin, Deborah Tannen and Heidi Hamilton (eds.), pp. 635-49.
- Jucker, Andreas (2002). Discourse markers in Early Modern English. In Richard Watts and Peter Trudgill (eds.), pp. 210-30.
- Jucker, Andreas and Sara Smith (1998). *And people just you know like 'wow': Discourse markers as negotiating strategies*. In Andreas Jucker and Yael Ziv (eds.) *Discourse Markers: Descriptions and Theory*. Amsterdam: John Benjamins, pp. 171-201.
- Jucker, Andreas, Sara Smith, and Tanja Lüdge (2003). Interactive aspects of vagueness in conversation. *Journal of Pragmatics* 35: 1737-69.
- Kautzsch, Alexander (2002). *The Historical Evolution of Earlier African American English: An Empirical Comparison of Early Sources*. Berlin: Mouton de Gruyter.
- Keenan, Edward (1985). Relative clauses. In Timothy Shopen (ed.) *Language Typology and Syntactic Description, Vol. II: Complex Constructions*. Cambridge: Cambridge University Press, pp. 141-70.

- Keenan, Edward and Bernard Comrie (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63-99.
- Kennedy, Graeme (1998). *An Introduction to Corpus Linguistics*. London : Longman.
- Kernan, Keith (1977). Semantic and expressive elaboration in children's narratives. In Susan Ervin-Tripp and Carol Mitchell-Kernan (eds.) *Child Discourse*. New York: Academic Press, pp. 91-102.
- Kerswill, Paul (1996). Children, adolescents and language change. *Language Variation and Change* 8: 177-202.
- Kerswill, Paul (2002). Models of linguistic change and diffusion: New evidence from dialect levelling in British English. *Reading Working Papers in Linguistics* 6: 187-216.
- Kerswill, Paul and Jenny Cheshire (2004-2007). Linguistic innovators: The English of adolescents in London. ESRC funded project (RES000230680).
- Klein, Wolfgang (1992). The present perfect puzzle. *Language* 68: 525-52.
- Klemola, Juhani (1996). Non-standard periphrastic *do*: A study of variation and change. PhD thesis, University of Essex.
- Kortmann, Bernd (2006). Syntactic variation in English: A global perspective. In Bas Aarts and April McMahon (eds.) *The Handbook of English Linguistics*. Oxford: Blackwell, pp. 603-24.
- Kortmann, Bernd, and Benedikt Szmrecsanyi (2004). Global synopsis: Morphological and syntactic variation in English. In Bernd Kortmann, Edgar Schneider, Kate Burridge, Rajend Mesthrie, and Clive Upton (eds.), pp. 1142-1202.



- Kortmann, Bernd, Edgar Schneider, Kate Burridge, Rajend Mesthrie, and Clive Upton (eds.) (2004). *A Handbook of Varieties of English. Vol 2: Morphology and Syntax*. Berlin: Mouton de Gruyter.
- Kortmann, Bernd, Tanja Herrmann, Lukas Pietsch, and Susanne Wagner (eds.) (2005). *A Comparative Grammar of British English Dialects: Agreement, Gender and Relative Clauses*. Berlin: Mouton de Gruyter.
- Kortmann, Bernd and Susanne Wagner (2005). The Freiburg English dialect project and corpus (FRED). In Bernd Kortmann, Tanja Herrmann, Lukas Pietsch and Susanne Wagner (eds.), pp. 1-20.
- Kotsinas, Ulla-Britt (1994). *Ungdomsspråk*. Uppsala: Hallgren & Fallgren.
- Kovac, Cecil and Douglas Adamson (1981). Variation theory and first language acquisition. In David Sankoff and Henrietta Cedergren (eds.) *Variation Omnibus* Edmonton, Alberta: Linguistic Research, Inc., pp. 185-92.
- Krug, Manfred (2000). *Emerging English Modals: A Corpus-Based Study of Grammaticalization*. Berlin: Mouton de Gruyter.
- Kurath, Hans and Sherman Kuhn (eds.) (1954). *Middle English Dictionary*. Ann Arbor, Michigan: University of Michigan Press.
- Kyrtziz, Amy and Susan Ervin-Tripp (1999). The development of discourse markers in peer interaction. *Journal of Pragmatics* 31: 1321-38.
- Kytö Merja and Suzanne Romaine (1997). Competing forms of adjective comparison in modern English: What could be more quicker and easier and more effective? In Terttu Nevalainen and Leena Kahlas-Tarkka (eds.) *To Explain the Present: Studies in the Changing English Language*. Helsinki: Mémoires de la Société Néophilologique de Helsinki, pp. 329-52.

- Labov, William (1966). *The Social Stratification of English in New York City*.  
Washington, DC: Center for Applied Linguistics.
- Labov, William (1970). The study of language in its social context. *Studium Generale* 23: 30-87.
- Labov, William (1972). *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.
- Labov, William (1980a). Introduction. In William Labov (ed.) *Locating Language in Time and Space*. New York: Academic Press, pp. xiii-xx.
- Labov, William (1982). Building on empirical foundations. In Winfred Lehmann, and Yakov Malkiel (eds.) *Perspectives on Historical Linguistics*. Amsterdam: John Benamins, pp. 17-92.
- Labov, William (1984a). Field methods of the project on linguistic change and variation. In John Baugh and Joel Scherzer (eds.) *Language in Use: Readings in Sociolinguistics*. Englewood Cliffs, N.J.: Prentice-Hall, pp. 28-54.
- Labov, William (1984b). Intensity. In Deborah Schiffrin (ed.) *Meaning, Form and Use in Context: Linguistic Applications*. Washington, D.C.: Georgetown University Press, pp. 43-70.
- Labov, William (1989). The child as linguistic historian. *Language Variation and Change* 1: 85-97.
- Labov, William (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2: 205-54.
- Labov, William (2001). *Principles of Linguistic Change Volume 2: Social Factors*. Oxford: Blackwell.

- Labov, William and Jack Waletzky (1967). Narrative analysis. In June Helm (ed.) *Essays on the Verbal and Visual Arts*. Seattle: University of Washington Press, pp. 12-44.
- Labov, William, Paul Cohen, Clarence Robins, and John Lewis (1968). *A Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City*. Vol. I : *Phonological and Grammatical Analysis*. (Co-operative Research Project No. 3288 ) New York City: Columbia University.
- Lass, Roger (1993). *Old English: A Historical Linguistic Companion*. Cambridge: Cambridge University Press.
- Lauwereyns, Shizuka (2002). Hedges in Japanese conversation: The influence of age, sex and formality. *Language Variation and Change* 14: 239-59.
- Lawrence, Helen (2001). Aspects of English: An examination of aspect within past temporal reference in northern British English. PhD thesis. University of York, York, UK.
- Leith, Dick (1995). Tense variation as a performance feature in a Scottish folktale. *Language in Society* 24: 53-77.
- Leitner, Gerhard (1991). The Kohlapur Corpus of Indian English - intravarietal description and/or intervarietal comparison. In Stig Johansson and Anna-Brita Stenström (eds.) *English Computer Corpora: Selected Papers and Research Guide*: New York: Mouton de Gruyter, pp. 215-32.
- Levey, Stephen (2003). Reported dialogue and pragmatic particles in the narratives of preadolescents. *World Englishes* 22: 305-21.
- Levey, Stephen (2006). Visiting London relatives. *English World-Wide* 27: 45-70.
- Levin, Magnus (2006). Collective nouns and language change. *English Language and Linguistics* 10: 321-43.

- Lindstedt, Jouko (2000). The perfect-aspectual, temporal and evidential. In Östen Dahl (ed.) *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter, pp. 365-83.
- Llamas, Carmen (2007). Age. In Carmen Llamas, Louise Mullany and Peter Stockwell (eds.) *The Routledge Companion to Sociolinguistics*. London : Routledge, pp. 69-76.
- Local, John (1983). How many vowels in a vowel? *Journal of Child Language* 10: 449-53.
- Long, Mary McDonald (1944). *The English Strong Verb from Chaucer to Caxton*. Menasha, Wis.: George Banta.
- Lucy, John (1993). Metapragmatic presentationals: Reporting speech with quotatives in Yucatec Maya. In John Lucy (ed.), pp. 91-125.
- Lucy, John (ed.) (1993). *Reflexive Language: Reported Speech and Metapragmatics*. Cambridge: Cambridge University Press.
- Macaulay, Ronald (1985). The narrative skills of a Scottish coal miner. In Manfred Görlach (ed.) *Focus on: Scotland*. Amsterdam: John Benjamins, pp. 101-24.
- Macaulay, Ronald (1991). *Locating Dialect in Discourse*. Oxford: Oxford University Press.
- Macaulay, Ronald (2001). *You're like Why not?* The quotative expressions of Glasgow adolescents. *Journal of Sociolinguistics* 5: 3-21.
- Macaulay, Ronald (2002). Discourse variation. In Jack Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), pp. 283-305.
- Macaulay, Ronald (2005). *Talk that Counts: Age, Gender and Social Class Differences in Discourse*. Oxford: Oxford University Press.

- MacWhinney, Brian (1991). *The CHILDES Project: Tools for Analyzing Talk*. Hillsdale, NJ.: Erlbaum.
- Mair, Christian, and Marianne Hundt (1995). Why is the progressive becoming more frequent in English? *Zeitschrift für Anglistik und Amerikanistik* 2: 111-22.
- Martinez Insua, Ana E., and Ignacio M. Palacios Martinez (2003). A corpus-based approach to non-concord in present day existential *there* constructions. *English Studies* 84: 262-83.
- Maschler, Yael (2001). Veke'ilu haraglayim sh'xa nitka'ot bifnim kaze (*and like your feet get stuck inside like*): Hebrew kaze (*like*), ke'ilu (*like*), and the decline of Israeli dugri (direct) speech. *Discourse Studies* 3: 295-326.
- Maschler, Yael (2002). On the grammaticalisation of ke'ilu '*like*' in Hebrew talk-in-interaction. *Language in Society* 31: 243-76.
- Mathis, Terrie, and George Yule (1994). Zero quotatives. *Discourse Processes* 18: 63-76.
- Matthews, William (1972). *Cockney Past and Present*. London : Routledge and Keegan Paul.
- Maybin, Janet (1999). Framing and evaluation in ten to twelve year old school children's use of repeated, appropriated, and reported speech in relation to their induction into educational procedures and practices. *Text* 19: 459-84.
- McCafferty, Kevin (2004). [T]hunder storms is verry dangese in this countrey they come in less than a minnits notice...: The Northern Subject Rule in Southern Irish English. *English World-Wide* 25: 51-79.
- McCoard, Robert (1978). *The English Perfect: Tense Choice and Pragmatic Inferences*. Amsterdam: North-Holland.

- McDaniel, Dana, Cecile McKee and Judy Bernstein (1998). How children's relatives solve a problem for minimalism. *Language* 74: 308-34.
- McMahon, April (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Meechan, Margaret and Michele Foley (1994). On resolving disagreement: Linguistic theory and variation – *There's bridges*. *Language Variation and Change* 6: 63-85.
- Meehan, Teresa (1991). It's like, 'What's happening in the evolution of like?' *Kansas Working Papers in Linguistics* 16: 37-51.
- Mendoza-Denton, Norma (2002). Language and identity. In Jack Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), pp. 475-99.
- Meyerhoff, Miriam (1992). A sort of something – hedging strategies on nouns. *Working Papers in Language, Gender and Sexism* 3. *AILA Commission on Language and Gender* 2: 59-73.
- Meyerhoff, Miriam (1994). Sounds pretty ethnic, eh?: A pragmatic particle in New Zealand English. *Language in Society* 23: 367-88.
- Meyerhoff, Miriam (2002). All the same? The emergence of complementizers in Bislama. In Tom Güldemann and Manfred von Roncador (eds.) *Reported Discourse: A Meeting Ground for Different Linguistic Domains*. Amsterdam : John Benjamins, pp. 341-59.
- Meyerhoff, Miriam (2006). *Introducing Sociolinguistics*. London: Routledge.
- Meyerhoff, Miriam and Nancy Niedzielski (1998). The syntax and semantics of *olsem* in Bislama. In Matthew Pearson (ed.) *Recent Papers in Austronesian linguistics*. (UCLA Occasional Papers in Linguistics 20): 235-43.

- Meyerhoff, Miriam and Nancy Niedzielski (2003). The globalization of vernacular variation. *Journal of Sociolinguistics* 7: 534-55.
- Miège, Georges (1688). *The English Grammar*. Menston, UK: Scolar Press.
- Miller, Jim and Weinert Regina (1995). The function of *LIKE* in dialogue. *Journal of Pragmatics* 23: 365-393.
- Miller, Jim and Regina Weinert (1998). *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Clarendon Press.
- Milroy, James (1992). *Linguistic Variation and Change*. Oxford: Blackwell.
- Milroy, James and Lesley Milroy (1977). Speech and context in an urban setting. *Belfast Working Papers in Language and Linguistics* 2: 1-85.
- Milroy, James and Lesley Milroy (1991). *Authority in Language* 2<sup>nd</sup> edn. London: Routledge.
- Milroy, James and Lesley Milroy (eds.) (1993). *Real English: The Grammar of English Dialects in the British Isles*. London: Longman.
- Milroy, James and Lesley Milroy (1997). Varieties and variation. In Florian Coulmas (ed), pp. 47-64.
- Milroy, Lesley (1987). *Observing and Analysing Natural Language*. Oxford: Blackwell.
- Milroy, Lesley (2002). Introduction: Mobility, contact and language change- Working with contemporary speech communities. *Journal of Sociolinguistics* 6: 3-15.
- Milroy, Lesley and Gordon, Matthew (2003). *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.

- Mitchell-Kernan, Claudia (1972). Signifying and marking: Two Afro-American speech acts. In John Gumperz and Dell Hymes (eds.) *Directions in Sociolinguistics*. New York: Rinehart and Winston, pp. 161-79.
- Montgomery, Michael (2001). Transatlantic connections for variable grammatical features. *University of Pennsylvania Working Papers in Linguistics* 7: 205-23.
- Moore, Emma (2004). Sociolinguistic style: A multidimensional resource for shared identity creation. *Canadian Journal of Linguistics* 49: 375-96.
- Mossé, Fernand (1952). *A Handbook of Middle English*. Baltimore, MA : John Hopkins University Press.
- Mufwene, Salikoko (1986). Restrictive relativization in Gullah. *Journal of Pidgin and Creole Languages* 1: 1-31.
- Murray, James A. H. (1873). *The Dialect of the Southern Counties of Scotland: Its Grammar and Historical Relations*. London: Philological Society.
- Mustanoja, Tauno (1960). *A Middle English Syntax*. Helsinki : Société Néophilologique.
- Nesselhauf, Nadja (2007). The spread of the progressive and its 'future' use. *English Language and Linguistics* 11: 191-207.
- Nevalainen, Terttu (2006). Vernacular universals? The case of plural *was* in Early Modern English. In Terttu Nevalainen, Juhani Klemola, and Mikko Laitinen (eds.) *Types of Variation: Diachronic, Dialectal and Typological Interfaces*. Amsterdam: John Benjamins, pp. 351-69.
- Nevalainen, Terttu and Helena Raumolin-Brunberg (2002). The rise of the relative *who* in Early Modern English. In Patricia Poussa (ed.), pp. 109-21.



- Nevalainen, Terttu and Helena Raumolin-Brunberg (2003). *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. London: Longman.
- Nordberg, Bengt (1987). The use of onomatopoeia in the conversational style of adolescents. In P. Lillius & M. Saari (eds.) *The Nordic Languages and Modern Linguistics. Proceedings of the Sixth International Conference on Nordic and General Linguistics in Helsinki*. Helsinki: Helsinki University Press, pp. 256-88.
- Norrby, Catrin (2002). Svenska påhängsuttryck av typen *å så* och *eller nåt*. En diskussion av deras förekomst och funktion(er) i ett samtida ungdomsmaterial. *Språk & Stil* 11: 183-210.
- Norrby, Catrin and Joanna Winter (2002). Affiliation in adolescents' use of discourse extenders. In Cynthia Allen (ed.) *Proceedings of the 2001 Conference of the Australian Linguistics Society*. <<http://linguistics.anu.edu.au/ALS2001/proceedings.html>>.
- Norrby, Catrin and Karolina Wirdenäs (2003). Swedish youth discourse: On performing relevant selves in interaction. In Jannis Androutsopoulos and Alexandra Georgakopoulou (eds.) *Discourse Constructions of Youth Identities*. Amsterdam: John Benjamins, pp. 247-78.
- Orton, Harold, Michael Barry, Wilfrid Halliday, Peter Tilling and Martyn Wakelin. (1963-1971). *Survey of English Dialects*. Vols. 1-4. Leeds: Edward Arnold.
- Orton, Harold, Stewart Sanderson, and John Widdowson (eds.) (1978). *The Linguistic Atlas of England*. London: Croom Helm.
- Östman, Jan-Ola (1981). *You know: A Discourse Functional View*. Amsterdam: John Benjamins.

- Östman, Jan-Ola. (1995). Pragmatic particles twenty years after. In Brita Wårvik, Sanna-Kaisa Tanskanen, and Risto Hiltunen. (eds.) *Organization in Discourse. Proceedings from the Turku Conference*. Anglicana Turkuensia, 14. Turku: University of Turku, pp. 95-108.
- Overstreet, Maryann (1999). *Whales, Candlelight, and stuff like that – General Extenders in English Discourse*. Oxford: Oxford University Press.
- Overstreet, Maryann (2005). And stuff *und so*: Investigating pragmatic expressions in English and German. *Journal of Pragmatics* 37: 1845-64.
- Overstreet, Maryann and George Yule (1997). On being inexplicit and stuff in contemporary American English. *Journal of English Linguistics* 25: 250-8.
- Overstreet, Maryann and George Yule (2002). The metapragmatics of *and everything*. *Journal of Pragmatics* 34: 785-94.
- Paolillo, John (2002). *Analyzing Linguistic Variation: Statistical Models and Methods*. Stanford, CA: CSLI Publications.
- Pawley, Andrew and Frances Hodgetts Syder (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics* 7: 551-79.
- Payne, Arvilla (1980). Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In William Labov (ed.) *Locating Language in Time and Space*. New York: Academic Press, pp. 143-78.
- Peitsara, Kirsti (2002). Relativizers in the Suffolk dialect. In Patricia Poussa (ed.), pp. 167-80.
- Petyt, Malcolm (1985). *Dialect and Accent in Industrial West Yorkshire*. Amsterdam: John Benjamins.

- Pietsch, Lukas (2005a). Some do and some doesn't: Verbal concord variation in the north of the British isles. In Bernd Kortmann, Tanja Herrmann, Lukas Pietsch and Susanne Wagner (eds.), pp. 125-209.
- Pietsch, Lukas (2005b). *Variable Grammars: Verbal Agreement in Northern Dialects of English*. Tübingen: Niemeyer.
- Pollock, Jean-Yves (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20: 365-424.
- Poplack, Shana (1989). The care and handling of a megacorpus: The Ottawa-Hull French project. In Ralph Fasold and Deborah Schiffrin (eds.) *Language Change and Variation*. Amsterdam: John Benjamins, pp. 411-44.
- Poplack, Shana (ed.) (2000). *The English History of African American English*. Oxford: Blackwell.
- Poplack, Shana (2006). How English became African American English. In Ans van Kemenade and Bettelou Los (eds.), pp. 452-75.
- Poplack, Shana and Sali Tagliamonte (1991). African American English in the diaspora: Evidence from old-line Nova Scotians. *Language Variation and Change* 3: 301-39.
- Poplack, Shana, and Danielle Turpin. (1999). Does the *Futur* have a future in (Canadian) French? *Probus* 11 : 133-64.
- Poplack, Shana, and Sali Tagliamonte (1999). The grammaticalization of *going to* in (African American) English. *Language Variation and Change* 11: 315-42.
- Poplack, Shana, and Sali Tagliamonte (2001). *African American English in the Diaspora*. Oxford: Blackwell.

- Poplack, Shana, Gerard van Herk and Dawn Harvie (2002). Deformed in the dialects: An alternative history of non-standard English. In Richard Watts and Peter Trudgill (eds.), pp. 87-110.
- Poussa, Patricia (1988). The relative *what* : Two kinds of evidence. In Jacek Fisiak (ed.) *Historical Dialectology*. Berlin: Mouton, pp. 443-74.
- Poussa, Patricia (1994). Norfolk relatives (Broadland). In Wolfgang Viereck (ed.) *1990 Verhandlung des Internationalen Dialektologen Kongresses Bamberg. ZDL Beiheft 76*. Stuttgart: Franz Steiner, pp. 418-26.
- Poussa, Patricia (ed.) (2002a). *Relativisation on the North Sea Littoral*. Munich: Lincom Europa.
- Poussa, Patricia (2002b). North Sea relatives: Introduction. In Patricia Poussa (ed.), pp. 1-23.
- Preston, Dennis (1991). Sorting out the variables in sociolinguistic theory. *American Speech* 66: 33-56.
- Prince, Ellen (1981). Toward a taxonomy of given-new information. In Peter Cole (ed.) *Radical Pragmatics*. New York: Academic Press, pp. 223-54.
- Przedlacka, Joanna (2001). Estuary English and RP: Some recent findings. *Studia Anglica Posnaniensia* 36: 35-50.
- Pyles, Thomas and John Algeo (1993). *The Origins and Development of the English Language* 4<sup>th</sup> edn. Orlando: Harcourt Brace.
- Quirk, Randolph (1957). Relative clauses in educated spoken English. *English Studies* 38: 97-109.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartik (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

- Radford, Andrew (1988). *Transformational Grammar: A First Course*.  
Cambridge: Cambridge University Press.
- Radford, Andrew (1997). *Syntactic Theory and the Structure of English*.  
Cambridge: Cambridge University Press.
- Rampton, Ben (1995). *Crossing: Language and Ethnicity among Adolescents*.  
London: Longman.
- Ravid, Dorit and Ruth Berman (2006). Information density in the development of  
spoken and written narratives in English and Hebrew. *Discourse Processes*  
41: 117-49.
- Reid, Euan (1978). Social and stylistic variation in the speech of children: Some  
evidence from Edinburgh. In Peter Trudgill (ed.), pp. 158-72.
- Richardson, Peter (1991). Tense, discourse, style: The historical present in Sir  
Gawain and the Green Knight. *Neuphilologische Mitteilungen* 92: 343-49.
- Rickford, John (1997). Unequal partnership: Sociolinguistics and the African  
American speech community. *Language in Society* 26: 161-97.
- Rickford, John (2006). Down for the count? The Creole Origins Hypothesis of  
AAVE at the hands of the Ottawa Circle, and their supporters. *Journal  
of Pidgin and Creole Languages* 21: 97-155.
- Rickford, John, Norma Mendoza-Denton, Thomas Wasow and Julie Espinoza (1995).  
Syntactic variation and change in progress: Loss of the verbal coda in topic-  
restricting *as far as* constructions. *Language* 71: 102-31.
- Rickford, John, and Christine Théberge-Rafal (1999). Preterite *had* + verb-*ed* in the  
narratives of African American preadolescents. In John R. Rickford (ed.) *African  
American Vernacular English*. Oxford: Blackwell, pp. 34-60.

- Rickford, John, Isabelle Buchstaller, Thomas Wasow, and Arnold Zwicky (2005).  
Intensive and quotative *all*: Something old, something new. Paper presented at  
NWAV 34, New York University.
- Rizzi, Luigi (1991). Residual verb second and the *wh*-criterion. *Technical Reports in  
Formal and Computational Linguistics 2*, University of Geneva.
- Roberts, Julie (1994). *Acquisition of Variable Rules: (-t,d) Deletion and (ing)  
Production in Preschool Children*. Institute for Research in Cognitive Science  
(IRCS), Report 96-09. Philadelphia: University of Pennsylvania.
- Roberts, Julie (1997). Hitting a moving target: Acquisition of sound changes by  
Philadelphia children. *Language Variation and Change 9*: 249-66.
- Roberts, Julie (2002). Child language variation. In Jack Chambers, Peter Trudgill  
and Natalie Schilling-Estes (eds.), pp. 333-48.
- Roberts, Julie (2005). Acquisition of sociolinguistic variation. In Martin Ball (ed.),  
pp. 153-64.
- Roberts, Julie and William Labov (1995). Learning to talk Philadelphian: Acquisition  
of short *a* by preschool children. *Language Variation and Change 7*: 101-112.
- Romaine, Suzanne (1975). Linguistic variability in the speech of some Edinburgh  
schoolchildren. M.Litt. thesis. University of Edinburgh.
- Romaine, Suzanne (1978). Post-vocalic /r/ in Scottish English: Sound change in  
progress. In Peter Trudgill (ed.), pp. 144-57.
- Romaine, Suzanne (1980). The relative clause marker in Scots English: Diffusion,  
complexity and style as dimensions of syntactic change. *Language in Society 9*:  
221-49.
- Romaine, Suzanne (1982). *Socio-historical Linguistics: Its Status and Methodology*.  
Cambridge: Cambridge University Press.

- Romaine, Suzanne (1984a). *The Language of Children and Adolescents: The Acquisition of Communicative Competence*. Oxford: Blackwell.
- Romaine, Suzanne (1984b). On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Linguistica* 18: 409-37.
- Romaine, Suzanne (1984c). Towards a typology of relative clause formation strategies in Germanic. In Jacek Fisiak (ed.) *Historical Syntax*. Berlin: Mouton, pp. 437-70.
- Romaine, Suzanne (1984d). Review of Nessa Wolfson *The Conversational Historical Present in American English*. *Language in Society* 13: 117-23.
- Romaine, Suzanne (1985). Grammar and style in children's narratives. *Linguistics* 23: 83-104.
- Romaine, Suzanne (1988). *Pidgin and Creole Languages*. London: Longman.
- Romaine, Suzanne (1992). *Language, Education and Development: Urban and Rural Tok Pisin in Papua New Guinea*. Oxford: Oxford University Press.
- Romaine, Suzanne (1999). *Communicating Gender*. New York: Lawrence Erlbaum Associates.
- Romaine, Suzanne (2000). *Language in Society*. 2<sup>nd</sup> edn. Oxford: OUP.
- Romaine, Suzanne and Deborah Lange (1991). The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66: 227-79.
- Rosewarne, David (1994). Estuary English: Tomorrow's RP? *English Today* 37: 3-8.
- Ross, Sarah, Janna Oetting, and Beth Stapleton (2004). Preterite *had* + V-ed: A developmental narrative structure of African American English. *American Speech* 79: 167-93.

- Roth-Gordon, Jennifer (2001). Slang and the struggle over meaning: Race, language and power in Brazil. PhD dissertation, Department of Cultural and Social Anthropology, Stanford University.
- Rupp, Laura (2004). The scope of the Northern Subject Rule: Grammatical conditioning of verbal –s. Paper presented at the 39<sup>th</sup> Linguistics Colloquium, Amsterdam, 25-27 August, 2004.
- Rupp, Laura (2005). Constraints on non-standard –s in expletive *there* sentences: A generative-variationist perspective. *English Language and Linguistics* 9: 255-88.
- Ruzaite, Jurate (2004). Academic precision reconsidered: A corpus-based account. *SKY Journal of Linguistics* 17: 217-47.
- Sag, Ivan (1997). English relative clause constructions. *Journal of Linguistics* 33: 431-83.
- Sampson, Geoffrey (2002). Regional variation in the English verb qualifier system. *English Language and Linguistics* 6: 17-30.
- Sankoff, David (1988). Sociolinguistics and syntactic variation. In Frederick J. Newmeyer (ed.) *Linguistics: The Cambridge Survey Vol. IV Language: The Socio-cultural Context*. Cambridge: Cambridge University Press, pp. 140-61.
- Sankoff, David and Pierrette Thibault (1981). Weak complementarity. Tense and aspect in Montreal French. In B. Johns and D. Strang (eds.) *Syntactic Change*. Ann Arbor: University of Michigan Press, pp. 205-15.
- Sankoff, Gillian (1973). Above and beyond phonology in variable rules. In Charles-James N. Bailey and Roger W. Shuy (eds.) *New Ways of Analyzing Variation in English*. Washington, D.C.: Georgetown University Press, pp. 44-61.
- Sankoff, Gillian (1990). The grammaticalization of tense and aspect in Tok Pisin and Sranan. *Language Variation and Change* 2: 295-312.



- Sankoff, Gillian, and Penelope Brown (1976). The origins of discourse in syntax: A case study of Tok Pisin relatives. *Language* 52: 631-66.
- Sankoff, Gillian, Pierrette Thibault, Naomi Nagy, H el ene Blondeau, Marie-Odile Fonollosa, and Lucie Gagnon (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9: 191- 217.
- Sapir, Edward (1921). *Language. An Introduction to the Study of Speech*. San Diego: Harcourt Brace Jovanovich.
- Schendl, Herbert (1996). The third person plural present indicative in Early Modern English: Variation and language contact in English historical linguistics. In Derek Britton (ed.) *Papers from the 8<sup>th</sup> International Conference on English Historical Linguistics*. Amsterdam: Benjamins, pp. 143-60.
- Scherre, Maria Marta Pereira and Anthony J. Naro (1991). Marking in discourse: 'Birds of a feather.' *Language Variation and Change* 3: 23-32.
- Schiffrin, Deborah (1981). Tense variation in narrative. *Language* 57: 45-62.
- Schiffrin, Deborah (1987). *Discourse Markers*. Cambridge: Cambridge University Press.
- Schiffrin, Deborah (1994). *Approaches to Discourse*. Oxford: Blackwell.
- Schiffrin, Deborah (1996). Narrative as self-portrait: Sociolinguistic constructions of identity. *Language in Society* 25: 167-203.
- Schiffrin, Deborah (2001). Discourse markers: Language, meaning, and context. In Deborah Schiffrin, Deborah Tannen and Heidi Hamilton (eds.), pp. 54-75.
- Schiffrin, Deborah (2006). *In Other Words: Variation in Reference and Narrative*. Cambridge: Cambridge University Press.
- Schiffrin, Deborah, Deborah Tannen and Heidi Hamilton (eds.) (2001). *The Handbook of Discourse Analysis*. Oxford: Blackwell.

- Schilling-Estes, Natalie (1998). Investigating 'self-conscious' speech: The performance register in Ocracoke English. *Language in Society* 27: 53-83.
- Schilling-Estes, Natalie and Walt Wolfram (1994). Convergent explanation and alternative regularization: *Were/weren't* leveling in a vernacular English variety. *Language Variation and Change* 6: 273-302.
- Schourup, Lawrence (1985). *Common Discourse Particles in English Conversation: Like, Well, Y'know*. New York: Garland.
- Schourup, Lawrence (1999). Discourse markers. *Lingua* 107: 227-65.
- Schreier, Daniel (2002). Past *BE* in Tristan da Cunha: The rise and fall of categoricity in language change. *American Speech* 77: 70-99.
- Schütze, Carson T. (1999). English expletive constructions are not infected. *Linguistic Inquiry* 30: 467-84.
- Schwenter, Scott (1994). Hot news and the grammaticalization of perfects. *Linguistics* 32: 995-1028.
- Seppänen, Aimo (1999). Dialectal variation in English relativization. *Lingua* 109: 15-34.
- Sharma, Devyani (2001). The pluperfect in native and non-native English: A comparative corpus study. *Language Variation and Change* 13: 343-73.
- Shinmura, Izuru (ed.) (1998). *Kojien Dictionary* 5<sup>th</sup> edn. Tokyo: Iwamura Shoten.
- Shorrocks, Graham (1999). *A Grammar of the Dialect of the Bolton Area*. Vol II: *Morphology and Syntax*. Berlin: Lang.
- Siegel, Muffy (2002). *Like*: The discourse particle and semantics. *Journal of Semantics* 19: 35-71.

- Sigley, Robert (1997a). Choosing your relatives: Relative clauses in New Zealand English. PhD dissertation, Victoria University of Wellington.
- Sigley, Robert (1997b). The influence of formality and channel on relative pronoun choice in New Zealand English. *English Language and Linguistics* 1: 207-32.
- Sigley, Robert (2003). The importance of interaction effects. *Language Variation and Change* 15: 227-53.
- Silva-Corvalán, Carmen (1983). Tense and aspect in an oral Spanish narrative: context and meaning. *Language* 59: 760-80.
- Singler, John Victor (2001). Why you can't do a VARBRUL study of quotatives, and what such a study can show us. *University of Pennsylvania Working Papers in Linguistics* 7: 257-78.
- Sivertsen, Eva (1960). *Cockney Phonology*. Oslo Studies in English, 8. Oslo: Oslo University Press.
- Smith, Jennifer (2000). Synchrony and diachrony in the evolution of English: Evidence from the far reaches of Scotland. D.Phil. dissertation, Department of Language and Linguistic Science, University of York, UK.
- Smith, Jennifer (2003). *Narrative Analysis Database*. University of York.
- Smith, Jennifer (2005). Caregiver, community and child in the acquisition of variable dialect features. Final report to the ESRC (RES000220447).
- Smith, Jennifer and Tagliamonte, Sali (1998). We were all thegither .... we was all thegither: *was* regularization in Buckie English. *World Englishes* 17: 105-26.
- Smith, Jennifer, Mercedes Durham and Liane Fortune (2007). Mam, my trousers is fain doon!: Community, caregiver and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change* 19: 63-99.

- Soares da Silva, Augusto (2006). The polysemy of discourse markers: The case of *pronto* in Portuguese. *Journal of Pragmatics* 38: 2188-2205.
- Sprott, Richard (1992). Children's use of discourse markers in disputes: Form-function relations and discourse in child language. *Discourse Processes* 15: 423-39.
- Stenström, Anna-Brita, Gisle Andseren and Ingrid K. Hasund (2002). *Trends in Teenage Talk*. Amsterdam: John Benjamins.
- Strang, Barbara M.H. (1970). *A History of English*. London: Methuen.
- Stubbe, Maria and Janet Holmes (1995). *You know, eh* and other exasperating expressions: An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language and Communication* 15 : 63-88.
- Suñer, Margarita (1998). Resumptive restrictive relatives: A cross-linguistic perspective. *Language* 74: 335-64.
- Tagliamonte, Sali (1996). Has it ever been 'perfect'? Uncovering the grammar of Early Black English. *York Papers in Linguistics* 17: 351-96.
- Tagliamonte, Sali (1998a) *Was/were* variation across the generations: View from the city of York. *Language Variation and Change* 10: 153-92.
- Tagliamonte, Sali (1998b). Roots of identity: Variation and grammaticisation in contemporary British English. Final report to ESRC (R000221842).
- Tagliamonte, Sali (1999). Modelling an emergent grammar: Past temporal reference in St Kitts Creole in the 1780s. In Philip Baker and Adrienne Bruyn (eds.) *St Kitts and the Atlantic Creoles*. London: University of Westminster Press, pp. 201-36.

- Tagliamonte, Sali (2000a). The grammaticalization of the present perfect in English. In Olga Fischer, Anette Rosenbach and Dieter Stein (eds.) *Pathways of Change: Grammaticalization in English*. Amsterdam: John Benjamins, pp. 329-54.
- Tagliamonte, Sali (2000b). English ...*and them*. Form and function in comparative perspective. *American Speech* 75: 405-8.
- Tagliamonte, Sali (2001). *Come/came* variation in English dialects. *American Speech* 76: 42-61.
- Tagliamonte, Sali (2002a). Comparative sociolinguistics. In Jack Chambers Peter Trudgill, and Natalie Schilling-Estes (eds.), pp. 729-63.
- Tagliamonte, Sali (2002b). Variation and change in the British relative marker system. In Patricia Poussa (ed.), pp. 147-65.
- Tagliamonte, Sali (2005). *So who? Like how? Just what?* Discourse markers in the conversations of English speaking youth. *Journal of Pragmatics* 37: 1896-1915.
- Tagliamonte, Sali (2006a). *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali (2006b). Historical change in synchronic perspective: The legacy of British dialects. In Ans van Kemenade and Bettelou Los (eds.), pp. 477-506.
- Tagliamonte, Sali (2006c). Sociolinguistics in the secular world: Variation, theory and observation. Paper presented at NWAV 35, Ohio State University, 9<sup>th</sup> November, 2006.
- Tagliamonte, Sali (*in press*). *So cool, right?* Canadian English entering the 21<sup>st</sup> century. In Peter Avery, Alex D'Arcy, Elaine Gold. and Keren Rice (eds.) *Canadian English in a Global Context*. Theme issue of *The Canadian Journal of Linguistics*.

- Tagliamonte, Sali (*to appear*). *Roots of English: Exploring the History of English Dialects*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali and Shana Poplack (1988). How Black English *past* got to the *present*. *Language in Society* 17: 513-33.
- Tagliamonte, Sali and Rachel Hudson (1999). *Be like et al.* beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics* 3: 147-72.
- Tagliamonte, Sali and Jennifer Smith (1999). Analogical leveling in Samaná English: The case of *was* and *were*. *Journal of English Linguistics* 27: 8-26.
- Tagliamonte, Sali and Jennifer Smith (2000). Old *was*, new ecology: Viewing English through the sociolinguistic filter. In Shana Poplack (ed.), pp. 141-61.
- Tagliamonte, Sali and Alex D'Arcy (2004). *He's like, she's like*: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8: 493-514.
- Tagliamonte Sali and Rosalind Temple (2005). New perspectives on an *ol'* variable: (t,d) in British English. *Language Variation and Change* 17: 281-302.
- Tagliamonte, Sali, Jennifer Smith and Helen Lawrence (2005). No taming the vernacular ! Insights from the relatives in northern Britain. *Language Variation and Change* 17: 75-112.
- Tagliamonte, Sali and Alexandra D'Arcy (2007). Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change* 19: 199-217.
- Tannen, Deborah (1982). Oral and literate strategies in spoken and written narratives. *Language* 58: 1-21.

- Tannen, Deborah (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In Florian Coulmas (ed.) *Direct and Indirect Speech*. Berlin: Mouton de Gruyter, pp. 311-32.
- Tannen, Deborah (1989). *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.
- Tidholm, Hans (1979). *The Dialect of Egton in North Yorkshire*. Göteborg: Bokmaskinen.
- Toolan, Michael (1988). *Narrative: A Critical Linguistic Introduction*. London: Routledge.
- Torgersen, Eivind, Paul Kerswill, and Susan Fox (2006). Ethnicity as a source of changes in the London vowel system. In Frans Hinskens (ed.) *Language Variation- European Perspectives. Selected papers from the Third International Conference on Language Variation (ICLaVE3)*, Amsterdam, June 2005. Amsterdam: Benjamins, pp. 249-63.
- Tottie, Gunnel (1997). Relatively speaking: Relative marker usage in the British National Corpus. In Terttu Nevalainen and L. Kahlas-Tarkka, (eds.) *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*. Helsinki: Mémoires de la Société Néophilologique de Helsinki 52, pp. 465-81.
- Tottie, Gunnel and Michel Rey (1997). Relativization strategies in Earlier African American Vernacular English. *Language Variation and Change* 9: 219-47.
- Tottie, Gunnel and Dawn Harvie (2000). It's all relative: Relativization strategies in Early African American English. In Shana Poplack (ed.), pp. 198-230.
- Traugott, Elizabeth (1972). *The History of English Syntax*. New York: Holt, Rinehart and Winston.

- Traugott, Elizabeth (1982). From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalisation. In Winfred Lehmann and Yakov Malkiel (eds.) *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins, pp. 245-71.
- Traugott, Elizabeth (1992). Old English Syntax. In Richard Hogg (ed.) *The Cambridge History of the English Language*. Vol 1., pp. 168-289.
- Traugott, Elizabeth (1995). The role of the development of discourse markers in a theory of grammaticalisation. Paper presented at the 12<sup>th</sup> International Conference on Historical Linguistics, University of Manchester, UK.
- Traugott, Elizabeth and Richard Dasher (2002) *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
- Trudgill, Peter (1974). *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Trudgill, Peter (ed.) (1978). *Sociolinguistic Patterns in British English*. London: Arnold.
- Trudgill, Peter (1984). Standard English in England. In Peter Trudgill (ed.) *Language in the British Isles*. Cambridge: Cambridge University Press, pp. 32-44.
- Trudgill, Peter (1990). *The Dialects of England*. Oxford: Blackwell.
- Trudgill, Peter (2002). *Sociolinguistic Variation and Change*. Edinburgh: Edinburgh University Press.
- Trudgill, Peter (2004). *The Inevitability of Colonial Englishes*. Edinburgh: Edinburgh University Press.
- Underhill, Robert (1988). *Like is like focus*. *American Speech* 63: 234-46.



- Van den Eynden Morpeth, Nadine (2002). Relativizers in the southwest of England. In Patricia Poussa (ed.), pp. 181-94.
- Van der Auwera, Johan (1985). Relative *that* – a centennial dispute. *Journal of Linguistics* 21: 149-79.
- Van Herk, Gerard and Shana Poplack (2003). Rewriting the past: Bare verbs in the Ottawa Repository of Early African American Correspondence. *Journal of Pidgin and Creole Languages*. 18: 231-66.
- Van Kemenade, Ans and Bettelou Los (eds.) (2006). *The Handbook of the History of English*. Oxford: Blackwell.
- Vincent, Diane and David Sankoff (1992). Punctors: A pragmatic variable. *Language Variation and Change* 4: 206-16.
- Virdee, Dev and Tricia Williams (2003). *Focus on London*. Norwich: HMSO
- Visser, Fredericus (1963). *An Historical Syntax of the English Language*. Leiden: Brill.
- Waksler, Rachelle (2001). A new *ALL* in conversation. *American Speech* 76: 128-38.
- Walker, James (2007). *There's bears back there*: Plural existentials and vernacular universals in (Quebec) English. *English World-Wide* 28: 147-66.
- Walker, Jim (2007). The narrative perfect in contemporary British English- Oddity or emergent form? Poster presented at the 4<sup>th</sup> International Conference on Language Variation in Europe (ICLaVE), University of Cyprus, 19<sup>th</sup> June, 2007.
- Ward, Gregory and Betty Birner (1993). The semantics and pragmatics of *and everything*. *Journal of Pragmatics* 19: 205-14.
- Watt, Dominic and Jennifer Smith (2005). Language change. In Martin Ball (ed.), pp. 101-19.

- Watts, Richard and Peter Trudgill (eds.) (2002). *Alternative Histories of English*. London: Routledge.
- Webster, Noah (1789/1967). *Dissertations on the English Language*. Menston, UK: Scholar Press.
- Weiner Judith and William Labov (1983). Constraints on the agentless passive. *Journal of Linguistics* 19: 29-58.
- Weinert, Regina (2004). Relative clauses in spoken English and German- their structure and function. *Linguistische Berichte* 197: 3-51.
- Weinreich, Uriel, William Labov and Marvin Herzog (1968). Empirical foundations for a theory of language change. In Winfred Lehmann and Yakov Malkiel (eds.) *Directions for Historical Linguistics*. Austin: University of Texas Press, pp. 95-188.
- Wells, John (1982). *Accents of English. Volume 2: The British Isles*. Cambridge: Cambridge University Press.
- Wennerstrom, Ann and Andrew Siegel (2003). Keeping the floor in multiparty conversations: Intonation, syntax and pause. *Discourse Processes* 36: 77-107.
- Wilson, John and Alison Henry (1998). Parameter setting within a socially realistic linguistics. *Language in Society* 27: 1-21.
- Winter, Joanne (2002). Discourse quotatives in Australian English: Adolescents' performing voices. *Australian Journal of Linguistics* 22: 5-21.
- Winter, Joanne and Catrin Norrby (2000). Set marking tags *and stuff*. In John Henderson (ed.) *Proceedings of the 1999 Conference of the Australian Linguistic Society*. <<http://www.arts.uwa.edu.au/LingWWW/als99/proceedings.html>>.

- Wolfram, Walt (1993). Identifying and interpreting variables. In Dennis Preston (ed.) *American Dialect Research*. Amsterdam: John Benjamins, pp. 193-221.
- Wolfram, Walt (2000). Issues in reconstructing earlier African American English. *World Englishes* 19: 39-58.
- Wolfram, Walt and Sellers, Jason (1999). Ethnolinguistic marking of past *BE* in Lumbee Vernacular English. *Journal of English Linguistics* 27: 94-114
- Wolfram, Walt and Natalie Schilling-Estes (2003). Language change in conservative dialects: The case of past tense *BE* in southern enclave communities. *American Speech* 78: 208 -17.
- Wolfson, Nessa (1979). The conversational historical present alternation. *Language* 55: 168-82.
- Wolfson, Nessa (1982). *CHP: The Conversational Historical Present in American English*. Dordrecht: Foris Publications.
- Wray, Alison (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University.
- Wright, Joseph (1902). *The English Dialect Dictionary*. London: Henry Frowde.
- Wright, Joseph (1905). *The English Dialect Dictionary*. Vol.5: *The English Dialect Grammar*. London: Henry Frowde.
- Wright, Laura (2002). Third person plural present tense markers. *American Speech* 77: 242-63.
- Young, Robert and Robert Bayley (1996). Varbrul analysis for second language acquisition research. In Robert Bayley and Dennis Preston (eds.) *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins, pp. 253-306.

Young, Robin (2004). Is this the most irritating paragraph you will ever read?

*The Times*, March 24<sup>th</sup>.

Yule, George (2000). Review of Jim Miller and Regina Weinert *Spontaneous*

*Spoken Language: Syntax and Discourse*. *Lingua* 110: 375-78.

Zimmermann, Rüdiger (1973). Verbal syntax and style in *Sir Gawain and the Green*

*Knight*. *English Studies* 54: 533-43.