Personal Pronouns in the Dialects of England

A Corpus Study of Grammatical Variation in Spontaneous Speech

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der Philologischen Fakultät
der Albert-Ludwigs-Universität
Freiburg im Breisgau

vorgelegt von
Nuria Hernández y Siebold
aus Kehl am Rhein

Sommersemester 2010

Ursprünglicher Titel:
Personal Pronouns in the Dialects of England –
A Corpus-Driven Study of Grammatical Variation in Spontaneous Speech

Erstgutachter: Prof. Dr. Bernd Kortmann

Zweitgutachter: Prof. Dr. Dr. h.c. Christian Mair

Drittgutachter: Prof. Dr. Guido Seiler

Vorsitzender des Promotionsausschusses
der Gemeinsamen Kommission der
Philologischen, Philosophischen und Wirtschafts-
und Verhaltenswissenschaftlichen Fakultät: Prof. Dr. Hans-Helmuth Gander

Datum der Disputation: 14. März 2011

Contents

List of tables vi
List of figures vii
List of maps viii
Abbreviations ix
Acknowledgements xi

I Background

1 Contents and definitions . . . 1
1.1 Approach and objectives . . . 1
1.2 Grammatical categorisation and terminology . . . 7
1.3 Structure of the study . . . 11

2 Spoken-language research . . . 13
2.1 Preserving data of speech . . . 13
2.2 Transcribing data of speech . . . 14

3 Pronoun variation in the history of English . . . 17
3.1 Some historical facts . . . 18
3.1.1 Old English (450–1150 AD) . . . 18
3.1.2 Middle English (1150–1470 AD) . . . 23
3.1.3 Early Modern English to Modern English (after 1470) . . . 26
3.2 Standard English and the ideology of Standard English . . . 30
3.2.1 Desire for (linguistic) order . . . 30
3.2.2 Appreciation of linguistic diversity . . . 33
3.3 The standard Modern English paradigm . . . 34

II Empirical study

4 Data and methodology . . . 39
4.1 The Freiburg Corpus of English Dialects (FRED) . . . 39
4.1.1 Corpus design . . . 39
4.1.2 Subcorpus size and area coverage . . . 42
4.1.3 Speakers . . . 45
4.1.4 Recordings and transcripts . . . 46
4.2 Data extraction, coding and statistics . . . 49
4.3 CD-ROM . . . 55

5 Preliminary remarks **57**
5.1 The functional versatility of S-, O- and self-forms in the corpus 57
5.2 Speaker distribution patterns 59
5.2.1 Rare features 60
5.2.2 Speaker-specific features 60
5.2.3 Generalised features 61

6 Number and person features **63**
6.1 Plural forms with singular reference 64
6.1.1 Singular *us* 64
6.1.2 Third person plural forms with singular referent NPs 65
6.1.3 Areal distribution 68
6.2 Dummy objects: why not *walk it?* 68
6.3 Existential *it* 71
6.4 Generic question tags – *innit?* 73
6.5 Summary 77

7 Gendered pronouns **79**
7.1 Gendered pronouns in the literature 80
7.2 Gendered pronouns in FRED 81
7.2.1 Quantitative distribution 82
7.2.2 Referent categories, topicality, and the speaker viewpoint 84
7.2.3 Graphic representation on the Animacy Hierarchy 91
7.2.4 Switches between gendered and non-gendered forms 92
7.2.5 In-text distribution: semantic priming and semantic differentiation . . 93
7.2.6 Syntactic positions 94
7.2.7 Gendered pronouns and speaker sex 95
7.2.8 Areal distribution 97
7.3 Summary 99

8 Pronoun exchange **101**
8.1 Traditional definition 101
8.2 Subject forms in object function 103
8.2.1 Object and prepositional complement function 104
8.2.2 For–to and ECM constructions 106
8.2.3 Areal distribution 107
8.3 Object forms in subject function 108
8.3.1 Subject and subject complement function 111
8.3.2 Coordination 114
8.3.3 Question tags 115
8.3.4 No-verb utterances(1) 116
8.3.5 Relative clauses and clefts 119
8.3.6 Context hierarchy 119

8.3.7 Areal distribution . . . 120
8.4 Personal pronouns in reflexive function . . . 122
8.5 Independent self-forms . . . 124
8.5.1 Subject function . . . 126
8.5.2 Third person cases . . . 128
8.5.3 Subject complement function . . . 129
8.5.4 No-verb utterances(2) . . . 130
8.5.5 Object function . . . 131
8.5.6 Prepositional complement function . . . 131
8.6 Summary and revised definition . . . 133

9 Case variation in prepositional phrases . . . 135
9.1 Subject and self-forms in pronominal PPs . . . 135
9.2 'He's as tall as me' – *as*, *like, than*: prepositions or conjunctions? . . . 136
9.3 Snake sentences . . . 141
9.3.1 Discourse perspective . . . 143
9.3.2 Self-directed vs. other-directed verbs . . . 144
9.3.3 Corpus results . . . 145
9.4 Summary . . . 147

10 Gerunds and participles . . . 149
10.1 Structural and cognitive distinctions . . . 149
10.2 Distribution of case forms in the corpus . . . 153
10.3 Underlying differences in meaning . . . 154
10.3.1 Acting vs. action . . . 157
10.3.2 Hypothetical vs. actual . . . 159
10.3.3 'O-form + V-*ing*' for a more vivid mode of expression . . . 161
10.4 Syntactic constraints . . . 162
10.5 Adverbial participles . . . 164
10.6 Summary . . . 167

11 Qualified pronouns . . . 169
11.1 Constructions with qualifier nouns and quantifiers . . . 169
11.1.1 Qualified vs. non-qualified NPs . . . 172
11.1.2 Synthetically vs. analytically qualified NPs . . . 176
11.2 Second person plurals . . . 179
11.3 'Oh deary me': delexicalised exclamations . . . 182
11.4 Summary . . . 184

12 Pleonastic pronouns . . . 187
12.1 Overt *you* imperatives . . . 188
12.1.1 Structural subtypes . . . 188
12.1.2 Distribution in the corpus . . . 191

12.2 Pleonastic reflexives after standard intransitive verbs . . . 192
12.2.1 Historical and typological perspective . . . 193
12.2.2 Pleonastic reflexives in FRED . . . 194
12.2.3 Possible explanations . . . 195
12.3 Benefactive objects . . . 199
12.3.1 Non-pleonastic occurrences . . . 200
12.3.2 Pleonastic occurrences . . . 201
12.4 Ethic datives and ethic PPs . . . 203
12.5 Disjunctive pronouns: emphasis, topicalisation, specification . . . 206
12.5.1 Disjunctive uses in FRED . . . 207
12.5.2 Preference for object case . . . 209
12.5.3 Disjunction and coordination – two distinct mechanisms . . . 211
12.5.4 Discourse functions . . . 212
12.5.5 Areal distribution . . . 215
12.6 Resumptive pronouns . . . 215
12.7 Summary . . . 218

13 Genitive variation . . . 221
13.1 The two genitives of English . . . 221
13.2 Quantitative distribution in the corpus, with focus on 3SGn cases . . . 225
13.3 Modifier–head relations . . . 234
13.3.1 Subjective vs. objective readings . . . 234
13.3.2 Relationality of the head noun . . . 236
13.4 Collocational preferences . . . 237
13.5 Summary . . . 242

14 Possessive function . . . 245
14.1 *S*-genitives: pronouns in possessive determiner function . . . 246
14.1.1 Possessive *we* and *us* . . . 246
14.1.2 Possessive *thee* . . . 250
14.2 *Of*-genitives: pronouns in possessive complement function . . . 251
14.3 Rare variants . . . 252
14.4 Summary . . . 254

15 Demonstrative function . . . 255
15.1 *They* and *them* demonstratives . . . 255
15.2 Syntactic environments . . . 257
15.3 Quantitative distribution in the corpus . . . 258
15.4 Areal distribution . . . 261
15.5 Applying collexeme analysis to dialect data . . . 262
15.5.1 Distinctive collexeme test . . . 262
15.5.2 Covarying collexeme test . . . 264
15.6 Summary . . . 266

III Synopsis and discussion

16 The major empirical results **267**
16.1 Areal distribution of non-standard pronouns 267
16.2 Frequencies 271
16.3 Case assignment 273

17 Practical and theoretical implications **277**
17.1 Prioratisation of morphosyntactic categories 277
17.2 Determinants of pronoun variation: a combined approach 281
17.3 Pronoun pragmatics 290

18 Conclusion and outlook **295**

Appendix A – Historical information 299
Appendix B – Coding information 301
Appendix C – Supplementary analysis data 303
Appendix D – Maps 308

References 311

Deutsche Zusammenfassung 335

List of tables

3.1 Standard English paradigm of personal pronouns and reflexives 34
4.1 Size and coverage of the FRED corpus and subcorpus 42
4.2 Speaker distribution and text production by birth decade 46
4.3 Speaker distribution and text production by speaker sex 46
4.4 Spelling variants of the different pronouns in the corpus. 50
6.1 3SGn question tags 74
7.1 Gendered pronoun frequencies: animate referents 83
7.2 Gendered pronoun frequencies: inanimate referents 83
7.3 Non-human references in text sample, breakdown by syntactic function 95
7.4 Gendered pronouns and speaker sex 96
8.1 S-form proportions (%) in object and prepositional complement function 105
8.2 S-form proportions (%) in for–to and ECM constructions 107
8.3 O-form proportions (%) in subject and subject complement function 111
8.4 Case form distribution in no-verb statements and questions 118
8.5 O-form reflexives: absolute and relative frequencies 123
8.6 Self-form proportions (%) in subject and object functions 126
9.1 Case form distribution after *as*, *like* and *than* (non-coordinated occurrences) . . . 140
10.1 Transformation tests for gerund vs. participle classification 152
10.2 Case form distribution in gerund-participles: POSS vs. non-POSS 154
10.3 Case form distribution in adverbial participles (subject function) 165
11.1 Qualified pronouns, breakdown by qualifier type (synthetic cases) 170
11.2 Case form distribution in qualified vs. non-qualified pronouns (synthetic) 174
11.3 Qualified pronouns: synthetic vs. analytic variants 177
12.1 Case form distribution in pleonastic reflexives 195
12.2 Disjunctives, 4 structural subtypes 208
12.3 Simplex vs. coordinated disjunctives, breakdown by case form 208
12.4 O-form subjects in disjunctive vs. non-disjunctive occurrences 211
12.5 Self-form subjects in disjunctive vs. non-disjunctive occurrences 212
13.1 Genitives with pronominal modifiers: personal pronoun forms 227
13.2 Genitives with pronominal modifiers: possessive determiner/ pronoun forms . . . 227
13.3 Distribution of *s*- vs. *of*-genitives in FRED 228
13.4 Genitive structures and modifier–head relations (subjectiveness) 234
13.5 Genitives with head nouns of the ‘friends and family’ category 239

13.6 Grammatical variation in genitives with head noun *name* 241
14.1 Distribution of 1PL forms in prenominal determiner position 247
15.1 Multiple distinctive collexeme analysis: lexical variants in DEM contexts 263
15.2 Measured attraction for different DEM variants and collexeme *days* 265
16.1 Catalogue of non-standard pronouns and their areal distribution in FRED 269
16.2 Breakdown of non-standard phenomena by relative frequency range 272
A1 The Indo-European personal pronoun paradigm after Schmidt (1978) 299
A2–A10 The English personal pronouns from Old to Modern English 299
B1 Dialect area codes 301
B2 County codes 301
B3 Syntactic function table with corpus examples 301
C1 Function matrix: S- and O-forms in the FRED corpus 303
C2 Function matrix: self-forms in the FRED corpus 305
C3 Top attracted collexemes of demonstrative determiner *they* 306
C4 Top attracted collexemes of demonstrative determiner *them* 306
C5 Top repulsed collexemes of demonstrative determiner *these* 306
C6 Top repulsed collexemes of demonstrative determiner *those* 306

List of figures

1.1 A unified account of linguistic variation (Anderwald and Kortmann 2002: 159) . . 3
1.2 Patterns of variation and ambiguity (adapted from Anttila and Fong 2004) . . . 10
2.1 Recordings and transcripts on the standard–non-standard continuum 14
4.1 FRED on the orality–literacy continuum (after Koch and Oesterreicher 1994) . . 40
4.2 Header and first lines of a FRED transcript (text LAN_010) 48
4.3 Spreadsheet with concordances, POS codes and extra-linguistic variables 53
5.1 Speaker distribution pattern of very rare features 62
5.2 Speaker distribution pattern of speaker-specific features 62
5.3 Speaker distribution pattern of generalised features 62
5.4 Speaker distribution pattern of a generalised feature with outliers 62
6.1 Transport For London “Why not walk it?” campaign 70
7.1 Gendered pronouns vs. *it* on the Animacy Hierarchy 92
7.2 Geographic continuum: gendered pronouns, ‘non-human animate’ category . . . 98

7.3 Geographic continuum: gendered pronouns, 'inanimate count' category 98
8.1 Areal distribution: S-form objects (*I*, *he*, *she*, *we*, *they*) 108
8.2 Areal distribution: S-form prepositional complements (*I*, *he*, *she*, *we*, *they*) . . . 108
8.3 Tree structure for untriggered reflexives in co-ordinated NPs 114
8.4 Areal distribution: O-form subject complements (*me*, *him*, *her*, *us*, *them*) 120
8.5 Areal distribution: O-form subjects (*me*, *him*, *her*, *us*, *them*) 122
8.6 Binary PE system, traditional definition 134
8.7 Ternary PE system based on FRED 134
9.1 Tree structure for subjects with understood predicates (Emonds 1986: 99). . . . 141
10.1 Tree structure 'Me singing the Marseillaise' (Abney 1987: 144) 155
10.2 Tree structure 'My singing the Marseillaise' (Abney 1987: 144) 155
12.1 Areal distribution: pleonastic reflexives 194
12.2 Areal distribution: disjunctive pronouns 215
13.1 Animacy as a determinant of genitive choice 229
13.2 Interacting forces: animacy and topicality in 3SGn genitives 230
13.3 The increasing use of *s*-genitives with inanimate modifiers 232
14.1 Areal distribution: prenominal possessive *us* 249
15.1 Areal distribution: *they*, *them*, *these*, *those* in DEM positions 261
17.1 Determinants of pronoun choice in English 282
17.2 Pragmatic pronoun usage 290
C1 Linear regression model of syntactic functions and non-standard occurrences. . . 307

List of maps

3.1 The Anglo-Saxon kingdoms c. AD 650 (Trudgill 1999[2]: 36) 19
3.2 Middle English dialect areas (Kristensson 1997: 658) 24
4.1 England's traditional dialect areas (Trudgill 1999[2]: 34) 43
4.2 The FRED counties (by Benedikt Szmrecsanyi) 44
14.1 'Our/ us eyes' (LAE M75) 250
D1 England's modern dialect areas (Trudgill 1999[2]: 65) 308
D2 Possible future dialect areas in England (Trudgill 1999[2]: 83) 308
D3 Pronoun exchange areas according to Trudgill (1999[2]: 96) 308
D4 'SHE is' (LAE M68) 308

D5 ‘(is) SHE (married)’ (LAE M69) 309
D6 ‘US’ (WM 178) 309
D7 ‘just WE TWO’ (WM 185) 309
D8 ‘We have OUR own’ (WM 116) 309
D9 Action nominal constructions (WALS map by Maria Koptjevskaja-Tamm) . . . 310
D10 Third person pronouns and demonstratives (WALS map by D.N.S. Bhat) . . . 310
D11 Gender distinctions in independent personal pronouns (WALS map by Siewierska) 310
D12 Politeness distinctions in pronouns (WALS map by Johannes Helmbrecht) . . . 310

Abbreviations

Following the Leipzig Glossing rules (http://www.eva.mpg.de/lingua/resources):

1/2/3	first/second/third person
SG/PL/DU	singular/plural/dual
m/f/n	masculine/feminine/neuter
NOM	nominative
DAT	dative
ACC	accusative
GEN	genitive
REFL	reflexive
POSS	possessive
DEM	demonstrative
REL	relative

Pronoun forms and functions:

O-form	object case form
S-form	subject case form
self-form	*myself*, *yourself*, *ourselves*, etc.
POS	part-of-speech codes/ function codes, see Table B3

Dialect areas:

SE	Southeast
SW	Southwest
Mid	Midlands
N	North

Historical stages of English:

OE	Old English
MidE	Middle English,
EModE	Early Modern English
ModE	Modern English

English varieties:

StE	Standard English
AmE	American English
AAVE	African American Vernacular English
BrE	British English
IrE	Irish English
ScE	Scottish English

Speaker distribution patterns (ch. 5.2):

SN	Speaker Number (number of speakers who used a phenomenon in the corpus)
SP	Speaker Proportion (proportion of speakers who used the phenomenon)
MCR	Mean Case Ratio (total occurrences divided by SN)

Linguistic surveys, atlases, corpora and dictionaries:

BNC	The British National Corpus
CLAE	Computer Developed Linguistic Atlas of England (cf. Viereck 1991)
LAE	Linguistic Atlas of England (cf. Orton et al. 1978)
LSV	Lowman-Survey (cf. Vierck 1975)
OED	Oxford English Dictionary (cf. Murray et al. 1989[2])
SAND	Syntactische Atlas van de Nederlandse Dialecten (Barbiers et al. 2005)
SED	Survey of English Dialects (cf. Orton 1962; Orton et al. 1962–1971)
WALS	World Atlas of Language Structures (cf. Haspelmath et al. 2005)
WM	Word Maps: A Dialect Atlas of England (cf. Upton et al. 1987)

trba.	transcript revised by the author

Acknowledgements

I would like to express my gratitude to my colleagues, family and friends who have made this work possible. First and foremost, I want to thank my supervisor Bernd Kortmann for his continued encouragement and support over the years. I also gratefully acknowledge the generous award of a two-year postgraduate scholarship by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg.

In Freiburg, I want to thank the first generation of FRED researchers at the English Department of Freiburg University, especially Susanne Wagner, Tanja Herrmann and Lukas Pietsch. Thank you to the members of a small reading circle, Tamsin Sanderson, Iman Laversuch, Nadja Nesselhauf, Michelle Miles, Bianca Kossmann and Daniela Kolbe, for our afternoon teas and many stimulating discussions. In particular, I want to thank Benedikt Szmrecsanyi for his valuable advice and suggestions, and Melitta Cocan for always being a welcoming point of contact.

Sincere thanks to Glenda Leung for agreeing to proofread the study at the very last minute, and to Raymond Hickey for reviewing this last final draft before publication.

I am gratefully indebted to the School of English at Leeds University: Clive Upton, Kate Wallace and Esther Asprey, for welcoming me during my research stay, as well as the staff of the Brotherton Library, especially Oliver Pickering from Special Collections, for the privilege to access the SED fieldworker notebooks.

My deepest personal gratitude goes to my family and friends. Thanks to all friends in Freiburg and Cambridge who have made these last few years so much fun despite all the work. Thank you to mum and dad for all their love and support, and to Jordi and Mayte for their late-night e-mails and some much needed distraction.

My greatest thank you goes to my partner Claus, for always being there.

Part I

Background

Chapter 1
Contents and definitions

Personal pronouns such as *I*, *me*, *myself, she, her, herself* are among the most widely discussed topics in present-day linguistics. The use of personal pronouns in everyday conversation is often marked by the transgression of pre-determined categorial boundaries and is frequently in contradiction to the grammar of Standard English. The present study is an empirical investigation of the functional diversity of personal pronouns in 20th century spoken British English based on the England component of the Freiburg Corpus of English Dialects (FRED). The main objective is to offer a comprehensive account of personal pronoun behaviour from a variation perspective, focusing on non-standard uses.

1.1 Approach and objectives

The approach used in this study is descriptive and geared towards spontaneous language data. It follows the general ideas of John Sinclair, corpus linguistics pioneer and founder of the COBUILD project,[1] who advocated the simple searching for word forms in authentic, non-annotated corpus data as the most objective method for approaching real language (cf. Sinclair 1991). The large size of the database (also one of Sinclair's requirement) allows for an analysis of both frequent and rare phenomena. Great importance is given to a transparent presentation of the results, which can be reproduced by using the primary data on the enclosed CD-ROM. In terms of theoretical background, the study is best described as framework-free (cf. Haspelmath 2010). It is, however, not anti-theoretical since the analyses include detailed discussions of different theoretical aspects.

Within corpus linguistics, the corpus-based approach is distinguished from the corpus-driven approach. While the former typically avails itself of corpus data "mainly to expand, test or exemplify theories" (Tognini-Bonelli 2001: 65), the orientation in corpus-driven linguistics is even more towards the data rather than a specific theory. Although the present study is not corpus-driven in the strict sense of the word (since it makes certain theoretical presuppositions about grammatical structure), it follows the bottom-up approach advocated in corpus-driven studies. The implications of this is that the data itself forms the starting

1 See http://www.mycobuild.com/about-collins-corpus.aspx.

point for this study and that the investigation embraces all non-standard pronoun phenomena found in the data. This way, a comprehe2nsive picture is obtained of the scope and degree of grammatical variation manifest in the dataset. Tognini-Bonelli has described the corpus-driven approach as follows:

> In a corpus-driven approach, the commitment of the linguist is to the integrity of the data as a whole, and descriptions aim to be comprehensive with respect to corpus evidence. The corpus, therefore, is seen as more than a repository of examples to back pre-existing theories or a probabilistic extension to an already well-defined system. The theoretical statements are fully consistent with, and reflect directly, the evidence provided in the corpus. [...] Examples are normally taken verbatim, in other words they are not adjusted in any way to fit the predefined categories of the analyst; recurrent patterns and frequency distributions are expected to form the basic evidence for linguistic categories; the absence of a pattern is considered potentially meaningful. (Tognini-Bonelli 2001: 84)

In compliance with this definition, the overarching aims of the present study are:

1. to investigate the full functional range of personal pronoun behaviour;
2. to identify dominant distributional tendencies and patterns of variation, and to translate these into concise linguistic hierarchies;
3. to identify and account for the most influential determinants of variation of each individual phenomenon as well as the pronominal paradigm as a whole.

All empirical observations in this study are complemented by quantitative and qualitative analyses, including the revision of linguistic principles and hierarchies proposed in the literature. Also included are typological parallels and historical precursors, in accordance with the unified approach to linguistic variation shown in Figure 1.1. Although the study is not typological in itself, it recognises the benefits of typology as an additional frame of reference. Similarly, reference will be made to historical evidence for the different phenomena under discussion (see 3.1 and the individual analyses). This additional information will not only help to position the empirical observations on the language time line, it will also help to distinguish traditional dialect phenomena from linguistic innovations.[2]

2 The benefit is mutual. Rot (1994), for example, mentions the importance of dialect studies for the

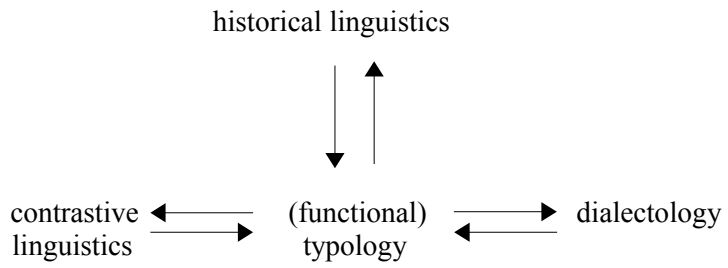


Figure 1.1: A unified account of linguistic variation (Anderwald and Kortmann 2002: 159)

Why focus on dialects and spontaneous speech?

Motivations for studying non-standard varieties vary from the recognition of societal differences in linguistic expression to practical applications in language teaching. Ever since the advent of dialectology as a linguistic discipline, the great importance of non-standard language data has been recognised for the investigation of 'real language': starting with Wenker's *Sprachatlas* in the late 19th century, Wright's *English Dialect Dictionary* in 1905, the *Linguistic Atlas of the United States* in the 1930s, the *Survey of English Dialects* in the 1950s, and the sociolinguistic studies of William Labov since the 1960s. Nowadays, variation in empirical data is no longer regarded as a problem for theoretical linguistics. Instead, the dialectological perspective is being integrated in major frameworks like generative linguistics, the leading school of thought in many academic institutions since the 1960s (e.g. Black and Motapanyane 1996).

Cheshire and Stein have pointed out the importance of dialect research and the investigation of spoken language in the following words:

> The syntax of the standard variety often seems to result from an explicit attempt by grammarians to tidy up the inherent fuzziness and indeterminacy of spoken syntax. This sometimes makes a structure more amenable to systematic description, but at the same time it becomes impossible to describe and explain the precise nature of that structure. (Cheshire and Stein 1997b: 6)

understanding of historical texts. According to Rot, words and grammatical structures which remain incomprehensible despite a detailed knowledge of both Standard English and earlier stages of English can be decoded with the help of similar words and structures still used in regional varieties.

A similar view is held by Van Marle (1997: 29), who states in the same volume that "standard languages are the product of scrutinizing by generations of grammarians with regard to the question of whether the constructions they contain are 'logical'." Standard varieties are usually oriented towards medial and conceptual literacy, demonstrating a heightened awareness of norms and 'correctness' which logically entails that "the purely systematic forces experience much more resistence from socio-cultural forces than in dialects" (Weiß 2004: 27). In other words, standard varieties are to a certain extent artificial whereas dialects represent the language in its natural state.

Why study morphosyntactic variation?

Studies on linguistic variation are traditionally concerned with phonological or lexical variation. The current study deliberately takes a different approach, emphasising the appeal of morphosyntactic variation.

Over the last decade, various major projects on dialect syntax in different European countries have opened new prospects for dialect studies and dialect-based cross-linguistic research. The FRED corpus used for this study, for example, resulted from an initiative of the European Science Foundation to investigate the morphology and syntax of regional varieties and spoken language. Other projects that started around the same time include the *Syntactic Atlas of Northern Italy* (ASIS), the *Syntactic Atlas of the Dutch Dialects* (SAND), the *Syntactic Atlas of Swiss German* (SADS), and the *Romani Morpho-Syntax Database and Dialect Survey* (RMS).[3] Nevertheless, the general focus in dialectology remains on phonological and lexical features whereas dialect grammar – from a syntactician's point of view – has only started to receive the attention it deserves.

Recent research has shown how important it is to widen the focus. New insights from dialect grammar have already challenged some of the classifications and definitions established on phonological and lexical grounds. Wagner (2004a: 184), for example, found that West Cornwall – which in non-syntactic studies is often excluded from the so-called West Country dialect area – exhibits the most pronounced use of the typical West Country phenomenon of gender diffusion, a gender system sensitive to mass–count distinctions. Findings from syntax and morphology can thus lead to a qualification of established isoglosses.

3 http://asis-cnr.unipd.it; http://meertens.nl/projecten/sand/sandeng.html; www.ds.unizh.ch/dialektsyntax/index.html; www.llc.manchester.ac.uk/Research/Projects/romani.

Why another pronoun study?

Variation in pronoun use has been a point of debate ever since Chomsky (1981) postulated Binding Principles A and B for the delimitation of pronouns against anaphors.[4] In the literature, many proposals have been made for the correct use of pronominal forms in specific syntactic contexts, and varied approaches have tried to explain the variation found in actual performance data, leading to an extensive exchange between formalist and functionalist linguistics. The ongoing discussion is an example for the sometimes tedious attempt to justify why language users behave the way they do and not the way they 'should'. In his cross-linguistic study on anaphora, Huang (2000: 22) concludes that the "distributional complementarity between anaphora and pronominals [...] seems to be a generative syntactician's fantasy world," thus giving credence to the view that structural principles alone cannot account for pronoun use in all its variety.

Research on pronoun variation has reached a complex state, making it all the more important to define what a specific approach and selection of data are able to contribute to the discussion. The present study situates itself within the current trend of contributions that aim at a descriptive, cross-dialectal perspective (e.g., Rohdenburg and Mondorf 2003; Kortmann et al. 2004). The study uses an empirically-founded holistic approach which allows for a detailed description of distributional tendencies in quantitative and qualitative terms. It brings together a variety of non-standard phenomena and the different frameworks in which they have so far been discussed: formalist and functionalist, generative, discourse-analytic, sociolinguistic, and pragmatic. This integrative approach allows to expand on those claims and hypotheses that are most compatible with the empirical observations in an unprepossessed way, and it allows to identify determinants of variation which are common to apparently unconnected phenomena.

An important incentive for the present investigation is also the fact that studies on pronoun variation often restrict themselves to smaller geographic areas (e.g., Trudgill 1974, 2004a; Ihalainen 1985, 1991; Wakelin 1986; Coupland 1988; Harris 1993; Beal 1993, 2004; Henry 1995; Pietsch 2005; but consider Szmrecsanyi's current work in English dialectometry, forthcoming). Regionally restricted studies naturally run the risk of

4 Principles which constrain relations of pronominal expressions to possible antecedents, based on government and binding relationships. Principle A states that "An anaphor is bound in its governing category."; Principle B states that "A pronominal is free in its governing category." (Chomsky 1981: 188). Note that 'anaphor' refers to reflexive and reciprocal pronouns only. Binding is defined as: α binds β iff α is in an argument position, α c-commands β, and α and β are coindexed; α c-commands β iff the first branching node dominating α dominates β, α does not dominate β and β does not dominate α. According to Principle A, anaphora must not only be bound, but must be bound in a local domain. In the simplest case, the anaphor and antecedent are clause-mates, as in $John_i$ *hurt* $himself_i$.

misinterpreting locally substantiated phenomena as regional phenomena if their presence in other varieties is ignored. The present study, therefore, takes a cross-dialectal perspective at pronoun use in the different parts of England: the Southeast, the Southwest, the Midlands and the North (see 4.1.2 for more details). Regionally restricted dialect features can thus be distinguished from supraregional features of spoken English.

Finally, the comprehensive approach used in this study not only facilitates the comparison of different phenomena, but also the comparison of different morphosyntactic categories such as person, number, gender and case, establishing the relative importance of each category for the correct processing of pronominal expressions (see 17.1).

A priori expectations and hypotheses

As mentioned above, a priori conceptualisations are avoided in this study. Nevertheless, some general expectations can be formulated based on previous studies and-well known characteristics of the English language. These can be used to guide us in the overall layout of the analysis of the corpus data in Part II. During the course of the study, these expectations will also be implicitly reconfirmed.

1. As a fairly fixed word order language, English can be expected to allow for a comparatively high degree of flexibility in the marking of syntactic relations, i.e. variation in case.
2. Morphosyntactic phenomena tend to be geographically widespread, necessitating a cross-regional approach to grammatical variation.
3. The traditional dialect data used in this study are not expected to exhibit linguistic innovations, nor are the detected phenomena expected to be language-specific.[5]
4. Based on previous studies, object case forms are expected to be functionally more versatile than subject case forms. A predominance of object case forms has been claimed for certain syntactic constructions described as default case environments; these environments will be monitored more closely.
5. Distributional tendencies of frequent phenomena can be translated into linguistic hierarchies, either in terms of morphological variants (hierarchies of the form 'S-form

5 The low-contact data used in this study are also expected to exhibit more variation than more formal data or written data, and they are expected to exhibit more variation than higher-contact L1 varieties (cf. Szmrecsanyi and Kortmann 2009). A comparison with these varieties, however, is beyond the scope of this study.

> O-form'), determinants of variation (e.g., 'simplex NP > coordinated NP', 'emphatic > non-emphatic', etc.), or geographic differences (e.g., South > North).

In Part II, we will specifically test a great variety of identifiable phenomena, such as pronoun exchange or gendered pronouns, against influences like geographical or syntactic factors. In addition, some more general hypotheses will also be put to the test:

1. Pronoun use in spontaneous conversation is non-categorical, counteracting the structuralist principle of 'one form – one meaning'.
2. Pronoun use is influenced by a number of different variation determinants. In addition, a certain degree of inherent variability is to be expected (speaker variability).
3. Individual phenomena may have different possible determinants, i.e. more than one explanation.

The overall results regarding these expectations will be summarised in chapters 16 and 17. Empirical examples which contradict hypotheses put forward in other studies will be handled with caution. Rather than invalidating a given hypothesis, counter-evidence should lead to revaluation and a closer investigation of statistical tendencies.[6] The absence of phenomena known from other studies is not interpreted as evidence of non-existence.[7]

It is sometimes said that man is a
categorizing animal... (Labov 1966: 20)

1.2 Grammatical categorisation and terminology

In modern Standard English, pronominal expressions have clearly delimited functions. Each function is encoded by a specific type of pronoun, generally with a telling name such as interrogative, possessive or demonstrative. Regarding the use of personal pronouns and reflexives, the standard paradigm is straightforward and unambiguous: subject functions are encoded by subject forms (*I*, *he*, etc.), object functions are encoded by object forms

6 In other words, a weak approach to falsifiability in the Popperian sense. For a discussion on what constitutes empirical evidence in linguistics, including Popper's principle of falsification, consider Penke and Rosenbach (2004).

7 This concerns, in particular, phenomena which are triggered by specific situational contexts. So-called picture noun phrases as in *This is myself...*, for example, can be expected to appear when people sit together looking at photographs.

(*me*, *him*, etc.), and reflexivity is encoded by self-forms (*myself*, *himself*, etc.). This neat classification is convenient and sometimes indispensable, for example in language teaching. However, it leaves no room for the variation observed in other varieties. Here, pronoun forms can be observed to cross the standard categorial boundaries. Interrogative forms, for example, appear in relative function (*the man what I saw*), possessive forms are replaced by definite articles (*I told the wife*), personal pronouns can have reflexive function (*I'll buy me a car*), and subject forms can have object function and vice versa (*her told me*; *I gave it to he*). Different technical terms from the literature could be used to describe these phenomena: 'transcategorization' (Ježek and Ramat 2009), 'refunctionalization' (Lass 1990), or 'functional reinterpretation' (Howe 1996).[8] In the present study, however, it seems more adequate to speak of functional variability, especially considering the fact that historical evidence exists for almost all phenomena found in the corpus (see 3.1 for details).

In order to measure and assess the correlation between formal and semantic properties of pronouns in the corpus, a strict terminological distinction will be maintained between pronoun form and pronoun function. Consequently, I define these terms as follows:

Pronoun form

Pronoun form refers to the word form irrespective of its syntactic function. The analyses focus on three different form types: S-forms are the subject forms *I*, *he*, *she*, *we*, *they*; O-forms are the object forms *me*, *him*, *her*, *us*, *them*; and self-forms are all pronouns ending in *-self* and *-selves*, i.e. *myself*, *yourself*, *himself*, *herself*, *itself*, *ourselves*, *yourselves*, *themselves*. In the following, analyses which compare different case forms only include pronouns with overt case distinctions, which is why *you* and *it* will be excluded most of the time.

It should be noted that definitions of 'personal pronoun' vary in the literature. Among the major reference grammars of English, some only include S- and O-forms (Greenbaum and Nelson 2002), others include possessive pronouns (Leech and Svartvik 1994; Quirk et al. 1985), or reflexive pronouns (Young 1984), or both (Huddleston and Pullum 2002). Based on the shared person (1, 2, 3), gender (m, f, n) and number features (SG, PL), subject, object, reflexive and possessive pronouns are often regarded as central pronouns (Quirk and Greenbaum 1984: 101; Quirk et al. 1985: 6.13), primary pronouns (Greenbaum

8 'Transcategorization' so far only includes the transgression of lexical items from one part-of-speech to another (e.g. verb to noun, as in *he likes to run > he went for a run*). 'Refunctionalization': contrary to Smith (2006), Lass (1990) describes refunctionalization as an essentially discontinuous and random process. 'Functional reinterpretation': compare Howe (1996: 176) on pronoun exchange and emphasis.

1996: 91) or as the core members of the personal pronoun category (Huddleston and Pullum 2002: 426). The four form types are also connected historically, possessives having developed from the genitive case of personal pronouns, and reflexives having developed by addition of *self*, later *-self*, to different personal and possessive pronoun forms. Bhat (2004: 13) supports the view of a continuum between personal pronouns and other pro-forms, with the 1SG personal pronoun *I* as the central prototypical form.

Furthermore, it should be noted that some studies distinguish between first and second person pronouns on the one hand, and third person pronouns on the other hand. Bhat (2004: 153), for example, distinguishes between personal pronouns, comprising first and second person S- and O-forms, and pro-forms, comprising third person S- and O-forms plus other pronouns. Elsewhere, pro-form is used as a cover term for all persons (e.g., Quirk et al. 1985). In the present study, no distinction will be made between first/ second and third person pronouns, but it is of course acknowledged that only the former denote speech act participant roles.

Pronoun function

Pronoun function denotes the syntactic function of a pronoun in a specific clause or sentence irrespective of its form. The focus in the analyses will be on three functions: subject, object and reflexive.

Subject function denotes the sentence subject, irrespective of the sentence type (statement or answer), the noun phrase structure (simplex or coordinated) and the presence or absence of the corresponding verb.

Object function denotes the sentence object or objects, both direct and indirect, irrespective of the sentence type, the noun phrase structure and the presence or absence of the corresponding verb. In several analysis chapters, object function and prepositional complement function will be discussed separately.

Reflexive function denotes pronouns which encode a concept of self-relatedness, or which redirect the action or effect described by the verb towards the agentive subject (cf. Wierzbicka 1996: 415). Reflexivity is seen as a special type of anaphoricity (cf. Huang 2000: 1). A broad definition of reflexivity is applied to prepositional phrases to include all cases where the pronoun is coreferential with the preceding subject and bears a beneficiary or recipient thematic relation. Note that non-reflexive self-forms are not automatically classified as emphatic. Adnominal and adverbial intensifiers do not form part of this study (see ch. 3.3).

Linguistic variables and linguistic variants

Following the sociolinguistic terminology of Labov (1966: 15), syntactic functions which allow for different surface realisations will be referred to as *linguistic variables*. Different pronoun forms occurring in the same function without a major change in meaning will be referred to as *linguistic variants*. The two types of variation considered in this study are visualised in Figure 1.2. In the following, *variation* will be used to refer to those cases where one meaning (M) can be encoded by different formal variants (F1, F2,...), whereas *ambiguity* will be used to refer to those cases where a specific pronoun form (F) carries several potential meanings (M1, M2, ...).

Figure 1.2: Patterns of variation and ambiguity (adapted from Anttila and Fong 2004)
M = meaning/syntactic function/variable, F = pronoun form/variant

Standard vs. non-standard

Occurrences which conform to the norms and regulations of Standard English will be referred to as *standard*. The standard paradigm in Table 3.1 will serve as the tertium comparationis in the analyses. It represents the prescriptive norms, in their strictest sense, for the correct use of personal pronouns in written Standard English. All occurrences which depart from this paradigm will be interpreted as *non-standard* (synonymously *dialectal*).

The terms standard and non-standard will be used for descriptive purposes and carry no qualitative connotations ('substandard' is not used due to its negative connotations). Neither of the two terms implies a judgement of grammatical choices. Also, non-standard is not to be confused with ungrammatical. The present study follows the view of Henry Sweet, who stated more than a century ago that "whatever is in general use in a language is for that very reason grammatically correct" (1900: 5).

Regarding the formality level of the investigated conversations, it should be noted that the interview situation may have induced speakers to use a variety closer to Standard

English than the one they would normally use privately. Consider the following extract from text WES_019, where the speaker was asked to produce some words typical of his own dialect:

> {<u IntSF> Are there particular dialect words which are peculiar, is it just Langdale you ’re talking about or this part of Langdale?} [...]
> <u WesZb> Well, it ’s Cumbria now, but there ’s quite a difference between the Westmorland dialect and the Cumberland dialect. There ’s lots of different words, where, for instance, a hole, well, in Cumberland, they ’d say a warl, and we would say a hoal which is quite different, isn’t it? Because I know Mary Birkett, she ’s broad Cumberland, and she says quite a lot of words different to what we say. But if you talk really broad, your broad Westmorland dialect, it ’s not good to understand, I should think, by town people.
> {<u IntSF> Can you tell me in broad Westmorland dialect, for example, just about collecting the sheep for shearing or dipping, just tell me something in broad Westmorland.} [...]
> <u WesZb> Oh well, you ’re gaan to gidder t’ sheep, that ’s what you ’d say. Gidder for gather.

The interview effect is difficult to determine. Nevertheless, passages like this one provide important clues that such an effect exists. The conversations under investigation position themselves somewhere between broad dialect and an informal spoken standard (cf. Van Marle 1997: 19).

1.3 Structure of the study

The present study is divided in three major parts: background (Part I), empirical study (Part II), and synopsis and discussion (Part III). In the remaining two sections of Part I, chapter 2 addresses the importance of spoken-language research, and chapter 3 provides some historical background information relevant for the subsequent analyses.

Part II forms the empirical core of the study. Chapter 4 gives an introduction to the data and methodology. Chapter 5 contains some preliminary observations on the overall versatility of different pronoun forms as well as the different speaker distribution patterns discernible in the corpus. The empirical analyses start with variation in number and person in chapter 6, followed by variation in gender in chapter 7. Variation in case takes centre stage in chapters 8 through 15. Individual sections investigate different non-standard phenomena, including pronoun exchange, case variation in gerunds and participles, so-called qualified pronouns which form part of a more complex noun phrase, pleonastic pronouns judged unnecessary in the standard variety, and the distribution of the two genitive structures in English. The last two empirical sections investigate the use of

personal pronoun forms in possessive and demonstrative functions. Each empirical section closes with a brief summary.

Part III, which covers chapters 16 through 18, completes the study with a synopsis and discussion of the major empirical results and an outlook towards possible areas of interest for future research in pronoun use. Reading back to front, the synopsis can also be used as an introduction.

Chapter 2
Spoken-language research

The next two sections draw attention to some issues related to the investigation of spontaneous speech. Section 2.1 addresses the importance of data preservation; section 2.2 adverts to potentially problematic issues in the transcription of speech with reference to the data at hand.

2.1 Preserving data of speech

Language corpora are often but a positive side-effect of the need for empirical foundation. However, collections like the FRED corpus also represent a means of language documentation. The growing interest in spoken language and the rising popularity of corpus linguistics over the last few decades have resulted in the collection of large amounts of spoken language data. While the primary incentive naturally lies in the investigation of specific linguistic phenomena, the linguistic community is currently in a new and special position. It has at its disposal the technology, not only to record large stretches of natural conversation, but to preserve what to future generations will be historical stages of spoken English. This consideration is all the more important in view of the progressive levelling of dialects (cf. Kerswill and Williams 1999)[9] and in view of the typical progression of language change from speech to writing (cf. Kortmann 2001).

Over the last century, England, like most other countries, has experienced a general increase in mobility accompanied by widespread migration from the rural areas to the urban centres. Even the most remote communities are nowadays in daily contact with Standard English via educational institutions and the media. Collections of spontaneous conversations such as the FRED corpus could lead the way to a systematic preservation of data of speech, i.e. a continuous documentation of the language and the way it is changing. With the available technology, both research-related and documentation-related exigencies can be met with a sufficiently high standard of elicitation techniques, quality of the recordings, and geographical and societal representativity. Speech can finally be preserved directly without the interference of writing.

9 Data preservation is also important in view of the continuous disappearance of languages world-wide. Compare the UNESCO Atlas of the World's Languages in Danger on http://www.unesco.org/culture/en/endangeredlanguages/atlas.

2.2 Transcribing data of speech

Researchers working with spoken-language transcripts are confronted with some very specific issues such as the representation of dialectal variants. Miethaner (2000: 534–535) has pointed out that "written representations of a given variety can acquire 'quasi-authoritative' status", and that written representations can "influence readers' perceptions of the linguistic status of a given variety." In order to interpret corpus findings correctly, the end-user should therefore be familiar with the corresponding transcription policy, especially any standardisation procedures applied during the transformation from original linguistic performance to machine-readable text.

In the current study, this concerns in particular the representation of dialectal variants such *mi* 'me'/ 'my', *thee*, *thou*, *'im*, *ourself*, and so forth, in the orthographic transcripts of the FRED corpus (see 4.2 for more details). Transcribers of spontaneous conversations inevitably have to deal with the question of whether and how to represent such variants without compromising the searchability and readability of the data. One solution is to include variants in the transcripts through in-line annotation (*there was myself [miself] and my [mi] sister*; partially implemented in FRED). Another is to hold available the corresponding soundfiles and to enable direct access through text and sound alignment (not yet implemented in FRED).

Regularisation and standardisation always entail the risk of losing potentially interesting information. By artificially raising the formality level of the transcript, these two procedures can open a perceived gap between the formality levels of the text and the original recording. Figure 2.1 visualises what happens if the written representation of a conversation is pushed towards the standard pole of the standard–non-standard continuum. In the worst case scenario, the two versions no longer represent the same variety, and investigations ignoring the formality gap provide inaccurate results.

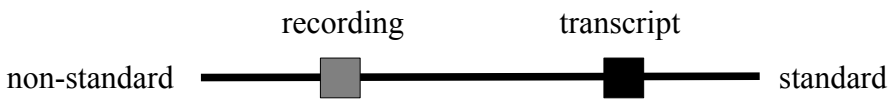


Figure 2.1: Recordings and transcripts on the standard–non-standard continuum

Another challenge in corpus linguistics is the fact that transcription methods vary widely (cf. Siren and Wilcox 1990; Jaffe 2000; Buchholtz 2000). For every collection of data, decisions concerning the text markup and orthography are made by researchers with their own personal interests and experiences. It, therefore, seems more useful not to "distribute praise and blame nor to offer a definite guide to 'correct transcription practices' but to make researchers aware of the complexities of the transcription process" (Buchholtz 2000: 1453). Some basic requirements that are necessary for the interpretation of all orthographic transcripts include (i) a basic understanding of the options which writing systems in general offer for the representation of speech; (ii) an awareness that characteristic properties of the standard orthography may influence the way that orthographic modifications are used;[10] and (iii) a distinction between scholarly and popular interpretations of 'respellings', specifically the use of so-called 'phonetic graphemes' (cf. Miethaner 2000). An ideological dimension may also filter into the transcript:

> The motivation to capture linguistic differences from the standard variety orthographically depends both on the transcriber's trust in the linguistic efficiency of modification strategies and on his or her ideas about the linguistic salience of specific features [...] This choice is especially relevant for the representation of features whose linguistic status is unclear or has attracted conflicting interpretations... (Miethaner 2000: 541)

Finally, there is always the human factor which needs to be kept in mind as soon as there is more than one transcriber. In the FRED corpus, for example, different students and members of staff were involved in the transcription of the data over several years, their language proficiency ranging from advanced (e.g., non-native with an M.A. in English linguistics) to native British, American or Canadian. The application of the predetermined guidelines varies to a certain extent. Dialectal pronoun forms, for example, were only standardised consistently in some texts. It is therefore advisable to take a look at the information available in the text headers as well as the corpus user guide (Hernández 2006; Szmrecsanyi and Hernández 2007).

10 For example, the general tendency in English to avoid homographs explains the decision to transcribe /mi/ 'my' as <mi> in FRED.

Chapter 3
Pronoun variation in the history of English

This chapter contains some basic information about the historical development of personal pronouns in English which is considered essential for a correct classification of the phenomena under investigation. This additional knowledge will help position the empirical observations on the language time line. It will also facilitate the distinction between traditional phenomena and more recent developments.

Edward Sapir once explained:

> The linguistic drift has direction. In other words, only those individual variations embody it or carry it which move in a certain direction, just as only certain wave movements in the bay outline the tide. [...] This direction may be inferred, in the main, from the past history of a language. (Sapir 1921: 165–166)

For the present study, a very interesting insight in this respect is that English personal pronouns have a long tradition of functional variability, and that this variability still appears to be growing. Almost all of the non-standard phenomena discussed in chapters 6 through 15 are found in historical stages of English. The fact that these phenomena continued to be used despite different regulatory forces and structural changes in the English language is in line with the generally accepted notion that (morpho-)syntactic features tend to stay stable over long periods of time.

The availability of historical evidence, of course, depends to a large extent on the acceptance of the respective phenomena in writing. Evidence may be found in early documents but disappear in the decades or centuries to follow, especially during the 17th and 18th century prescriptive period. Evidence may then reappear much later once a certain linguistic use is allowed back into writing (cf. Stein 1994). In the present study, any historical evidence is therefore considered proof of previous existence. The temporary absence of the same phenomenon from writing does not justify its classification as a linguistic innovation.

3.1 Some historical facts

This section serves as a reference chapter for the analyses presented in Part II. The goal is not a detailed description of earlier stages of English but a concise overview of the historical precursors found for the different non-standard phenomena discussed later on. The earliest historical records in this section date from the Old English period. A reconstruction of the Indo-European pronoun system as well as the personal pronoun paradigms from Old to Modern English are shown in Tables A1–A10 in Appendix A.

The historical developments outlined in this section illustrate that pronoun use in English has always shown a certain degree of functional variability. The persistent variation and, in later stages of the language, deviation from a prescriptive standard is in line with the fact that the acquisition of categorical assignment rules is "to a large extent a conscious process" (Tieken-Boon van Ostade 1994: 226). According to Tieken-Boon van Ostade, it is only "after continued exposure to standard adult usage and as the result of persistent correction that the grammatically correct forms begin to appear more regularly" (ibid.). Instances of variation and ambiguity can, therefore, not be dismissed as linguistic 'aberrations'. They form part of the language naturally.

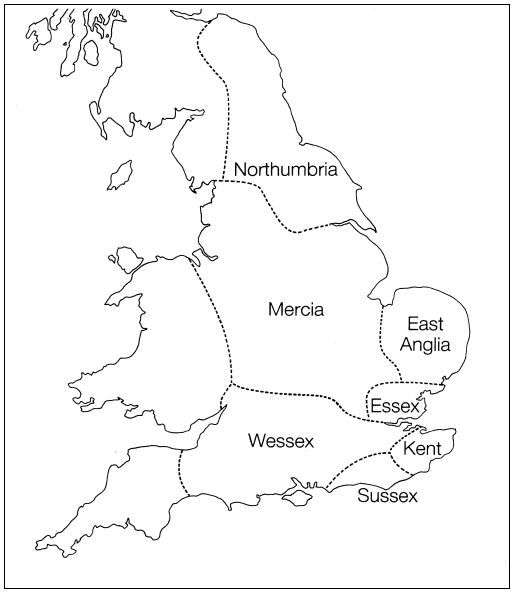
3.1.1 Old English (450–1150 AD)

The personal pronoun paradigm of Old English is shown in Tables A2–A10 in Appendix A. The main differences compared to Modern English consist in the following features: two additional dual forms *wit* 'we two' and *ġit* 'you two' (cf. Mitchell 1985: 110), overt number distinction between 2SG *þu* and 2PL *ġe* (ibid., p. 107), and a fourfold case distinction in NOM, GEN, ACC and DAT forms. In the third person, the same three genders are distinguished as in Modern English, but the system is one of grammatical gender (cf. Curzan 2003).

In Old English, pronouns can appear in different syntactic positions and still be assigned the correct function. A Modern English sentence like *My name is Maria*, for instance, can be translated into Old English as *Mīn nama is Maria*, or *Nama mīn is Maria*, or *Maria mīn nama is*. In addition, the Old English paradigm is characterised by a relatively large number of homographic case forms. We find identical ACC and DAT forms (1PL *ūs*, 2PL *ēow*, etc.), identical NOM and ACC forms (3SGn *hit*, 3PL *hīe*), and identical DAT and GEN forms (3SGf *hire*, compare ModE *her*). Two other ambiguous

forms are *him* (3SGm/ 3SGn/ 3PL DAT) and *hīe* (3SGf ACC, 3PL NOM/ ACC). In many Old English sentences, the correct assignment of morphosyntactic features must therefore be deduced from the context.

The regional varieties of Old English – Northumbrian, Mercian, Kentish and West Saxon – are associated with different independent Anglo-Saxon kingdoms, as shown in Map 3.1 (also cf. Campbell 1962: 4). After the unification of these kingdoms under Alfred the Great in 878, most Old English texts were written in the Wessex variety.



Map 3.1: The Anglo-Saxon kingdoms c. AD 650 (Trudgill 1999[2]: 36)

The following non-standard uses in FRED have historical precursors from this period:[11]

<u>O-form subjects (compare ch. 8.3)</u>

Example (1) shows two subjects, one DAT *him*, one NOM *he*; the pronoun *hine* could be ACC or REFL so that the second part of the sentence could be translated as '...and he threw himself to the ground and he heard voices' or '...and it threw him to the ground and he heard voices'. The contrary phenomenon, S-form objects, does not seem to appear in Old English texts.

11 Unified formatting was used in the examples.

(1) **him** com færlice to micel leoht, and **hine** astrehte to eorðan,
3SGm-DAT came suddenly into great light, and 3SGm-ACC/ 3SGm-REFL threw to earth,
and **he** gehyrde stemne
and 3SGm-NOM heard voices

(Ælfric *Catholic Homilies* i.386.6, quoted in Mitchell 1985)[12]

<u>O-form reflexives (compare ch. 8.4)</u>

In Old English, reflexivity is encoded by different means: simple O-forms (both ACC and DAT), absolute *self*, and complex 'personal pronoun + *self*' reflexives, as shown in (2)–(4).[13] Example (1) showed how the use of O-form reflexives can lead to ambiguity, in this case in the assignment of *hine* due to the lack of an overt subject in the second clause. While the use of *self* is not mandatory for a reflexive interpretation, its addition as a marker of coreference can help disambiguate the pronoun's function (note that Old English inflected *self* agrees with the corresponding pronoun in number, gender and case). This is illustrated in example (2): *selfne* functions as an anti-obviation marker which redirects the personal pronoun towards the central agent of a verb which would otherwise be interpreted as other-directed by default.

(2) Judas$_i$ gewræc hine selfne.
other-directed $j \neq$ **i** ← **i**

'Judas punished himself'

(adapted from König and Siemund 1997: 104)

The situation in Old English stands in contrast to the complementary use of personal pronouns and reflexives in standard Modern English. The historical evidence is, however, surprisingly similar to the situation in FRED, where we also find a considerable amount of O-form reflexives, and a continued absence of S-form equivalents. It may well be that present-day vernaculars are preserving an inherited constraint (cf. Kiparsky 2008: 32).

(3) swa hwa swa eadmedaþ **hine**
whoever humiliates 3SGm-ACC
'whoever humiliates himself'

(quoted in Faltz 1985)

12 *Catholic Homilies*: full text available on http://www.archive.org/details/sermonescatholic02aelfuoft.
13 It would be wrong to say that Old English had no reflexives, as is occasionally argued in the literature. Also compare Mitchell (1985: 113/ 121/ 192) for the use of genitive reflexives.

(4) Se Hælend geceas **him**, to eacan þam twulf apostolum,
the Saviour chose him-DAT, besides the twelve apostles,

twa and hundseofontig leorning-cnihta, and sende hi twam and twam ætforan
seventy-two disciples, and sent them by two and two before

him to ælc þæra byrig and stowe þe **he sylf** toweard wæs
him-DAT to each of those cities and places where he self-EMPH towards was

(Ælfric *Catholic Homilies* ii.528.25, quoted in Mitchell 1985)

Independent self-forms (compare ch. 8.5)

Highly interesting in the light of the present study are certainly those Old English occurrences which resemble the so-called independent self-forms in present-day English. These are cases where *self* or 'pronoun + *self*' is used instead of a personal pronoun subject, object or prepositional complement. The use of *self* in subject function is rarely mentioned in the literature, but consider the following examples:

(5) swa **þu** **self** talast
like you-NOM self-NOM say

(*Beowulf* 594, quoted in Mitchell 1985)[14]

(6) **selfe** ofersawon, ða ic of searwum cwom
'themselves saw, as I came from the battle'

(*Beowulf* 418, quoted in Mitchell 1985)

(7) þa wundrade Alexander hwy hit swa æmenne wære,
then wondered Alexander why it so deserted was

hrædlice þone weall **self** overclom
quickly that wall himself over-climbed

(King Alfred's *Orosius*, 134.12, quoted in Mitchell 1985)[15]

(8) Sittan læte ic hine wiþ **me** **sylfne**
remain let I him-ACC with me-ACC self-ACC
'I let him stay with me'

(*Junius*, *Genesis* 438, quoted in Van Gelderen 2000a)[16]

(9) Nu ic þæs tacen wege, sweotol on **me** **selfum**
now I the sign carry, clear on me-DAT self-DAT
'And now I clearly bear the sign upon me'

(*Junius Manuscript*, *Genesis* 885, quoted in Mitchell 1985)

14 *Beowulf*: full text available on http://www.archive.org/details/beowulf00981gut.
15 *Orosius*: full text available on http://ota.ahds.ac.uk/headers/0815.xml.
16 *Junius Manuscript, Genesis*: full text available on http://www.gutenberg.org/ebooks/618.

Pleonastic reflexives with inherently reflexive verbs (compare ch. 12.2)

Overt reflexives are commonly used with inherently reflexive verbs. Mitchell (1985: 114) refers to them as pleonastic reflexives.

(10) Simon to ðisum wordum **hine** gebealh
Simon at these words himself enraged
'Hearing these words, Simon got angry'

(Ælfric *Catholic Homilies* i.372.17, quoted in Mitchell 1985)

(11) Hwæt þa Egeas **him** ondred ða menigu
'Hereupon Ægeas feared the multitude'

(Ælfric *Catholic Homilies* i.596.35, quoted in Mitchell 1985)

Overt *you* imperatives (compare ch. 12.1):

Old English verbs are conjugated in three moods: indicative, subjunctive and imperative. Despite the use of distinct imperative verb forms, some sentences show overt second person pronouns as seen in example (12).

(12) Gang **ðe** nu on sibbe
go you now in peace

(Ælfric *Catholic Homilies* ii.394.22, quoted in Mitchell 1985)

Qualified pronouns (compare ch. 11)

Old English personal pronouns can be used in apposition with participles, adjectives and nouns. The construction with participles, which is still found in languages like German (*ich an Gott glaubender Mensch* 'I in God believing human'), is no longer found, but combinations with qualifier nouns are very common in spoken English today.

(13) Se ðe æfre is god, he brincð **us yfele** to godum mannum
he who ever is good, he brings us evil [people] to good men
'He who ever is good makes us who are evil be good men'

(Ælfric *Catholic Homilies* i.254.9, quoted in Mitchell 1985)

(14) An andwyrde for manegum, forðan ðe annys wæs on **him manegum**
one answered for many because the unity was in them many
'one answered for many, because unity was in the many'

(Ælfric *Catholic Homilies* ii.388.34, quoted in Mitchell 1985)

Missing gender distinctions (compare ch. 7 and 13.2)

The missing gender distinction in Old English *him* (DAT) and *his* (GEN), which are both masculine and neuter, provides a potential explanation for the use of so-called gendered pronouns in present-day informal English (cf. Lass 1999b: 148). In addition, Old English *he* 'he' and *hēo* 'she' occasionally refer to both male and female human referents (cf. Traugott 1992: 177–178). The absence of a distinctively neuter possessive before the 16th century (nowadays *its*) could explain the difference in distribution between 3SGn *its*/ *of it* and other pronominal modifiers in the two English genitive structures. This aspect will be discussed in more detail in chapter 13.2.

Demonstrative *þam* (compare ch. 15)

The widespread use of demonstrative *them* in Modern English possibly has its roots in the Old English 3PL definite article *þam* (DAT) and the Middle English homonymous 3PL demonstrative. The Old English example in (15) can be translated into Modern English as 'to the/ those heathen gods they obeyed', or, in vernacular style, 'to them heathen gods they obeyed'.

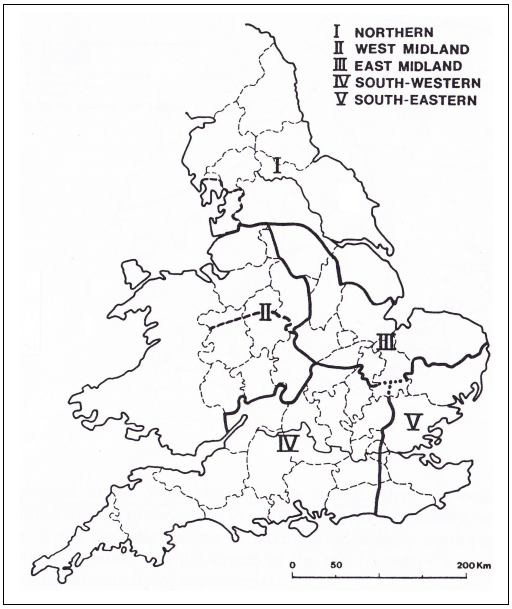
(15) to þam hæþenȝilde buȝon
to the-DAT heathen-gods-DAT obeyed-3PL
(*Numbers* 25,2, quoted in Van Gelderen 2000a)

3.1.2 Middle English (1150–1470 AD)

The Middle English period is characterised by a general simplification of the pronominal paradigm. The rising number of homonyms illustrates the language's "drift toward the invariable word" (Sapir 1921: 180). This involves the loss of dual forms (cf. Henderson 1985: 305) and the reduction of object case forms. The collapse of Old English ACC and DAT forms can be seen in Tables A2–A10. Genitives cease to function as true case forms, retaining only their possessive function (cf. Lass 1999b: 119). In the second person, the different pronouns are reduced to one form, *you*. The only differentiations which take place affect the third person. In the 3PL, the Old English homonymous NOM and ACC form is replaced by distinct case forms in *th-* (Old Norse; cf. Morse-Gagne 2003), and in the 3SGf,

OE *hēo* is gradually replaced by new forms in /ʃ/ (cf. Duncan 1972; Maps D4 and D5 in Appendix D).

The ambiguity resulting from the morphological reduction of the paradigm is partly absorbed by the grammaticalisation of word order (cf. Sapir 1921) and the development from inherent to structural case (cf. Van Gelderen 2000a). Word order starts to play a crucial role in the assignment of syntactic roles such as subject vs. object, direct vs. indirect object, object vs. reflexive, and reflexive vs. intensifier.



Map 3.2: Middle English dialect areas (Kristensson 1997: 658)

Regarding England's linguistic division during this period, Kristensson identifies 5 major dialect areas, as shown in Map 3.2 (based on spellings of place and personal names in lay subsidy rolls in the Survey of Middle English Dialects, 1290–1350). These dialect areas are surprisingly similar to the traditional dialect areas of Modern English shown in Map 4.1, demonstrating the general stability of dialect boundaries (cf. Wakelin 1983).

In Middle English texts we find more examples for the phenomena listed under Old English above. Complex reflexives of the form 'pronoun + *self*' start gaining ground but continue to coexist with simple reflexives. At the same time, *self* develops into a bound morpheme -*self*, losing its gender and case agreement features. According to König and

Siemund (1997: 101), 80% to 90% of all Middle English reflexives are still simple personal pronouns as seen in example (16):

(16) He that buckles **him** in my belt
(Shakespeare *Henry IV*, I, iii, quoted in Van Gelderen 2000a)[17]

We also continue to find overt reflexives with inherently reflexive verbs,

(17) I may **me** wel dres, for I haf sen a selly, I may not forsake.
(*Sir Gawain and the Green Knight* 474, quoted in Van Gelderen 2000a)[18]

(18) I dresse **my selfe** handsome (Shakespeare *Henry IV*, II, iv, quoted in Van Gelderen 2000a)

(19) He xuld repent **hym** (*Paston Letters* 129, quoted in Van Gelderen 2000a)[19]

and independent self-forms in subject and prepositional complement function:

(20) if **yowreself** lykez
(*Sir Gawain and the Green Knight* 1964, quoted in Van Gelderen 2000a)

(21) **Himself** drank water of the wel/ As dide the knight Sire Percivel
(Chaucer *Canterbury Tales*, *The Tale of Sir Thopas*, quoted in König and Siemund 1997)[20]

(22) Ah scupte him nome, after **him-seluan**
'But (Brutus) gave him a name after himself'
(Layamon *Brut* 977, Caligula manuscript, quoted in Van Gelderen 2000a)[21]

<u>Second person plurals (compare ch. 11.2)</u>

During Middle English, the original number distinction in the second person develops into a sociolinguistic distinction between familiar *thou*/ *thee* and polite *ye*/ *you*. Over time, the polite variants become the dominant form of address, and object *you* starts taking over the subject domain (cf. Barber 1976: 204).[22] Nowadays, the need for a 2SG–2PL distinction is noticeable in different varieties of English. In informal speech, this need is reflected in

17 All works by Shakespeare quoted in the current chapter are available on http://www.gutenberg.org.
18 *Sir Gawain and the Green Knight*: full text available on http://www.luminarium.org. Modern translation on the same site.
19 *Paston Letters*: full text available on http://www.luminarium.org.
20 *Canterbury Tales*: full text available on http://www.librarius.com.
21 Layamon's *Brut*: full text available on http://www.gutenberg.org/etext/14305.

different colloquial 2PL forms which compensate for the missing distinction (see ch. 11.2), including *youse* and qualified pronouns like *you all* or *you guys*. The reintroduction of a distinct 2PL form is among the most likely candidates for a future grammatical change in English (cf. Kortmann 2001; Wright 1997).

Singular *us* (compare ch. 6.1.1)

The use of *us* with singular reference is a well-known dialect feature of Modern English. In the *Oxford English Dictionary* (OED), examples of singular *us* date back as far as 1258. The first colloquial uses, for example in kind requests such as (24), seem to appear much later in the 19th century. Around the same time, singular *us* starts to be noticed in the dialectological literature (e.g., Lowsley 1888: 6; Wright 1905: 271).

(23) Tell our army from **us** that...
(Queen Elizabeth I in Moryson's *Itinerary* II, year 1601, quoted in OED online)[23]

(24) Tell **us** something more about the pea-shooting.
(Thomas Hughes *Tom Brown's School Days* I, iv, year 1857, quoted in OED online)[24]

3.1.3 Early Modern English to Modern English (after 1470)

It is during the Early Modern English period that an English linguistic standard evolves, a supraregional variety with clear morphological and syntactic constraints. Most changes in the pronominal paradigm are by now complete. The remodelling of the second person progresses towards completion, establishing *you* as the only second person form.[25] Among the few linguistic innovations after the year 1500 is the introduction of a new 3SGn possessive *its*.

22 The loss of social semantics in English agrees with the typological tendency that the presence of social deixis implies personal deixis, but not vice versa. According to Mühlhäusler and Harré (1990: 264), it is thus "not surprising that in the reduction of pronoun systems those signalling social deixis (e.g. honorifics or distancers) tend to be lost." However, the development in English does not support Mühlhäusler's and Harré's observation that "when this happens, those pronoun forms signalling respect are replaced typically by those that do not" (ibid.).
23 *Morison's Itinerary*: full text available on http://www.archive.org.
24 *Tom Brown's School Days*: full text available on http://www.gutenberg.org.
25 *Thee* and *thou* were still common in Shakespeare, but Ellis (1889) already observed that people found *thou* "too personal." Ihalainen (1994: 230) still reports occurrences of subject *thee* in the 1970s and 80s "in intimate style, when you talk to someone you are 'pally wi'."

The standardisation process bears down on the use of orthographic variants. In the standard variety, all pronouns except *you* and *it* are left with one S-form and one O-form (variation continues to exist in speech, e.g. h-dropped *'im* and *'er*, 1SG /mi/, 3SGn /hit/).[26] Grammatical variation is strongly curtailed, too, especially in the prescriptive literature of the 17th and 18th centuries. Overt subjects are now obligatory in most contexts, leading to the emergence of dummy *it*. Among the newer developments which are of interest to the present study are the use of S-forms in object function, the spreading of O-forms into additional syntactic functions, and the use of different types of pleonastic pronouns.

S-form objects and prepositional complements (compare ch. 8.2)

In addition to the continued use of O-form subjects as found in Old and Middle English, S-forms start appearing in object and prepositional complement function, but not reflexive function. In the dialectological studies of the 18th and 19th centuries, the bi-directional exchangeability of subject and object case forms is already referred to as 'pronoun exchange' (cf. Marshall 1789; Lowsley 1888; Barnes 1886; Ellis 1889; Halliwell 1887). The use of S-forms in object and prepositional complement function is attributed to emphasis (cf. Wright 1905: 271; Kruisinga 1905: 35–36). Interestingly, S-form objects appear to cluster in one dialectal area, the Southwest of England (compare the results in 8.2.3).[27] There appears to be a geographically restricted loosening of a constraint 'no S-form objects', which, in earlier stages, seemed to apply to English as a whole.

(25) Yes, you have seen Cassio and **she** together.
(Shakespeare *The Tragedy of Othello, the Moor of Venice* IV, ii; quoted in Vallins 1952)

(26) ...and since in paying it, it is impossible I should live, all debts are cleared between you and **I**, if I might but see you at my death.
(Shakespeare The Merchant of Venice III, ii, quoted in Bauer and Trudgill 1998)

Spreading of O-forms (compare chapters 8.3, 9.2, 12.5.2 and 16.3)

In addition to the continued use of O-form subjects, O-forms start spreading to other syntactic functions. This includes, for example, subject complements (probably after 1600, compare 8.3.1):

26 The Southwestern 1SG variants *ich* and *utch* are now obsolete (cf. Ellis 1889: 4/ 43/ 44/ 84).
27 Many early studies focus on the Southwest and Midlands, which makes it difficult to ascertain the geographic restriction.

(27) I am the dog: no, the dog is himself, and I am the dog –
oh, the dog **is me**, and I am myself; ay, so, so.

(Shakespeare *The Two Gentlemen of Verona* II.iii)

Variation in case is also noticed and criticised in disjunctive pronouns of the type *I told you, I* (Shakespeare *The Taming of the Shrew* III, ii, 12) which, according to Visser (1963: 56), were "used with striking frequency in the sixteenth and seventeenth centuries, especially in the dramatic works of Ben Jonson, Shakespeare and contemporaries." In bare elliptical responses of the *Not me* type, O-form disjunctives date back at least as far as 1848, taking over in the decades to follow (cf. Denison 1998: 108). Nowadays, the widespread use of O-form disjunctives, even in educated speech, is acknowledged in major descriptive grammars (e.g. Quirk et al. 1985: 6.4).

Another much debated question is whether pronouns should have object or subject case after comparative *as*, *like* and *than*, depending on whether these expressions are regarded as prepositions or conjunctions. This particular issue will be discussed in chapter 9.2.

(28) I shall always have a great regard for the Miss Martins, especially Elizabeth, and should be very sorry to give them up, for they are quite as well educated **as me**.

(Jane Austen *Emma*, I, iv, quoted in Visser 1963)[28]

Pleonastic pronouns (compare ch. 12)

Additional pronouns which do not change the fundamental proposition of the sentence are frequently referred to as pleonastic pronouns (compare the pleonastic reflexives described for Old English). It can be difficult to pinpoint the exact function of these pronouns, especially in early texts. They are often used to stress the emotional involvement or opinion of the speaker.

Among the pleonastic pronouns used in Early Modern English, we find overt *you* imperatives similar to those described for Old English. In Early Modern English, overt subjects have become mandatory everywhere except in imperatives. Nevertheless, overt pronouns frequently occur in positive and negative commands and requests. They are not restricted to second person examples, even if these are naturally the most frequent.

28 In Austen's novel, the young girl Harriet Smith, who has become the target of Emma's attempts at matchmaking, is reflecting on her prospective behaviour towards the man she adores and his sisters. It may well be that the O-form (*as me* 'as I am') was used by Austen to underline the sweet and docile, but "not clever", nature of the girl (Emma I, iv). Full text available on http://www.gutenberg.org.

Rissanen (1999) identifies the following patterns: 'V + pronoun' (as seen in sentences (29) and (30)), 'pronoun + V' (as in *You go*, common in Old English and early Middle English, reappears at the beginning of Modern English), 'do + pronoun + V' (as in *Do you tell us*, early Middle English until the 19th century; polite contrastive use), 'don't + pronoun + V' (as in *Don't you go away*). Interestingly, all of these patterns also appear in the FRED corpus, as will be shown in chapter 12.1.

(29) ...this staffe is my sister, for, **looke you**, she is
as white as a lilly, and as small as a wand...
(Shakespeare *The Two Gentlemen of Verona* II.iii)

(30) ...**retyre we** to our chamber, a little water clears us of this deed...
(Shakespeare *Macbeth* II.ii)

Pleonastic reflexives with intransitive or inherently reflexive verbs are known from Old and Middle English. They continue to be found in Early Modern English texts, mainly with verbs of motion. As their frequency declines, they seem to be more and more clearly associated with involvement, emphasis and a familiar style (cf. Rissanen 1999: 256).

(31) Good Margaret runne **thee** to the parlour...
(Shakespeare *Much Ado about Nothing* III,i, quoted in Rissanen 1999)

(32) O, but I fear **me** nothing can reclaim him!
(Marlowe *The Tragical History of Doctor Faustus*, quoted in Poutsma 1929)[29]

In drama, the so-called ethic dative becomes a popular stylistic device. It guarantees higher intensity of expression and a more vivid, colloquial flavour (compare 12.4; note the puns in (34)):

(33) a. I seeing that, tooke him by the leg, and never rested pulling, till I had
pul'd **me** his leg quite off.
(Marlowe *The Tragical History of Doctor Faustus* 1248, 1616 edition, quoted in Rissanen 1999)

b. He did; and with an absolute 'Sir, not I!' the cloudy messenger turnes **me** his back...
(Shakespeare *Macbeth* III.vi, quoted in Rissanen 1999)

29 Marlowe *Faustus*: full text available on http://www.gutenberg.org.

(34) Petruchio ...Heere sirra Grumio, knocke I say.
Grumio Knocke sir? whom should I knocke?
Petruchio Villaine I say, **knocke me** heere soundly.
Grumio **Knocke you** heere sir? Why sir, what am I sir,
that I should **knocke you** heere sir
Petruchio Villaine I say, **knocke me** at this gate,
And **rap me** well, or Ile knocke your knaues pate.
Grumio My master is grown quarrelsome:
I should **knocke you** first,
And then I know after who comes by the worst.
Petruchio Will it not be?
'Faith sirrah, and if you'l not knocke, Ile ring it,
Ile trie how you can Sol-Fa, and sing it.
He rings him by the eares.
(Shakespeare *The Taming of the Shrew* I.ii, 1590s, quoted in Rissanen 1999)

3.2 Standard English and the ideology of Standard English

Different theories have been proposed in the literature for the evolution of the variety nowadays known as Standard English (e.g., Wright 2000; Hope 2000; Riley and Parker 1998). Most of them describe the codification of a written standard which, more often than its oral counterpart, reflects 'changes from above', i.e. changes promoted by the dominant social class.[30]

3.2.1 Desire for (linguistic) order

It is generally agreed that Standard English as we know it today developed from some Midland varieties in connection with the large-scale migration of speakers from the English Midlands to London during the Middle Ages. The dominant theory is that the language in its written form was strongly influenced by the appearance of the first government documents in English during the time when English started to gain ground as an official language.

30 For a discussion of the naturalness model, according to which Standard English came into existence naturally, see Hope (2000). For the effects of literacy on grammar, see Olson (1994).

In the 14th century, the first spelling rules developed in the writing offices of King Henry V at Westminster, probably based on a central Midlands variety.[31] The language of these documents would later be known as the Chancery Standard, named after one of the offices. This standard soon developed into a local norm with supraregional prestige. One of the main objectives of the responsible clerks was to reduce orthographic variation in order to put English on the same level as the previous administrative languages Latin and French.

In the century to follow, the diffusion of the Chancery Standard was strongly accelerated by the introduction of the printing press, which was set up at Westminster in 1476. The new large-scale reproduction of texts required further codification. The fixation of English orthography in official documents based on the printers' standard at the press was practically complete around 1650. The spelling partly followed phonemic principles, partly logographic principles in order to keep homophones apart (e.g., *flower* vs. *flour*; cf. Nevalainen 2006: 32).

In the 17th century, the English people experienced various political changes and disasters. The Civil War of 1642–1649 and the decapitation of King Charles I were followed by a new Commonwealth under Cromwell, which was then followed by the Restoration under Charles II. The constant upheavals caused a general sense of insecurity and instability resulting in an increased desire for order, regulation, sameness, control and correctness, in both the political and social life (cf. Riley and Parker 1998: 20–21). This atmosphere soon affected the language in use, bringing with it a further reduction of linguistic variation. At the same time, a novel vernacular nationalism emerged among philologists, their interests shifting from classical Latin and Greek to the national languages of western Europe. The erudite public witnessed the appearance of the first textbooks for the reading and writing of English, and in 1662 the foundation followed of a Royal Society which only a few years later established a committee for the improvement of the language. Among the committee's members were well-known poets like John Dryden, who, for example, corrected a number of pronoun 'mistakes' in Shakespeare (cf. Tieken-Boon van Ostade 1994: 219).

With the dawn of the neoclassical period in the 18th century, a renewed interest surged in classical Latin literature, regarded as representative of a golden age and an era of linguistic purity. Several writers, including Dryden, Defoe and Swift, initiated an intellectual movement to regulate English, to impose a system of rules like the ones

31 The decision to have all administrative documents written in English was probably motivated by political reasoning rather than an interest in the language itself (cf. Richardson 1980).

available for classical Latin, and to refine and purify the language by removing its defects. However, the idea of an authoritative dictionary by an English Academy, modelled on the Italian, French, Spanish and Swedish examples of the time, was abandoned.

It was not until the second half of the 18th century that prescriptivism started to take a stronger effect on English (cf. Stein 1994: 5). Many rules found in English grammars today were drawn up during this period (cf. Tieken-Boon van Ostade 1994; Sundby, Bjørge and Haugland 1991). In 1755, Samuel Johnson published his *Dictionary of the English Language*. Johnson observed that "tongues, like governments, have a natural tendency to degeneration; we have long preserved our constitution, let us make some struggles for our language" – a comment which illustrates the felt connection between the political and linguistic situations of the time (cf. Riley and Parker 1998: 25). Johnson's dictionary was soon followed by other major grammars such as *A Short Introduction to English Grammar* by Robert Lowth in 1762. This popular prescriptive opus illustrates the correct and incorrect use of English by means of 'right' and 'wrong' examples. Like Sapir once said, "sometimes we can feel where the drift is taking us even while we struggle against it" (1921: 166).

Non-standard pronouns were one of many undesirable linguistic features to attract attention in English grammars of the time. In his *Reflections on the English Language* from 1770, for example, Robert Baker described non-reflexive *themselves* as "mere Shopkeepers Cant" and predicted that it would "always sound contemptible in the Ears of Persons of any Taste" (cf. Tieken-Boon van Ostade 1994: 231).

The 18th century concept of correctness was closely linked to that of education and politeness (cf. Klein 1994). Whoever used the correct linguistic form was regarded as well-versed in polite conversation; who didn't ran the risk of being judged socially and intellectually inferior.[32] Dialectal, in this view, was synonymous with uneducated and uncivilised (cf. Stein 1994; Crowley 1989). At the same time, however, a very different standpoint was already on the rise:

> In modern and living languages, it is absurd to pretend to set up the compositions of any persons whatsoever as the standard of writing, or their conversation as the invariable rule of speaking. With respect to customs, laws, and every thing that is changeable, the body of a people [...] will certainly assert their liberty, in making

32 In Lindley Murray's *English Grammar* from 1795, 'good' grammar was associated with virtue and religious propriety (cf. Riley and Parker 1998: 28). New grammars were also being compiled in the United States with the desire to distinguish an independent American standard, for example by the lexicographer Noah Webster.

> what innovations they judge to be expedient and useful. (Priestley 1762, quoted in Riley and Parker 1998: 28)

3.2.2 Appreciation of linguistic diversity

The 19th century saw another important step towards the standardisation of English through the introduction of the Elementary Education Act in 1870. This act set the standard for the schooling of all children in England and Wales. It was followed ten years later by the introduction of compulsory school attendance until the age of 12.

Around the same time, the first systematic descriptions of English non-standard varieties started to appear. Most of them were publications of the newly formed English Dialect Society founded in 1873. The goals had shifted: this time, the interest did not lie in the purification and unification of the language but in its diversity. The English Dialect Society did not exist for very long. It dissolved in 1896, after completing a large-scale collection of data that would later become Wright's *English Dialect Dictionary* (1898–1905). The society's publications, however, represented the first major systematic contributions to dialect research in the United Kingdom. Among these publications were the regional dialect glossaries and grammars of Barnes (1886) and Elworthy (1975–1886), and the two cross-dialectal works of Ellis (1869–1889) and Wright (1905).[33]

In these new dialectological studies, it was the rural varieties which received the most attention. They were attributed a special status as prestigious dialects, said to be "in many ways 'purer', of greater ancestry and better pedigree" than Standard English (Ihalainen 1994: 206). The criticism turned towards urban dialects:

> The metropolitan county presents little in its dialect worthy of remark, being for the most part merely a coarse pronunciation of London slang and vulgarity. (Halliwell 1887, Vol.I, xxiv)

Although the idea of equality of treatment for all linguistic varieties was still in its infancy, studies of English were becoming more descriptive.

33 Earlier accounts were less comprehensive. Early references to dialect features in the travel literature include, for instance, Daniel Defoe (1724–1726) *Tour thro' the whole Island of Great Britain*, which probably contains the first mention of the Northumbrian burr.

3.3 The standard Modern English paradigm

According to Stein (1997: 35), the study of dialect grammar "can never be the study of dialect grammar alone, but always the study of dialect grammar in its dialectic with the study of standard grammar." In the present study, the general point of reference will be the Standard English paradigm of personal pronouns and reflexives presented in Table 3.1.[34] This paradigm reflects the prescriptive use of subject, object and self-forms as found in official documents, grammars, dictionaries, news broadcasts and education (cf. Greenbaum and Nelson 2002[2]: 3).

The standard paradigm is characterised by a strict form–function correlation in all syntactic contexts, obeying the 'one form – one meaning' ideology which typically lies at the heart of standardisation (cf. Stein 1997: 36–39; Milroy and Milroy 1993: 3–4). Standard English pronouns agree with their antecedents or referents in person (1/2/3) and number (SG/PL), and in the 3SG also gender (m/f/n). Each form encodes a specific syntactic function (subject/object/reflexive). The only two homonyms are *you* and *it*: *you* is both singular and plural, and *you* and *it* have both subject and object function.

Table 3.1: Standard English paradigm of personal pronouns and reflexives

Standard English		subjects (form = function)	objects (form = function)	self-forms (reflexive/ intensifier function)
SG	1	I	me	myself
	2	you	you	yourself
	3 m	he	him	himself
	3 f	she	her	herself
	3 n	it	it	itself
PL	1	we	us	ourselves
	2	you	you	yourselves
	3	they	them	themselves

Person

There are three persons which denote different speech roles: the speaker (1), the hearer (2) and third entities (3; cf. Forchheimer 1953: 5–6; Bühler 1934: 79). First and second person pronouns differ from third person pronouns in so far as only the former refer to speech act

34 Second person *th-* forms such as *thyself*, *thy* and *thine* could be included as "part of a general 'elevated' register" (Wales 1996: 167). However, even the Queen of England has stopped using these forms in her official speeches (cf. Wales 1994).

participants. The empirical findings in this study will show that person is the morphosyntactic category with the least variation, due to its importance for the correct processing of pronominal expressions. The only non-standard phenomena in the corpus are the use of existential *it*, dummy object *it*, and the use of generic question tags (discussed in ch. 6).

Number

All pronouns except *you* have distinct SG and PL forms. From a semantic perspective, plural formation in the first and second person involves conjunction, and plural formation in the third person involves plurality (cf. Bhat 2004: 10). In other words, *we* and *us* always refer to one speaker plus one or more non-speakers (with no overt inclusive–exclusive distinction), and plural *you* refers to several hearers, or one hearer plus one or more non-hearers. Third person plurals, on the other hand, refer to more than one referent of the same type: *boys* 'boy_1 + boy_2 (+ boy_3 +...boy_n)'. While conjunction naturally emphasises the differences between the conjunct elements, plurality emphasises their similarities. Some grammarians have therefore suggested that *I* and *we* are in fact distinct lexemes (cf. Huddleston and Pullum 2002: 1465).

Gender

Gender is only distinguished in the third person singular, where it is a means of minimising referential and coreferential ambiguity. Male referents are referred to by masculine pronouns (*he*, *him*), female referents are referred to by feminine pronouns (*she*, *her*), and referents without an assigned gender are referred to by neuter *it*. Animacy distinctions are also found in other pronominal categories, for example the relative pronouns *who* vs. *which*. Compare Map D11 for a typology of gender distinctions in independent personal pronouns.

Case

Standard English has lost its morphological case distinctions everywhere except in pronouns. The language has structural case which is licenced by the pronoun's syntactic position. This becomes apparent under A-Movement: *he called her* becomes *she was called by him* (not so in earlier stages, cf. Van Gelderen 2000a: 208). In other languages such as Icelandic, non-structural case does not change with movement, as seen in examples

(35a) and (35b), or it is selected lexically, as seen in the 'quirky subjects' in (35c) and (35d) (cf. Sigurðsson 1992).

(35) a. Þeir skiluðu **Maríu** bókinni. (Woolford 2006: 118)
they returned Mary-DAT the-book-DAT

b. **Maríu** var skilað Þessari bók. (Woolford 2006: 118)
Mary-DAT was returned this book-DAT

c. **Bátinn** rak á land. (Woolford 2006: 121)
the-boat-ACC drifted to the-shore

d. **Barniu** batnaði **veikin**. (Woolford 2006: 122)
child-DAT recovered-from disease-NOM

Reflexivity in Standard English is encoded exclusively by anaphoric pronouns in *-self/ -selves*, which are frequently treated as a separate pronominal category. The standard paradigm contains both adjectival forms, like *him-self*, and possessive forms, like *my-self*. In the vernacular, these forms are frequently regularised to 'POSS + self', as in *hisself*, *theirselves*, and so forth (e.g. Miller 1993; Beal 1993).

Standard reflexives occur in object or prepositional complement position. They require a coreferential, sentence-internal antecedent as seen in (36a):

(36) a. She$_i$ (Mary) saw herself$_i$ (Mary) in the mirror.

b. *She$_i$ (Mary) saw herself$_j$ (Jane) in the mirror.

c. *Herself saw nothing.

In the analyses, case will be shown to be the most fragile category with the largest amount of empirical variation.

Self-form intensifiers

In their emphatic function, appositional self-forms are used to highlight the corresponding noun (*he himself...*). Even if these intensifiers are often referred to as 'emphatic pronouns', they do not, strictly speaking, belong in this grammatical category. Self-form intensifiers are more adequately described as 'focus particles' (cf. König 1991) which "relate a periphery of alternative values (Y) to a centre X, identified by the expression they combine with" (König and Siemund 1997: 96).

Formally identical reflexives and intensifiers are usually distinguished distributionally (for a typological perspective, see König and Siemund 2000). In Standard English, self-form intensifiers can be divided into adnominal, adverbial inclusive and adverbial exclusive cases as shown in the following examples from König and Siemund (1997: 96):

(37) a. The director **himself** will talk to us. (adnominal)

b. I am a little short of cash **myself**. (adverbial inclusive 'me too')

c. Mary earned all that money **herself**. (adverbial exclusive 'alone')

In non-standard data, the sentence structure occasionally varies; consider the FRED example *And he never drunk beer himself -- he couldn't afford it* (text SAL_018). An investigation of self-form intensifiers, however, is beyond the scope of this study. The empirical analyses in Part II will only consider self-forms in reflexive and personal pronoun functions.

Part II

Empirical Study

Chapter 4
Data and methodology

All analyses presented in chapters 5 through 15 are based on the England component of the Freiburg Corpus of English Dialects. The text selection used in this study is described in detail in 4.1.2–4.1.4. More information about the corpus is available at http://www2.anglistik.uni-freiburg.de/institut/lskortmann/FRED.

4.1 The Freiburg Corpus of English Dialects (FRED)

4.1.1 Corpus design

The FRED corpus was compiled at the University of Freiburg, Germany, under the guidance of Professor Bernd Kortmann during the period 2000 to 2005. The primary incentives for this new corpus were the research team's interest in morphosyntactic variation in spoken British English and the need for a geographically well-balanced and machine-readable database.[35] The end result is a 2.5 million word corpus of spontaneous conversations.

FRED is a monolingual spoken-language corpus. It consists of full-length interviews with native senior citizens from England, Wales, Scotland, the Hebrides and the Isle of Man. Most of the interviews were recorded during the 1970s and 1980s, representing the traditional regional varieties of spoken British English in the second half of the 20th century. The conversations selected for this corpus are so-called oral history interviews in which the interviewees talk about their life memories. Since this type of interview originally aims at the elicitation and preservation of historical information and folklore, it provides an ideal basis for conversations in which the interviewees stay distracted from their own linguistic behaviour. According to Labov (1997: 395), the elicitation of narratives of personal experience is the most effective method in making people forget to pay attention to the way they speak and hence switch to a more casual register. This makes oral history interviews particularly valuable for the study of spontaneous speech. An additional advantage of the material collected in FRED is that all interviewers were native speakers of English, some of them of the same regional variety as the speakers themselves.

35 The author of this study was part of the research team.

All interviews in the corpus are face-to-face conversations in a private setting, usually the speaker's own home. Well-known difficulties for the elicitation of spontaneous speech, such as the observer's paradox (Labov 1972: 113), the tape recorder effect (ibid.) and effects of hyperadaptation (Trudgill 2004b: 62),[36] are thus minimised.

Regarding the formality level, the FRED interviews range somewhere between 'casual' style, used with friends and acquaintances, and 'consultative' style, used with strangers and distant acquaintances (cf. Joos 1967: 24). On a continuum of orality and literacy, the FRED interviews would be situated towards the orality end. This is visualised in Figure 4.1, which shows the position of FRED in the model of medial and conceptual orality and literacy by Koch and Oesterreicher (1994: 588). This model distinguishes two aspects of register: the medium (spoken/ phonic vs. written/ graphic) and the degree of conceptual orality or literacy. The FRED interviews are medially phonic and conceptually oral, situated close to the familiar conversations represented by letter a.

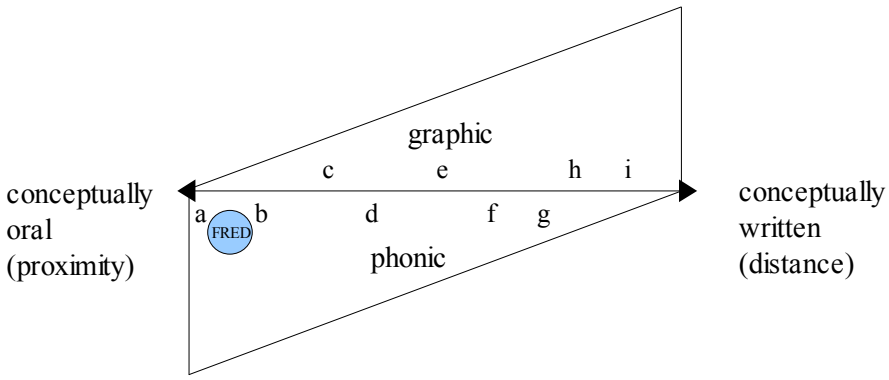


Figure 4.1: FRED on the orality–literacy continuum (after Koch and Oesterreicher 1994) (a 'familiar conversation'; b 'telephone conversation'; c 'private letter'; d 'job interview'; e 'newspaper interview'; f 'sermon'; g 'scientific presentation'; h 'editorial'; i 'legal document')

In Koch and Oesterreicher's model, conceptually oral language is the sort of language that expresses proximity between the participants, as opposed to conceptually written language

36 Speakers who are conscious of the way they talk are likely to adapt to the acrolect they have in mind. One side effect of this reaction is hypercorrection (cf. Boyland 2001), another is the loss of linguistic features that are associated with an informal style (cf. Labov 1966: 10, 'autoprescriptive norms'). Labov (1971: 450) has observed that "[w]henever a subordinate dialect is in contact with a superordinate one, linguistic forms produced by a speaker of the subordinate dialect in a formal context will shift in an unsystematic manner towards the superordinate. Although one can predict the general magnitude of the shift, social correction of this sort operates unpredictably upon particular forms, and furthermore, we have no guarantee that there will not be reverse hyper-correction [...] An educated informant who senses what the linguist is looking for may show 'hyper-incorrect' forms [...] Under 'formal context' we must include any situation where more than the minimal attention is paid to language, for whatever reason."

which expresses distance. Prerequisites for the former type are privacy, familiarity between the participants, face-to-face communication and a free development of conversation topics (cf. Koch and Oesterreicher 1985). The oral history material in FRED fulfils most of these conditions. It is hence much closer to the conceptually oral end of the continuum in Figure 4.1 than, for example, the job or newspaper interviews represented by letters d and e.

The FRED corpus is ideally suited for the present study for the following reasons:

- Spontaneous speech data: For the reasons specified above, oral history interviews provide one of the best types of material for an investigation of spontaneous speech.
- Traditional dialect data: The present study is specifically interested in traditional varieties, one reason being the gradual loss of these varieties due to progressive dialect levelling as mentioned in section 2.1. The low-contact data can be expected to exhibit more variation than higher contact varieties. Szmrecsanyi and Kortmann (2009: 76), for instance, have shown that it is the traditional, low-contact L1 vernaculars which "are on almost every count (ornamental complexity, grammaticity, irregularity) more complex than high-contact L1 varieties of English."
- Data volume: Morphosyntactic phenomena are usually less frequent than, for example, phonological phenomena. Variation in pronoun use is a particularly complex topic. It involves many different aspects and features, not all of which can be expected to appear frequently. Both the complete FRED corpus and the subcorpus selected for this study provide the quantity of data needed for an analysis of frequent and rare phenomena (see Table 4.1 for details).
- Area coverage: FRED is relatively well balanced regarding the amount of material available for each dialect area (cf. Hernández 2006; Szmrecsanyi and Hernández 2007). This property facilitates cross-dialectal comparisons of the phenomena under investigation and at the same time allows to zoom in on individual geographic areas.
- Extra-linguistic information: Although FRED was not designed for sociolinguistic studies, a limited set of variables was specified for each speaker, including the dialect area, age and sex. This information allows to evaluate the findings with respect to regional variation and gender and age differences.
- Format: The corpus consists of two parts, the sound recordings and the corresponding transcripts. The fact that the original recordings are available allows the researcher to countercheck dubious occurrences, which benefits the general credibility of the results.

4.1.2 Subcorpus size and area coverage

The present study is based on a special subcorpus compiled from the England component of FRED. Henceforth, all references to corpus data and results refer to this subcorpus.

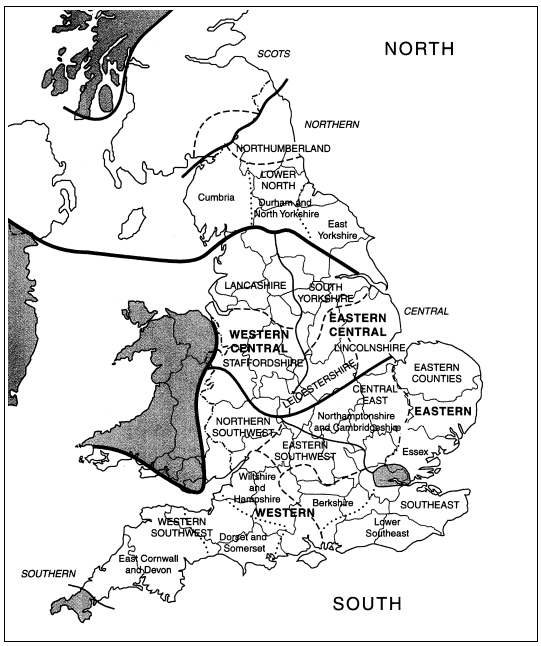
A general reference spreadsheet for the selected texts is available on the CD-ROM (speakers.xls). It contains the text ID for each interview as well as information on the dialect area, county and location, the speaker, the recording date of the interview and the storage location of the corresponding soundfile. The same spreadsheet also contains the text production in words for each interview, excluding text headers, transcriber's comments, tags and interviewer utterances.

Table 4.1: Size and coverage of the FRED corpus and subcorpus (subcorpus dialect areas: Southeast, Southwest, Midlands, North; additional dialect areas in complete corpus: Wales, Scottish Highlands, Scottish Lowlands, Hebrides, Isle of Man)

corpus size and coverage	**FRED subcorpus**	FRED complete corpus
words	**1.5 million**	2.5 million
texts	**176**	372
hours recorded	**180**	300
speakers	**210**	431
dialect areas	**4**	9
counties	**17**	43
locations	**83**	163

In order to facilitate the comparison of regional varieties, the England data were divided into four dialect areas: Southeast (SE), Southwest (SW), Midlands (MID) and North (N). These areas roughly correspond to the traditional dialect areas shown in Map 4.1[37] (historical boundaries were shown in Maps 3.1 and 3.2; modern dialect areas and possible future dialect areas in Maps D1 and D2, Appendix). The country's division into discrete dialect areas is necessary for practical purposes but it needs to be kept in mind that this is a construct used for orientation. In the words of Trudgill (1992: 7), "[w]e realize that dialects form a continuum, but for the sake of clarity and brevity, we divide this continuum up into areas at points where it is least continuum-like."

37 Trudgill's isoglosses are based on phonological criteria, including differences in pronunciation in words like *night*, *arm*, *few* or *milk*. The different areas roughly correspond to the ancient Anglo-Saxon kingdoms seen in Map 3.1; they are also similar to the dialect areas identified by Ellis (1869–1889).

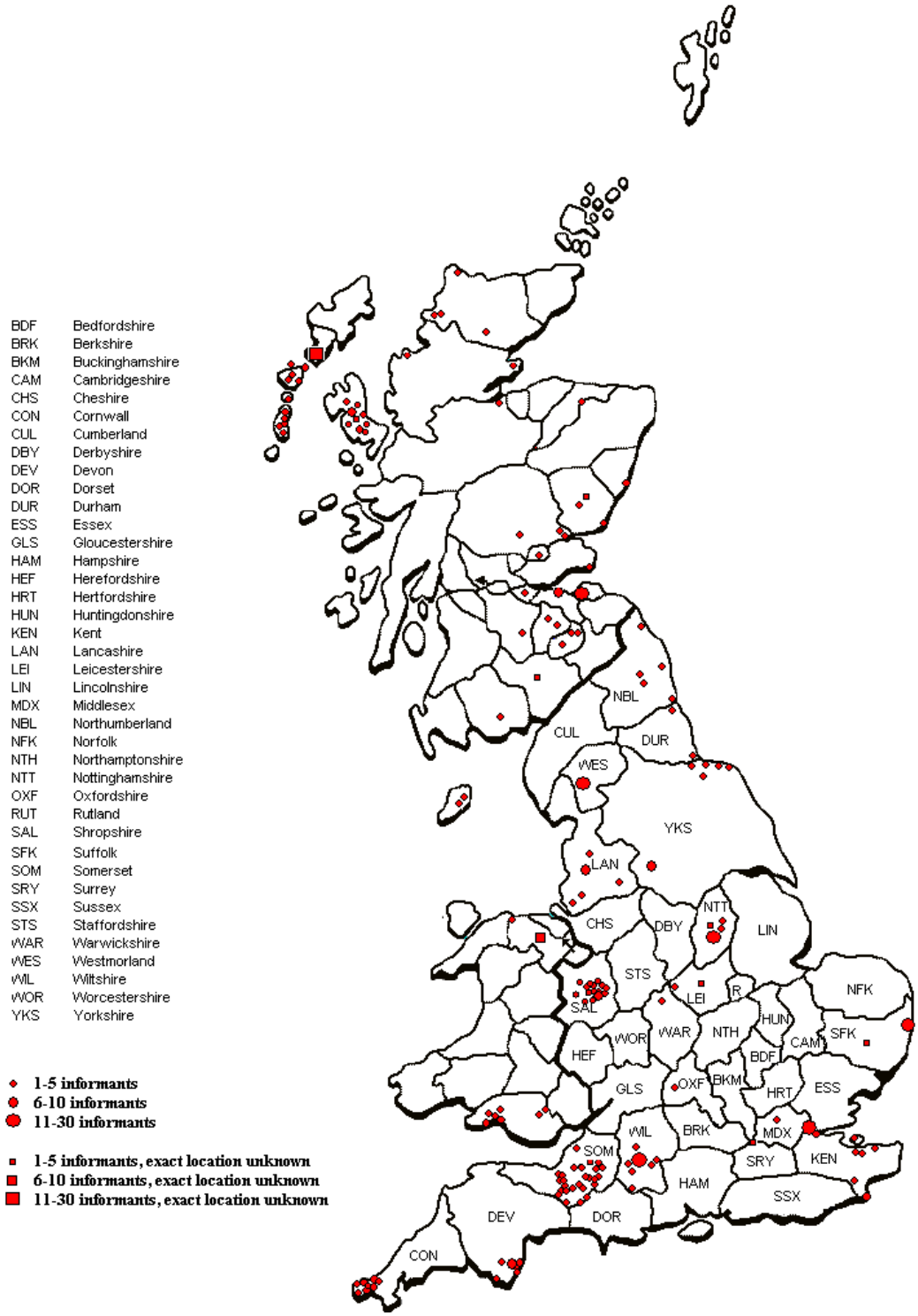


Map 4.1: England's traditional dialect areas (Trudgill 1999[2]: 34)

Map 4.2 visualises the data coverage of the complete corpus in the different counties. Given that administrative borders in the United Kingdom changed several times during the last century, county borders mentioned in different places do not always tally. Native speakers who talk about their county of origin often mention administrative entities of the past, or districts which are traditionally regarded as a distinctive region for historic or cultural reasons. One such example is East Anglia, which for some people consists of the counties of Norfolk and Suffolk, but for others includes parts of Cambridgeshire, Peterborough and Essex. The FRED team decided to use the so-called Chapman County Codes, a standard format for genealogical purposes.[38] Due to the age of the interviewees, it was resolved to apply the pre-1974 codes in use before the local government systems of England, Scotland and Wales were reorganised.

Regarding the selected subcorpus, it should be noted that only texts were included where information was available on both the county and location of the interview. The areal coverage in terms of counties was slightly balanced out by excluding some material from Suffolk, which in the complete corpus is much more strongly represented than the other southeastern counties.

38 For more information go to http://www.genuki.org.uk.



Map 4.2: The FRED counties (by Benedikt Szmrecsanyi)

4.1.3 Speakers

The subcorpus used in this study consists of 176 interviews with a total of 210 native speakers, excluding interviewers. The main speaker selection criteria for oral history interviews in general, and the present corpus in particular, were the speakers' age, a strong affiliation with the native surroundings and a generally low mobility. As a result, the corpus has a largely homogeneous speaker profile which is representative of the traditional dialects typically associated with a rural working-class population (cf. Ihalainen 1994: 252). The applied selection criteria also facilitate comparisons with former dialect surveys such as the Survey of English Dialects (SED).[39]

The typical FRED speaker spent most of his or her life in one particular geographic area so that his or her linguistic behaviour can be considered characteristic of that area. Most speakers grew up before World War II – a time which marks a "major cataclysmic event after which wide-ranging social and economic changes (with concomitant linguistic changes) came into effect" (Anderwald and Wagner 2007: 36) – and most of them left school at a very young age, usually 12 or 13, with no continuing education.[40] Most interviews were recorded during the 1970s and 1980s, which means that most speakers were in their 70s or 80s at the time of the interview.

For this study, the homogeneity of the speaker profile was further increased by including only those texts where information was available on the birth decade of the main interviewee (i.e. the person producing the largest amount of text in the interview), and where the main interviewee was born during the period 1890 to 1920. The oldest speaker in the text selection used for this study was born in 1877, aged 102 at the time of the interview. The mean age at recording date is 79.

Table 4.2 shows a breakdown of the overall text production in the subcorpus by dialect area and birth decade. The only speaker born in 1920 is included in the 1910 category. Due to the fact that some speakers were interviewed twice and some interviews have more than one speaker, there is a discrepancy between the total number of speakers (210) and the

39 Keeping in mind that the SED was collected slightly earlier than the FRED interviews (1950–1961, followed by recordings until 1974), and that it used a different elicitation technique (questionnaire; cf. Orton 1962; Orton et al. 1962; 1978; Upton et al. 1987; 1994). In the SED, pronoun questions which are of interest to the present study are mainly found in books VII, VIII and IX. The elicitation of data in the SED also led to prolonged stretches of natural conversation which were documented in the fieldworkers' notebooks. These notebooks not only contain the speakers' responses to the official Dieth–Orton questionnaire but supplementary notes, including remarks on linguistic features and local customs and traditions. The fieldworker notebooks are now available online from the Leeds Archive of Vernacular Culture at http://www.leeds.ac.uk/english/activities/lavc/IMdocs.htm.

40 Variables like language proficiency, occupation, and social and economic status were not considered in the selection process due to known uncertainty factors (cf. Sankoff and Laberge 1978).

total number of texts (176).

Table 4.3 shows an additional breakdown of the text production by speaker sex. FRED consists to a large extent of so-called NORMs, the non-mobile older rural male speakers typically selected for traditional dialect studies (cf. Chambers and Trudgill 1998[2]). The ratio of male to female speakers in the subcorpus is roughly 2:1 (137 vs. 65; 8 unknown); the ratio of text produced by the two groups is roughly 3.5 : 1 (78% vs. 22%).

Table 4.2: Speaker distribution and text production by birth decade (amount of text produced by speakers of different birth decades in the four England dialect areas: SE = Southeast, SW = Southwest, MID = Midlands, N = North)

speakers born in...	SE	SW	MID	N	TOTAL	text production in words
1890s	10	15	11	15	51	404,331 (28.3%)
1900s	9	23	25	21	78	618,251 (43.3%)
1910s	3	20	11	16	50	386,129 (27.0%)
unknown	5	18	0	8	31	19,161 (1.3%)
TOTAL	27	76	47	60	210	1,427,872 (100%)

Table 4.3: Speaker distribution and text production by speaker sex

speaker sex	SE	SW	MID	N	TOTAL	text production in words
male	20	50	34	33	137	1,114,293 (78.0%)
female	6	20	13	26	65	311,533 (21.8%)
unknown	1	6	0	1	8	2,046 (0.1%)
TOTAL	27	76	47	60	210	1,427,872 (100%)

4.1.4 Recordings and transcripts

The FRED corpus consists of entire conversations.[41] The recordings usually last between 30 and 90 minutes, the longest interview being almost 4.5 hours and 47,000 words long (text KEN_003). Consequently, the size of individual texts and the amount of text per dialect area varies (roughly between 300,000 and 426,000 words per area).

41 In a few cases the tape was interrupted at some point due to technical problems during the interview.

The interviews were transcribed orthographically as ASCII texts with no text-processor formatting in order to be compatible with text retrieval systems like TACT or WordSmith. Pre-existing transcripts made available to the research group by other institutions were included if they contained word-for-word transcriptions by an experienced linguist or a fieldworker with a linguistic background. Discourse features such as double negation, repetitions and truncations were reinserted where necessary and paralinguistic features like laughter were also marked.[42]

Overall, the transcripts read like normal writing with some minor exceptions: upper case, for instance, is used to indicate the beginning of direct speech citations; abbreviations and numbers are written out in full (*etcetera*, *nineteen hundred and sixty-four*); and auxiliary and modal verb contractions are divided by space (*he 's*, *they 'll*, etc.). Non-standard realisations are rendered most consistently for grammatical items, for example the reduced definite article *t'* which is distinguished from the 3SGn pronoun *'t*, contracted negations like *-nae* or *ain't*, reduced prepositions like *o'* 'of', and modals like *mun* 'must'. Pseudo-dialectisms were avoided: it is *fellow* not *feller*; *should 've done* not *should of done*; *maybe* not *mebbe*. Phonological dialect features are not rendered consistently, but transcribers were encouraged to mention recurrent features that appeared characteristic of the interviewee's speech. Regular non-standard pronunciations, such as initial *h*-dropping in 3SG pronouns, are either reflected in the text or mentioned in the text header.

Each interview in the corpus has its own text identification number consisting of a three letter county code followed by underscore and a running number, for instance the Yorkshire interviews YKS_001, YKS_002, and so forth. The text ID is the same for the recorded conversation and the corresponding transcript, for example YKS_001.wav and YKS_001.txt. Longer interviews may have more than one soundfile, for example NTT_006A.wav and NTT_006B.wav. The county where the interview took place can be easily inferred from the three letter county code in the file name: YKS for Yorkshire, LAN for Lancashire, SOM for Somerset, and so on.

Figure 4.2 shows the beginning of a typical FRED transcript. The text, LAN_010, is preceded by a text header with important editorial information, including copyright restrictions. Some headers contain short descriptions of the speaker's accent and dialect features. For each text, the header also includes a set of standardised COCOA-type tags (cf. Oakes 1998: 156) that specify the dialect area (<area *value*>), county (<county *value*>) and location (<location *value*>) of the interview. In addition to the geographical

42 According to Labov (1966: 9), laughter is a channel cue for casual speech style.

information, there are tags for different speaker variables: the speaker ID (<who id *value*>), sex (<who sex *value*>), birth date (<who dob *value*>), birth decade (<who dec *value*>) and age at the time of the interview (<who age *value*>). There are also two tags for the recording date and recording decade of the interview (<RECDAT *year*> and <RECDEC *3 digits decade reference*>). The speaker tags contain information on all interlocutors except the interviewer. Information is given where available; otherwise, the tag value is left empty.

header with editorial information

[full permission to use the material]
[Lancaster University; Centre for North West Regional Studies, oral history archive]
[Mr H3L, C60, Tape RSC/79/128; transcript pp. 1–14]
[interviewed February 1975 by Dr Elizabeth Roberts]
[dialect features include final g-, t- and d-dropping, initial h-dropping, occasional initial th- dropping, elides 'to' in 'used to', or weakens it to 't' ', elides definite and indefinite articles, uses 'o' ' instead of 'of']
[Informant: H3L born 5.8.1904; job: errand boy, munitions, stripper and grinder, hoistman, mill warehouse, silk winding, velvet finisher; education: started working part time at age 9; full time at age 12; working class; father: mill worker; mother: weaver in mill before marriage; dress maker]
[Informant: MH3L; his wife]

COCOA tags with text and speaker variables

<text LAN_010>	***text ID***
<area N>	***dialect area***
<county Lancashire>	***county***
<loc Preston>	***location***
<who id H3L,MH3L>	***speaker IDs (interviewer excluded)***
<who age 71,>	***speakers' age at recording date***
<who dob 1904,>	***speakers' birth years***
<who dec 190,>	***speakers' birth decades (190 for 1900s)***
<who sex m,f>	***speakers' sex***
<RECDAT 1975>	***recording date (year)***
<RECDEC 197>	***recording decade (here: 1970s)***

start of interview

{<u Int> Where were you born?}
<u H3L> #I was born in Lancaster. #I was born at Primrose, that 's just (trunc) a- (/trunc) around by the Bowerham Hotel.
{<u Int> Uh hm.}
<u H3L> #I were brought up there, and first thing I remember was living with grandfather and grandmother, mother lived there. #And times were bad then. #Mi father was a blacksmith. #And he worked, he worked at the Wagon Works which is on Caton Road which is Standfast Dyers and Printers now. #And I remember going down there with an aunt of mine taking dad's dinner down and sitting in what they called the tunnel fields because there 's a tunnel there goes under the, under the canal.

Figure 4.2: Header and first lines of a FRED transcript (text LAN_010)

4.2 Data extraction, coding and statistics

Preparing the data for the analysis involved two main steps: extracting all pronominal occurrences from the corpus and function-coding each case. The resulting spreadsheets for the individual pronoun forms are available on the CD-ROM.

Data extraction

Following a corpus-driven, form-based approach, all personal pronouns and self-forms were extracted from the data, amounting to over 132,500 individual occurrences. Interviewer utterances were excluded from the word search with WordSmith Concord by using the markup feature of curly brackets at the beginning and end of each interviewer passage (see Figure 4.2). Also excluded were cases marked as *unclear* in the transcripts, cases where the pronoun is part of a set phrase, noun or adjective, as in *do-it-yourself shop*, and verb-pronoun mergers as in *genny* 'gave him'.

Random sampling was used for *you* and *it*, the two pronouns with missing case distinction. For *you*, four samples were extracted of 2000 cases per dialect area. Additional searches were conducted in connection with specific phenomena such as '*you* + qualifier' plurals (ch. 11.2) and overt *you* imperatives (ch. 12.1). *It* samples were analysed in connection with the phenomenon of gendered pronouns (ch. 7).

Two types of irregularities had to be taken into account in the extraction process: pronunciation variants reflected in different spellings, and non-standard sentence structures and truncations. Regarding the first type of irregularities, only two pronominal variants are mentioned in the official transcription guidelines: the reduced possessive determiner /mɪ/ 'my', as in *mi brother*, which is orthographically distinguished from personal pronoun *me*, and the reduced forms *'em* 'them' and *'er* 'her', which are orthographically distinguished from pause fillers like *ehm* and *ehr*. In addition to these variants, personal pronouns and self-forms occur in a variety of different spellings in the corpus. Some of them are well-known dialect forms, for instance with initial *h-* deletion, while others are less predictable, for example *yuh* (2SG/ PL personal pronoun and 2SG/ PL possessive with final *-r* deletion). The problem was solved by using the asterisk option in WordSmith Concord: a word search for {m*se*}, for example, returns all instances of {*myself*, *miself*, *meself*, *m'self*}. At the same time, the results show the absence of other variants known from the dialectological literature, such as *mesell*, *thyself*, *theeself* and *hoo* 'she' (cf. Trudgill 1999[2]: 101). Table 4.4 lists the different spellings detected in the corpus. In the following

chapters, each pronoun will be referred to by its standard form: references to *myself*, for example, include all instances of *myself, miself, meself* and *m'self*. It should be noted that reduced forms like *en* and *um* were classified according to the respective referent.[43]

Table 4.4: Spelling variants of the different pronouns in the corpus

pronouns		spelling variants in the corpus
1SG	*I* *me* *myself*	*I*, *ahm* 'I'm' *me*, *mi* *myself*, *miself*, *meself*, *m'self*
2SG/ PL	*you* *yourself*	*you*, *y'*, *yuh*, *ya*, *ye*, *yeh*, *e*, *ee*, *oo*, *tha*, *thee*, *thou* *yourself*, *y'self*, *yerself*, *yoursel'*, *yuhself*, *yourselves*
3SGm	*he* *him* *himself*	*he*, *'e*, *a* *him*, *'im*, *em*, *um*, *en*, *un*, *'m* *himself*, *hisself*, *'iself*
3SGf	*she* *her* *herself*	*she* *her*, *'er* *herself*, *hersel'*, *'erself*
3SGn	*it* *itself*	*it*, *'t*, *-it* (*innit*, *wunnit*, etc.) *itself*, *itsel'*
1PL	*we* *us* *ourselves*	*we* *us*, *'s* *ourselves*, *ourself*
3PL	*they* *them* *themselves*	*they*, *thy*, *'ey*, *theys* *them*, *'em* *themselves*, *themself*, *theyselves*, *theyself*, *theirselves/ -self/ -sel'*

Regarding the delimitation of syntactic units, spontaneous speech is not always straightforward. In the data at hand, for instance, there is a fair amount of truncations, anacolutha and non-standard sentence structures (e.g. inversion of verb arguments). The identification of antecedents is complicated by the ability of personal pronouns to refer to chunks of discourse that vary in size and textual distance. Referential chains can be interrupted, and eventually picked up again, by the speaker or other discourse participants.[44] In many occurrences, the wider context was therefore taken into consideration. In the case of anacolutha, it was decided to include those occurrences where a subject–verb sequence was recognisable for the clause containing the pronoun. In the

43 Note the frequent use of pronominal /ən/ and /əm/, rendered as <en>, or <um> in the transcripts. Due to vowel reduction and changes in the final consonant, the distinction between *him*, *it* and *them* can become unclear.

case of verb elision, the syntactic role of the pronoun was inferred from the overall context. In sentence (38), for example, *me* is a coordinated subject. Furthermore, a distinction had to be made in coordinated examples between pronouns that form part of a coordinated NP, as seen in (38), and pronouns that are the subject of one of two main clauses coordinated by *and*, as seen in (39).

(38) That bed was ours. Eh, only us, eh, well, four of us, laid in that, see. [...] **Me** and Lil down the bottom, and Elsie and eh Edie up the top. (speaker FM, text LND_003)

(39) Tom was on the buoy and **us others** -- that left seven, didn't it? -- were in this little boat. (EF1, SFK_003)

Coding

After the extraction, all occurrences were coded according to their syntactic function. A complete list of the different part-of-speech codes (POS) with examples is shown in Table B3 in the Appendix. Figure 4.3 is an example of the pronoun spreadsheets that were used in the analyses. The spreadsheets are available on the CD-ROM. Each spreadsheet contains a concordance column and the corresponding extra-linguistic variables for each occurrence, including the dialect area and county (codes in Tables B1 and B2), the speaker of the utterance (see speakers.xls), the speaker's birth decade (three digits; e.g., 189 for 1890s) and the speaker sex (1 = male, 2 = female). Each spreadsheet also contains two columns for notes and specifications.

The grammaticality status can be easily obtained for each occurrence by combining the respective pronoun form and POS code. For example, S-forms like *I*, *he* or *they* with a subject function code (POS 10, 11, 12 or 13) are standard subjects in agreement with the Standard English paradigm shown in Table 3.1, whereas *I*, *he* or *they* with an object function code (POS 20, 21, 22 or 23) classify as non-standard pronouns. The empirical findings will be compared against the following prescriptive norms:

44 Compare Rocha (2000), on the analysis of anaphoric relations in empirical data. Rocha finds that "the frequency rate of cross-sentence anaphora is higher than the one for within-sentence anaphora. In fact, it is difficult to identify the structural sentence – as usually understood – when the corpus data are made up of real-life dialogues" (p. 83). Rocha addresses the problems of defining a) a discourse unit that represents a reliable basis for the analysis of dialogue data, and b) "associations that are stable enough to ensure correct identification of the antecedent" (p. 84). For these purposes, Rocha develops a complex annotation scheme which includes information on the topicality status of the antecedent, the type of knowledge used to process the anaphoric reference (collocational knowledge, full discourse knowledge, etc.), the anaphor type (subject pronoun, full NP, VP ellipsis, etc.), the antecedent type (explicit, implicit, no referent, etc.), and the global and local topics in the respective dialogue segment.

a) subject case is required if the pronoun is used in subject function, irrespective of the complexity of the subject noun phrase (simplex or coordinated NP) and the pronoun's position within the subject NP (initial or final);

b) subject case is required in subject complement function; linking verb *be* is followed by subject case (*It is I*); according to Phythian (1988: 41), linking verbs "are followed by complements, not direct objects. Logically, then, when pronouns form complements they should not be in the object-form.";

c) subject complements followed by a relative clause also require subject case (*it was she who...*);

d) object case is required if the pronoun is a direct object, indirect object or prepositional complement, irrespective of the complexity of the object NP (simplex or coordinated) and the pronoun's position within the object NP (initial or final);

e) reflexive case (self-form) is required if, and only if, the pronoun has reflexive function or is used as an intensifier; this rule applies to all reflexive arguments, including direct objects (*she hurt herself*), indirect objects (*I'm giving myself permission*) and prepositional complements (*he always draws attention to himself*); according to Phythian (1988: 42), there is "no good reason for misusing these pronouns by substituting them for personal pronouns: *The house belongs to my brother and myself* (should be ... *and me*) [...] Such misuse is common. No *-self*, *-selves* pronoun is correctly used if a personal pronoun can replace it.";

f) all nominal and pronominal constituents within a coordinated NP have the same function and thus require the same case; according to an early prescriptive grammar by Murray (1816[3]: 301), "[c]onjunctions connect the same moods and tenses of verbs, and cases of nouns and pronouns";

g) all pronouns must agree in person, number and gender with their antecedents and extra-linguistic referents (cf. Murray 1816[3]: 95); this means that singular pronoun forms should be used to refer to individual persons or objects, plural pronoun forms should be used to refer to more than one person or multiple objects, feminine pronoun forms should be used for female referents, masculine pronoun forms for male referents, and neuter forms for neuter referents;

h) the pronoun's case is determined by the pronoun's function in the respective clause or sentence even if the verb is omitted (cf. Murray 1816[3]: 315); this includes brief disjunctive questions and answers, as in *Who is it? – I.*, or *Did you go there? – I?*;

i) in comparative constructions, pronoun case does not change if the main verb is elided; it is *he is taller than I* since the complete sentence would be *he is taller than I am*; it is *they love him more than her* since the complete sentence would be *they love him more than they love her*;

j) double segmentalisations are regarded as superfluous; this applies to all pleonastic pronouns, including non-standard benefactives, overt reflexives used with standard intransitive verbs, overt indirect objects used with standard monotransitive verbs, disjunctive pronouns and resumptive pronouns.

1SG_me.xls.ods - OpenOffice.org Calc

Datei Bearbeiten Ansicht Einfügen Format Extras Daten Fenster Hilfe

	A	B	C	D	E	F	G	H	I
1	pronounform	POS	Concordance	dialectarea	county	speaker	decade	sex	specification
2	me	10	I once dropped in a muck, **me** did there. (v 'interviewer laughs'). I were a little girl.	4	14	LanD1P	190	2	
3	me	10	But **me** sat here, who have been brought up as a boy, to believe: though shalt not kill, and him sayin' that, and he 's leader of the church, what does he make of me?	4	14	LanEG	190	1	
4	me	10	and then we used to get we towel and soap out, have a jolly good wash into spring water, and then **me** for flowers, I don't think (trunc) th- (/trunc) , buttercups, daisies, dandelions I 'll pick -- eh, I love flowers, me.	4	14	LanD1P	190	2	verb elision
5	me	11	But before Christmas, sometimes before Christmas **me and him** used to get a week's work up, house up the, big house.	1	1	HM	189	1	me and him
6	me	12	He said eh, **Our Ernie and me** is going to make some sandwiches we get them ready like and then we 'll call you in when we 're ready, so eh, goose, goose skin, (v 'laughs') pork rind, and father had thickened them with pepper, you know.	4	14	H3L	190	1	our *name* and me
7	me	17	Well that didn't make any difference because the basic thing was done **by me** and the firm was going a get the orders,	2	9	TrbrEGC	190	1	by me
8	me	19	There used to be about -- **oh deary me**, there was a housekeeper, the butler who was absolutely the kingpin	4	16	WesBX	190	1	deary me
9	me	20	So, when he come next time, he 'd give me, went to give **me**, y' know, give **me** the half a crown	1	2	FM	191	2	
10	me	21	on the Bank holiday these three went on holiday and left **me and another fellow from Hilltop** to come and get the water out and we had a tub tub and we used to run it down	3	11	NotA25	189	1	me and NP
11	me	24	{<u Int> I think you said your mum pushed you up there in a, in a bassinet?} <u JG> **Not me**. {<u Int> Ah-hah.}	1	2	JG	191	1	
12	me	30	(mname) John (/mname) said, No (reg sic=dammit) damn it (/reg), that isn't me, boy.	2	5	CAVA_WW	190	1	
13	me	34	**Me being light**, I used to jump right at the top. If you never all got on on one go, y' see, the last one would be hanging onto the arse of the other kid.	4	17	YksJV	189	1	me V-ing
14	me	36	I said, excuse **me asking** your age but, I said, you know when anyone is older they feel as if it 's ten years or more don't they.	4	16	WesBP	191	2	me V-ing

me

Figure 4.3: Spreadsheet with concordances, POS codes and extra-linguistic variables

Statistics

Numerical comparisons in this study focus on the statistical significance of the appearance of specific phenomena, and on comparisons of relative frequencies with respect to a certain criterion (mostly the syntactic context and geographical differences). The simple chi-square test, computed as $(O - E)^2/E$ (with O being the observed and E being the expected value), works well in the limit of similar populations. In the present study, however, we are also confronted with comparisons of populations which vary greatly: for example, when comparing the same phenomenon in different dialect areas, or in simplex vs. coordinated NPs, or when comparing the same phenomenon with first/ second person pronouns vs. third person pronouns. It was therefore decided to compute the poisson errors of the frequencies separately, taking the correlation of x and n into account:

$f = x/n$ population frequency; n being the total number of pronominal occurrences in a defined syntactic function in a defined population; x being the total number of non-standard cases in the same function and population as n

$\delta_f = \sqrt{f*(1 - f)/n}$ error on population frequency f

The independent frequency errors are then combined by adding them in quadrature. The distance of the frequencies divided by this combined error is the number of standard deviations between the two frequencies:

$$\Delta_{f1,f2} = \frac{f1 - f2}{\sqrt{\delta_{f1}^2 + \delta_{f2}^2}}$$

The corresponding *p*-value reflects the likelihood that both frequencies are just statistical fluctuations from a common frequency (that would be located at the weighted mean between them). In the case of widely differing populations, this approach yields more reliable results. It is typically somewhat more conservative than the simpler chi-square test.

It could of course be argued that even one occurrence is a significant occurrence (assuming strong falsifiability; cf. Penke and Rosenbach 2004). The analyses, however, aim to establish whether the findings describe more than just individual deviations from the standard grammar. A quantifiable degree of significance offers an objective means of

assessment which can guide us in how strong the overall deviation is. Where to set the threshold for observation is subject to discussion. For consistency reasons, the minimum confidence level applied in all analyses will be 95%, or $p < .05$.

4.3 CD-ROM

The study is accompanied by a CD-ROM which contains a complete copy of the selected texts, plus various soundfiles for examples shown in the different analysis chapters. The soundfiles can, for instance, be played with PRAAT, which is downloadable for free from http://www.fon.hum.uva.nl/praat. In order to guarantee the transparency and traceability of the analysis results, the CD-ROM also provides the main spreadsheet with all concordances for each pronoun form. On the CD-ROM you will find the following folders:

'Corpus': contains all transcripts used in this study; contains the general reference file speakers.xls, a complete listing of the interviews with text and speaker-related variables;

'Soundfiles': contains the soundfiles for selected examples in different analysis chapters;

'Spreadsheets': contains the main spreadsheet for each pronoun form with all cases extracted from the corpus (1SG_me.xls, 1PL_us, etc.), plus some some other spreadsheets used in the analyses.

Chapter 5
Preliminary remarks

5.1 The functional versatility of S-, O- and self-forms in the corpus

Besides frequency criteria, behavioural criteria are "the primary source of evidence for markedness within language structure" (Croft 1990: 77). In syntax, evidence that one variant is grammatically or functionally more versatile than another variant indicates the former variant's unmarked status ('syntactic criterion'); in morphology, behavioural evidence pertains to the number of morphological distinctions that a particular grammatical category possesses ('inflection criterion'); and in phonology evidence for markedness concerns the number of phonological environments in which a specific phoneme can occur. According to Croft (1990: 81), "[t]he element which occurs in a larger number of constructions is the less marked one."[45]

In the present study, it will be shown that the least marked personal pronoun forms are object forms. Before we start with the investigation of specific pronoun phenomena, it is worth taking a closer look at the functional diversity of the different formal types (S-forms, O-forms, self-forms). The function matrices provided in Tables C1 and C2 (Appendix) give a first impression of the functional range of each pronoun.

A first comparison shows that almost all of the investigated pronouns occur in standard as well as non-standard functions (except *itself* and *yourselves*, which are both extremely rare). The greatest functional diversity can be observed among 1SG forms, *me* being the overall most versatile personal pronoun form, and *myself* the most versatile self-form. In the corpus, *me* is used in 35 out of the 50 syntactic functions distinguished in Table B3, more than half of which are non-standard uses. In various functions, *myself* is the only self-form to appear in the data, an observation which fits the frequent references to non-standard *myself* in the dialectological literature.

An even closer look at the two matrices shows that all S-forms were occasionally used as objects and prepositional complements, and all O-forms occasionally had subject function. Furthermore, all self-forms were used as independent personal pronouns somewhere in the interviews, both in object and subject functions (except the very rare *yourselves*). In line with the historical facts described in 3.1, we find O-forms with

45 Croft, for instance, mentions verbs which occur in active but not passive voice.

reflexive function but no comparable S-forms.

An overall comparison of the three formal types identifies O-forms as the functionally most versatile and least marked type. Gaps in the function matrix do not invalidate this generalisation. Among all pronouns with distinct S- and O-forms, such as *I–me* or *he–him*, the latter form can be expected to appear in a greater variety of syntactic contexts. Of special interest are non-object functions where O-forms play a substantial role, and those functions where O-forms are the default choice (e.g., disjunctive subjects as in *Me, I go twice every Sunday*, or *I love flowers, me*; see ch. 12.5). The results can be summarised in a simple hierarchy:

(40) Functional Diversity Hierarchy

O-forms > S-forms > self-forms

This Functional Diversity Hierarchy is partially correlated with another distributional hierarchy depicted in (41). The Non-standard Frequencies Hierarchy describes the comparative likelihood of the three formal types to appear in non-standard occurrences. It illustrates the higher overall frequency of O-form and self-form occurrences with non-standard functions, as compared to the lower amount of S-form occurrences with non-standard functions (the vast majority of S-forms in the corpus being simplex subjects of tensed verbs).

(41) Non-standard Frequencies Hierarchy

O-forms/ self-forms > S-forms

In the data at hand, the two pronouns with the highest non-standard proportions are 3PL *they* and *them*, which is largely due to their frequent use as demonstrative determiners (see ch. 15). It is important to note that the variety of functions in which a specific pronoun occurs is not necessarily correlated with its overall non-standard frequency. The two hierarchies are, for example, correlated regarding *myself*, the functionally most diverse self-form (20 contexts, see Table C2) with the highest non-standard proportion of all self-forms (> 22%). Other pronouns, however, have a high functional diversity index but a relatively low non-standard frequency (e.g., 1SG *I*), and some pronouns are functionally

less versatile but exhibit high non-standard proportions in connection with one specific, highly frequent phenomenon (e.g., *they* demonstratives).

In the following, it should be kept in mind that whilst the two hierarchies provide a good first impression of behavioural tendencies, their strength lies in their simplicity as generalised descriptions. For example, it is true that O-forms are generally more versatile than S-forms, yet it will also become clear in the analyses that the pronouns of the 1SG are comparatively more versatile than other pronouns. The combination of both tendencies explains, for instance, why *I* occurs in a similar number of functions as some O-forms (compare *her* and *us*). In the discussion of individual phenomena, therefore, the two hierarchies presented in (40) and (41) will serve as a background for more specific observations.

5.2 Speaker distribution patterns

In spontaneous oral conversation, morphosyntactic choices are not only influenced by syntactic and extra-syntactic variables, they are also inextricably mixed with the speaker's subjective interpretation of, and reaction to, these variables. The speaker's wish to be seen in a certain light by other discourse participants, for example socially, may have an unforeseeable effect on his or her linguistic choices, including the production of hypercorrected variants. Inter-speaker and intra-speaker variation are additional variables which form an integral part of the variation observed in natural language data. Evaluative norms are, however, extremely hard to isolate, especially for data where the speakers can not be asked to comment on their choices in retrospect.

What we can identify in data like FRED are the underlying speaker distribution patterns of different phenomena, meaning the proportion of speakers who used a certain phenomenon in the corpus, and the mean relative frequency of this phenomenon in the interviews. Two phenomena may have similar frequencies but different distribution patterns, suggesting different interpretations. The present study makes a plea for the consideration of such patterns, revealing their importance in the interpretation of other variables (variation will, however, not be investigated on an idiolectal level).

In the individual analyses, the *Speaker Number* (SN: the number of speakers who used a specific phenomenon in the corpus), *Speaker Proportion* (SP: the proportion of speakers who used the feature out of the 210 speakers in the corpus) and the *Mean Speaker–Case*

Ratio (MCR: the total occurrences of the feature divided by SN) will be used to shed more light on the composition of simple frequencies. In combination, the SN, SP and MCR values represent a straightforward means of comparison for the different phenomena under investigation.

Three basic distribution patterns are identifiable in the data. These are described in some detail in 5.2.1–5.2.3 and visualised in Figures 5.1–5.3 in schematic charts based on an imaginary group of 25 speakers and a cut-off point at 10 occurrences per speaker.

5.2.1 Rare features

A rare feature, as defined in this study,[46] is typically used by few speakers and has but a few traceable occurrences in a given set of data. In Figure 5.1, the fact that most speakers do not use the feature is represented by a strong 0 baseline. A rare feature has characteristically low SN, SP and MCR values. Extremely rare phenomena naturally run the risk of being mistaken for a slip of the tongue. This assumption, however, becomes significantly less probable as soon as the SN and MCR values are larger than 1.

In the literature, rare phenomena are not necessarily mentioned less frequently, since rarities often attract attention (e.g., the relatively rare appearance of S-form objects). One example of a very rare feature in this study is the use of possessive marker *'s* on possessive determiners, of which there are only 4 cases by 4 different speakers (*mi's mother's aunt*, *them's daughters*, *ours nightfighters*, *ours Lancaster*; see 14.3). Rare features are of particular interest if they connect to other properties of the language in some way. The use of possessive *'s* on possessive determiners, for example, can be interpreted as an interference with the possessive marking of nouns in English. The use of determiner *mine* (also discussed in 14.3) could be a vestige of a determiner which in earlier stages of English existed alongside the younger and shorter form *my* for some time after *my* emerged, especially before words beginning with a vowel.

5.2.2 Speaker-specific features

More frequent features can be divided into two basic types which will be referred to as speaker-specific and generalised features. Speaker-specific features are used more frequently than rare features but only by some speakers, whereas generalised features are

46 A quantitative definition of 'rare' is not proposed. It would of course be interesting to see whether there is agreement among linguists as to where the line between 'rare' and 'not rare' should be drawn.

used frequently and by a substantial number of speakers.

In Figure 5.2, the underlying distribution of speaker-specific features is reflected in a wider range of values and individual peaks on the y-axis. A typical example in FRED is the use of subject *thee* (see 8.3.1). Subject *thee* is used 43 times by only 4 speakers, with the following distribution: speaker IBJP, 39 cases; WflsWGP, 2 cases; SRLM_PT, 1 case; SRLM_Y, 1 case. The results show that subject *thee* is highly speaker-dependent but not idiolectal.

5.2.3 Generalised features

Generalised features have the highest SN and SP values of all features. They show a relatively great homogeneity in the linguistic behaviour of the speaker group, as seen in Figure 5.3. Since generalised features occur with a certain regularity, they are more likely to be perceived as part of the language system than rarer or highly speaker-dependent phenomena.

In statistical terms, the generalised pattern is the only distribution pattern which can be used for quantitative comparisons, based on a sufficient number of speakers and occurrences. A typical example in FRED is the use of gendered *him* (ch. 7). More than 12% of all *him* in the corpus have non-human referents; gendered *him* was used 385 times by 72 different speakers; its distribution in the corpus is as follows: 22 speakers x 1 case, 10 x 2, 10 x 3, 5 x 4, 2 x 5, 4 x 6, 4 x 7, 3 x 8, 1 x 9, 1 x 10, 2 x 11, 1 x 12, 2 x 13, 3 x 15, 1 x 35, 1 x 47. Based on this distribution, gendered *him* classifies as a generalised feature.

It is not uncommon for generalised features to have a few high-frequency peaks indicating speakers who use them particularly often, as shown in Figure 5.4. One such example will be discussed in 8.3.7: the high proportion of subject *us* in data from the Southwest can be attributed to two speakers who use subject *us* exceptionally often (almost 50% of all 1PL subjects in each text). Here, it is a matter of interpretation whether these speakers are treated as outliers or whether they are still considered representative of their regional variety. Regarding the regional comparison of subject *us*, for example, the SW area has a significantly higher O-form proportion than the rest of the country, but if we exclude the two high-frequency speakers, the differences in frequency become statistically insignificant.

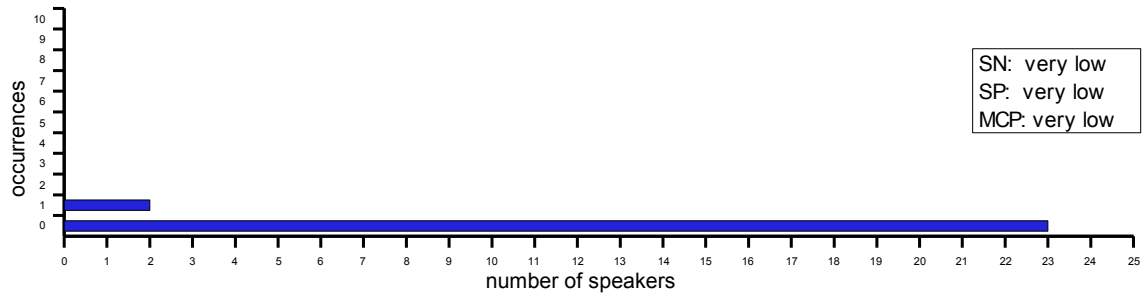


Figure 5.1: Speaker distribution pattern of very rare features

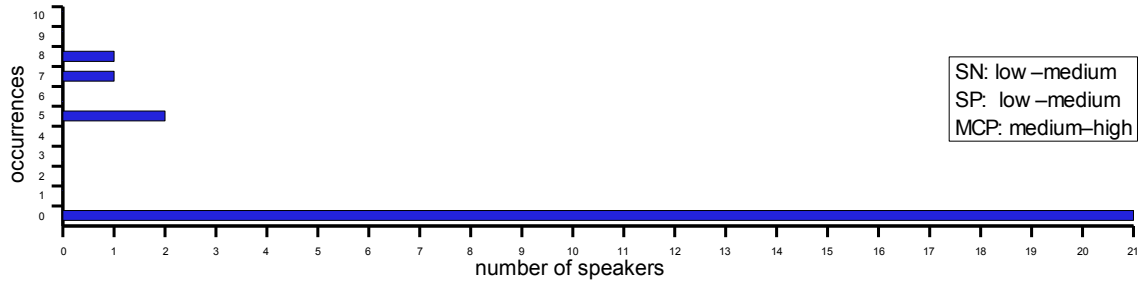


Figure 5.2: Speaker distribution pattern of speaker-specific features

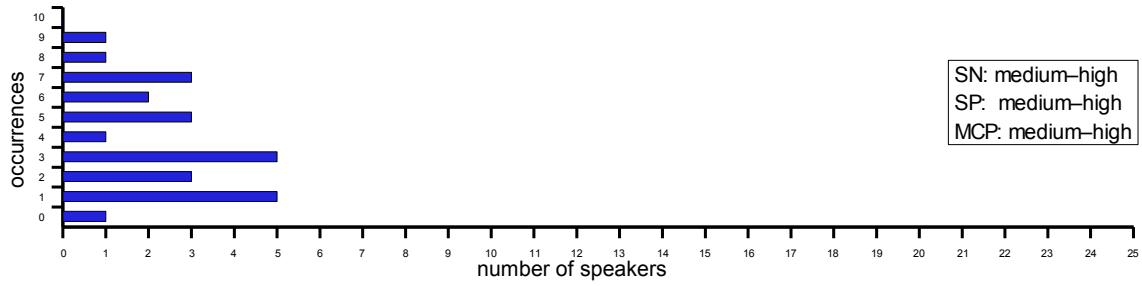


Figure 5.3: Speaker distribution pattern of generalised features

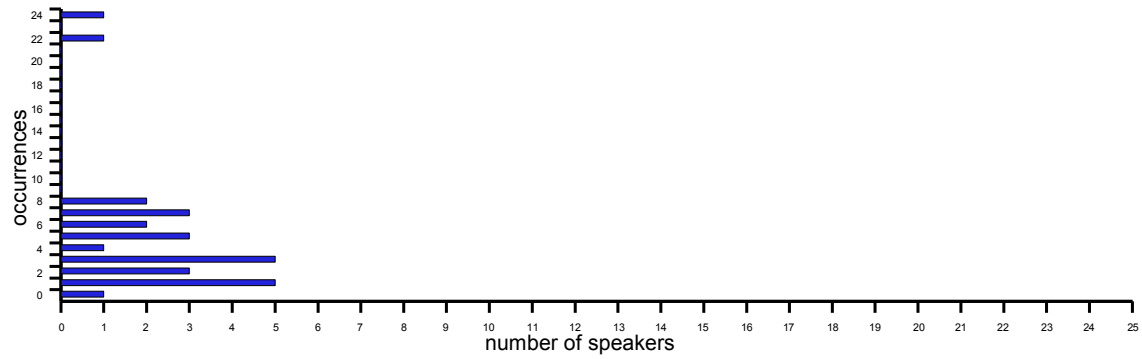


Figure 5.4: Speaker distribution pattern of a generalised feature with outliers

Chapter 6
Number and person features

In this first analysis chapter we will be looking at the formal realisation of number and person features in the corpus data. In terms of pronoun variability, number and person are the more robust categories compared to gender and case, which is why they occupy a relatively small part of the study. The low degree of variation can be attributed to the fact that number and person have the highest priority in the correct processing of pronominal expressions. However, it will be shown that some variation within these categories is possible, with no a negative effect on pronoun attribution, referent identification, or the overall understandability of the respective utterances.

The main non-standard phenomenon with regard to number is the use of plural forms with singular reference, which will be discussed in 6.1. Person is the most robust category of all, with even less variation than number. In the data, first person pronouns always refer to the speaker (plus one or more others in the plural), and second person pronouns always refer to the hearer (plus one or more others in the plural), except for indefinite *you*.

The usual correspondence between person and speech role can of course be violated intentionally. A speaker can, for example, exclude other conversation participants by referring to them in the third person. The deliberate violation of maxims of conversation and politeness, however, is not considered non-standard. Another phenomenon not included in the analysis are unintentional shifts in person. These are extremely rare and restricted to compatible pronominal expressions. In the corpus, some speakers were, for example, observed to switch between 1PL and 3PL pronouns, as seen in (42a), or between 1SG and 1PL, as seen in (42b). Similar switches can be attributed to a spontaneous change in perspective in the informal dialogue. In both examples, coreference relations remain traceable and the meaning of the individual pronouns remains clear.

(42) a. {Who looked after the garden? Did your dad or was he was it --} No, he was very good at that, my father was. He never bodged to bother us kids. **We kids**$_i$ used to have a little bit on **their**$_i$ own, little plot on **their**$_i$ own, just for **ourselves**$_i$ but eh only tiny plots, otherwise he done all the main. (pronoun switch, speaker KentAW1, KEN_009b.wav)

b. They were better days then than they are today, I ’ll tell you. **I**$_i$ ’m used to enjoy **ourselves**$_i$ with threepence more than they do now with thirty pound.
(pronoun switch, LAN002, LAN_002c.wav)

6.1 Plural forms with singular reference

6.1.1 Singular *us*

It is widely known that British English speakers occasionally use *us* instead of *me* to refer to themselves in informal conversations.[46] In earlier stages of English, the use of 1PL pronouns with singular reference usually carried connotations of authorship and majesty (compare 3.1.2; Levinson 1997: 70, 'editorial we'). Mitchell has described such occurrences as follows:

> OE poets used both *ic* and *we*. In most of those with *we*, the poet could be including his audience; there are similar examples in the prose. But we are left with a few sentences in both prose and poetry where we may have the 'plural of authorship'. (Mitchell 1985: 107)
>
> There are also places where a single individual other than an author seems to use the first person plural. […] the so-called 'plural of majesty'. (ibid.)

A different connotation which comes closer to the colloquial use known today is the expression of modesty (cf. Mitchell 1985: 251). In present-day English, the use of singular *us* in sentences like *Give us a kiss* is known all over the country (cf. Beal 1993: 205; 2004: 117; Trudgill 1999[2]: 88; Miller 1993: 108; Petyt 1985: 233; *Word Maps*, map 102 'with me').[47] Singular *us* makes requests, in particular, sound somehow more friendly and familiar. On the phonological level, the de-stressed pronoun, which is often pronounced /əs/, has weakened to the point of becoming enclitic, supporting the alleged modesty or reservedness of the speaker. (Note that this discourse-licensed usage is not captured by feature-geometric approaches; cf. Harley and Ritter 2002a: 507.)

While oral history interviews are not the type of data where we would expect to find many direct requests, such requests occasionally appear where speakers repeat conversations from the past. In addition to the extraction and coding procedures described in 4.2, supplementary word searches were conducted for all 1PL pronouns in combination with

46 The use of singular forms with plural reference, on the other hand, is extremely rare. Two occurrences were found in the data, one with *him* and one with emphatic *itself*: *...to become a pieceworker, or the professionals as we called* ***him****, one had to work for a number of years...* (SAL_032); *...they turned the tubs and pushed them along the road into the pipeworks* ***itself****.* (SAL_021)

47 SED informants who had to choose between *me* and *us* often chose *us*. For instance: Q VIII.8.12 *back us/ me up*; Q IX.8.2 *give me/ us*; Q IX.8.3 *given it me/ us*; Q IX.8.4 *come with me/ us*.

typical verbs of request (*will you*/ *can you*/... *bring*/ *get*/ *give*/...). The searches returned 11 cases, mostly requests, where the speaker used *us* to refer to himself or herself alone, as shown in examples (44) to (46). Only one example contains *we*, shown in (43). No examples were found for (royal) *ourself* with singular reference.

(43) When I come back there 's a big Yankee walking up and down. (v 'laughs') And **we**, Can I help you? Said, uh, I 'm the new manager. I said, You what? (singular *we*, CA, LND_001f.wav)

(44) ...and my mother said, Will you go down and **get us** a pound o' thin arrowroot biscuits? (singular *us*, YksSL, YKS_003b.wav)

(45) Yes, **bless us** what did they call him. (singular *us*, WesBX, WES_009)

(46) ...and, you see, there were several miners used to come in then and they would walk straight through. [...] There was one of them, he had one of the best dogs I 've ever seen in my life. He used to just open his coat out, Here 's a couple of rabbits, Fred, **give us** a couple of pints. (singular *us*, YksHS, YKS_006)

6.1.2 Third person plural forms with singular referent NPs

In the third personal plural, we find occasional examples where *they*, *them* and *themselves* are coreferential with a singular NP. Unlike colloquial singular *us*, these generic 3PL pronouns offer a gender-neutral alternative to *he*, *she* or *it*. They have a long-standing tradition in English and are well documented in the literature (e.g., Givón 2001: 435). Some accounts also include the morphological variant *themself*, which has come to be known as a colloquial reflexive with gender-neutral antecedents in sentences like *Someone has hurt themself* (no similar examples in the corpus). In present-day descriptive grammars, the alleged restriction of 3PL forms to plural antecedents has been abandoned:

> The use of *they* with a singular antecedent goes back to Middle English, and in spite of criticism since the earliest prescriptive grammars it has continued to be very common in informal style. In recent years it has gained greater acceptance in other styles as the use of purportedly sex-neutral *he* has declined. It is particularly common with such antecedents as *everyone*, *someone*, *no one*; indeed its use in examples like *No one*$_i$ *felt that they*$_i$ *had been misled* is so widespread that it can probably be regarded as stylistically neutral. (Huddleston and Pullum 2002: 493)

Feminist critics of the prevailing androcentrism in prescriptive grammars have observed with obvious relief that:

> despite almost two centuries of vigorous attempts to analyse and regulate it out of existence, singular *they* is alive and well. Its survival is all the more remarkable considering that the weight of virtually the entire educational and publishing establishment has been behind the attempt to eradicate it. (Bodine 1998[2]: 126)[48]

Generic pronouns are different from simple referential anaphora. Pinker (1994: 391–392) has even proposed to treat *they* as a homonym with two functions: (a) as plural referential expression, and (b) as indefinite generic non-referential anaphor (compare *They returned to their seats* vs. *Everyone returned to their seats*).

From a different perspective, 3PL forms with generic reference agree, in their own plurality, with the plurality which is implicit in the number-indeterminate referent NP. This holds for *they* with quantified antecedents, as seen in example (54), but it also holds for those *they* in the corpus which have collective referents, usually different types of food, tools or vessels, as seen in examples (47)–(53).

(47) If **they** were **a soft herrin'** like from up the north of Scotland, the Shetlands, you (reg sic=hefta) have to (/reg) give 'em a little bit more salt. (*they* with singular referent NP, EAA1, SFK_002)

(48) There was tons o' herrin' dumped up there, 'cause **they** were **a very soft, oily fish** and they wun't keep -- specially off the Fair Isle grounds. (*they* with singular referent NP, EAA1, SFK_002)

(49) **They** were **a nice little herrin'** off Shields. (*they* with singular referent NP, GCS, SFK_010)

(50) The Duke of York's, **they** were **a round potato**. (*they* with singular referent NP, ICS_RM, CON_010)

(51) ...'cause **they** was **a long, more like a shepherd's crook** with a, a knife at th' top end, shaped like, more like a shepherd's crook really... (*they* with singular referent NP, KentAW1, KEN_009a.wav)

In examples with collective referents, *they* has group reference even if this reference is encoded by a singular noun. The use of a 3PL pronoun in these examples can be attributed to the fact that the respective nouns – *fish*, *herring*, *potato*, *crook*, *boat* – represent collective referents despite the singular article. They can also be regarded as abbreviations of '*a*/ *the type of* + N' constructions, such as *a round [type of] potato*.

Some passages in the corpus show particularly flexible encodings which are possibly attributable to subtle changes in the speaker perspective. In (52), for example, several

48 According to Greenberg (who does not use the term himself), androcentrism is a universal property of human language: "Where masculine and feminine genders exist with or without further genders, the masculine usually appears to be the unmarked gender." (Greenberg 1966: 39)

changes in the plurality of the noun can be taken to reflect changes in the perceived individuality or genericness of the combined referents ('*a* and *b* were good boats' vs. '*a* and *b* were a good type of boat'). In this passage, the speaker shifts from plural reference (*good boats*/ *good old boats*) to singular reference (*they were a good sea boat*) and back to plural reference (*lovely boats*). Despite these switches, the central proposition remains clear: both boats were good sea boats.

(52) The Margaret and Sarah Hide were **good boats** in a blow, good boats. Yeah, good boats. **Good ol' boats**, **they** were. They had a good beam and **they** were **a good sea boat**. **They** were **lovely boats**. I liked the ol' Strive; I was in her, she was a good ol' boat. [...] The Margaret and the Sarah Hide had a Duffy's engine, twenty-two inch stroke. **They** were **a good engine**; oh yes, they 'd just tick over like a clock. (*they* with singular and plural referent NPs, ELF, SFK_004)

(53) The Courronne -- LT twenty-one her number was, I can remember that. I didn't know of any that foundered under a heavy load. They 'd carry some fish! **They were a good ol' sea boat**, you know. (*they* with singular referent NP, ELF, SFK_004)

(54) I think **each house**$_i$ had its wash-house to **themselves**$_i$. (*themselves* with singular quantified antecedent, IBCD, SAL_021)

Finally, 3PL pronouns with singular reference are also found in demonstrative determiner position. Generic and demonstrative uses of *they* and *them* are not usually discussed together in the literature, and they are also treated separately in the present study. Nevertheless, it should be noted that a definition of *they* and *them* as binary homonyms (as suggested by Pinker) is incomplete since it ignores their additional function as generic non-referential deictics. These occur with group nouns such as *lot*, *sort* or *type*,

(55) They were scranners. Ol' Jack Lay and ol' Teo and ol' Jasper and all **them ol' lot**. (3PL demonstrative before *lot*, EAA1, SFK_002)

(56) Oh, sums, and do dictation, and arithmetic, and all, all **they sort of things**... (3PL demonstrative before *sort*, TCA_EA, DEV_007)

(57) They called some boats white elephants -- the ol' Meg, the Ivy, the Covent Garden and **them sort o' boats**, you know, they never did earn a good lot o' money. (3PL demonstrative before *sort*, WT, SFK_012)

(58) Let 's take Calridges as an example -- they had **them Bermuda type**, four hundred horsepower. (3PL demonstrative before *type*, WT, SFK_012)

with singular count nouns,

(59) Well yes, and they bought it, **they house there** at the time. (3PL demonstrative before singular N, TCA_PR, DEV_011)

and, very rarely, also with mass nouns:

(60) It was eh what they call the tunnel pit, eh eh where you could walk down, but he used to

-- when he was younger he used to break, break **them ice** up... (3PL demonstrative before mass noun, Not A25, NTT_002)

(61) Course, a lot on ʻem were brass and if you didn't watch, ʻ**em brass** ʻd heat up. (3PL demonstrative before mass noun, JRS, SFK_011)

The use of *they* and *them* demonstratives will be discussed in more detail in chapter 15.

6.1.3 Areal distribution

In qualitative terms, the two topics discussed in this chapter – singular *us*, and 3PL forms with singular referent NPs – represent two distinct phenomena. This is supported by a notable difference in their geographical distribution. Even if there are not enough cases in the corpus for a statistically valid analysis, it is interesting to see that the majority of examples with singular *us* appear in data from the North (2 SE, 0 SW, 1 MID, 9 N). All 3PL forms with singular referent NPs, on the other hand, are from the South (7 SE, 1 SW, 0 MID, 0 N; excluding demonstrative function), a tendency which could be worth investigating in more detail.

6.2 Dummy objects: why not *walk it?*

English third person pronouns are typically used to refer to entities outside the speaker–hearer relation. Among these entities, neuter *it* covers the realm of inanimate objects (but consider the use of gendered pronouns, ch. 7). In addition, special non-referential uses of *it* entail a change in meaning in certain verbs, such as *to make it* ʻsucceed/ achieve/ manage to come' or *to get it* ʻunderstand/ comprehend'. When used as a non-referential impersonal dummy, *it* is neither a proper pronoun nor a proper syntactic object.[49]

In dictionaries and grammars, combinations of ʻV + dummy *it*' are usually listed as colloquial or informal uses of the corresponding verb. This section focuses on one combination which does not usually appear in dictionaries: ʻ*walk* + *it*'. Overall, there are 24 cases in the corpus, used by 11 different speakers (CD-ROM: 3SGn_walkit.xls). A closer look at the respective interviews shows that most of these speakers also used simple *walk*. The difference seems to be that *walk* is a simple motion verb whereas *walk it* can be

49 One could of course argue that the *it* in *get it* is referential (ʻget the idea/ the meaning/ what somebody said'), in which case *get it* would be synonymous with *understand it*. In the present study, however, *get it* is regarded as a colloquial non-referential alternative to *understand* (*Don't you get it?* ≙ *Don't you understand?*).

used to put additional stress on the fact that the person in question walked, for example instead of driving, or walked a long distance.

(62) And mi father then was livin' in Prescott, you see. Got a house here. And he had to walk back, to eh, **walked it**, no thought about jumpin' on a, well there was no, on a bus or aught like that, to Victoria Hotel [...] And mi father made excellent job at these, this work. And he **walked it** to Warrington. Told us often about this. And he **walked it** to Warrington... (*walk it*, LanEG, LAN_012e.wav and LAN_012f.wav)

(63) This place as we took over, it was an old bone yard, [...] we used to get down there on a tram car, a tram car from the centre of Wigan, but we had to **walk it** oftener than not to go to work. [...] I wish I 'd as many pennies now, I would have **walked it** from Wigan, I 'd got my fare, you know, t' get on t' tram. I used t' **walk it** down the Valley to Astley, Astley Pits, so that I could have a pint of beer coming back, **walk it** there and back so I could have a pint of beer at a pub in Tyldesley, it was fivepence a pint then and theirs were fourpence. (*walk it*, WigWHJ, LAN_020)

(64) But when we come to live down here we used to belong t' Saint Matthews and eh, we went for so long at Sunday morning even at eh it were eight o'clock Holy Communion. We 'd **walk it**, there were no buses running, we 're having to walk, and then **walk it** back again because they didn't start running while at ten. (*walk it*, Lang1P, LAN_003e.wav)

(65) {<u IntSH>And how would (trunc) y- (/trunc) how would you get from factory to factory, would you would you go by car or --} Around (trunc) wh- (/trunc) there was no cars then, you had to **walk it**. (*walk it*, NotA25, NTT_002a.wav)

Unlike other dummy objects with light verbs such as *make it*, *get it* or *take it*, the *it* in *walk it* does not effectuate a dramatic change in the meaning of the verb. Instead, it adds another meaningful layer to the sentence, similar to other pleonastic pronouns which will be discussed in this study later on (compare ch. 12, (464)).

Walk it is our first example in this study for pleonastic pronouns that can be attributed functions linked to emotive language use. Due to these functions, the *walk it* examples in the corpus differ clearly from thematically void dummies such as weather *it* (cf. Alba-Salas 2004, for lexically selected expletives). The additional *it* in sentences like (62)–(65), for example, is used to emphasise the strenuousness of covering long distances on foot, or to stress the fact that somebody walked the distance instead of driving. In both meanings, *walk it* could be paraphrased by something like 'walk all the way'. In some cases, the use of *walk it* in combination with heavy stress is a good indicator of emotivity, reflecting a meaning similar to 'Just imagine, he/ she walked all the way'. This can be heard very clearly in example (62), which has the corresponding soundfiles on the CD-ROM.

Unlike other dummy-object verbs such as *make it*, *walk it* seems to be perceived as highly colloquial (which explains its absence from English grammars and dictionaries). In 2007, this fact was exploited by the Transport For London company in a campaign entitled "Why not walk it?". The campaign was designed to encourage people to walk in the

London area, especially on their shorter journeys to work or school. Both the poster, shown in Figure 6.1, and the radio ads, shown in (66) and (67), make a light-hearted and friendly suggestion which is cleverly formulated in informal English. While most occurrences in the FRED corpus were found in data from the North (0 SW, 2 SE, 5 MID, 17 N), the London campaign shows that *walk it* enjoys supraregional popularity.

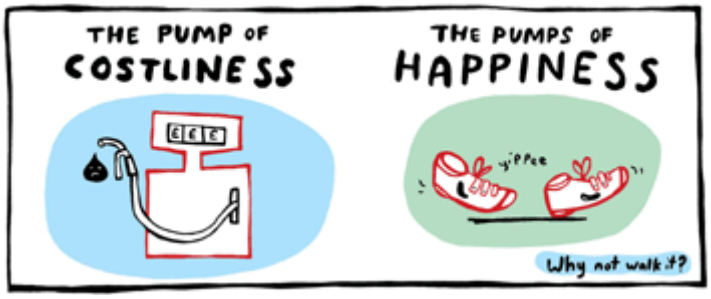


Figure 6.1: Transport for London "Why not walk it?" campaign (picture from TFL website at http://www.tfl.gov.uk/corporate/media/newscentre/archive/4939.aspx)

(66) TfL radio advert "School run"

Early morn, close the gate, bye bye house, can't be late.
Hurry up mum, see that cat? Ate Billy's budgy, that's why it's so fat.
Watch me climb up this tree, statues, the park, it's amazing what you see.
Over the bridge, down past the shop, not far to go, there's Mr Lollipop.
Hello School, it's half past eight, told you mum, we wouldn't be late.
Taking your kids to school by car just isn't the same.
The school run – **why not walk it?**

(Transport for London radio advert; transcript and soundfile retrievable from http://www.tfl.gov.uk/corporate/media/newscentre/archive/4939.aspx; my bold type)

(67) TfL radio advert "Work"

Wake up, Monday morning, leave for work, bleary eyes, half asleep.
Walk through park, funny tree, smiling dog, sun is shining, brain is waking.
Past the pond, hello swan, hello world, feeling better, feeling good.
Over the bridge, work in sight, nearly there, bring it on.
Going to work? **Why not walk it?**

(Transport for London radio advert; transcript and soundfile retrievable from http://www.tfl.gov.uk/corporate/media/newscentre/archive/4939.aspx; my bold type)

6.3 Existential *it*

A different phenomenon involving non-referential *it* is its use as a dummy subject in existential clauses, where it replaces standard *there*. As a non-pro-drop language, English uses *it* as a syntactic expletive in sentences which have no logical subject but nevertheless require an overt subject NP (e.g., weather *it*).[50] Similarly, presentational *there* followed by copula *be* functions as a dummy subject in existential clauses of the type *There is no coffee left*. In modern Standard English, the two expletives are not interchangeable, but occurrences of existential *it* can be observed in other varieties.

Existential *it* has not received much attention in the literature, probably due to its very low general frequency.[51] Kjellmer (2001) has investigated its use in modern, written British English phrases such as *it comes time to leave* (e.g., in *The Observer*), but the use of existential *it* plus copula *be* usually only gets noticed in earlier stages of English, or in connection with American vernaculars. Visser (1963: 42) has observed that "[t]he type '*It* is no man (who) can discourage me' was formerly quite common. It remained so in archaic ballad style." According to Visser, existential *it* plus copula *be* continues to be used in African American Vernacular English and among the white working-class population of the southern United States. This coincides with the *American Heritage Dictionary of the English Language*, where we find the following description:

> Existential *it* is hardly a recent innovation – it appears in Middle English; in Elizabethan English, as in Marlowe's *Edward II*: "Cousin, it is no dealing with him now"; and in modern American literature as well. Although most British and American varieties no longer have this historical feature, it still occurs in some Southern-based dialects and in African American Vernacular English. Use of existential *it* may actually be increasing in some places, such as Smith Island, Maryland, a historically isolated community. While older Smith Islanders sometimes use existential *it* rather than *there*, younger islanders almost always do. (Pickett 2000[4]: 929)

50 Bolinger (1977) and Chomsky (1981) argue in favour of an interpretation of weather *it* as verb argument.

51 More attention has been given to *that* used instead of *it*. In the SED, for example, this is a characteristic feature of central and eastern Norfolk (incidental material 21,8). Trudgill (2004a: 146) has mentioned the use of *that* for stressed *it* in East Anglia in sentences like *That's raining*, *That's cold in here* or *That was me what done it*. He has also mentioned the use of dummy *that* in the traditional dialects of Eastern England, e.g. in the *Essex Ballads* "When tha's wet the corn git läid" (Trudgill 1999[2]: 89), and in the modern varieties of East Anglia and the South Midlands, in sentences like *I don't like it when that's raining*. According to Trudgill (2004a: 147), the phenomenon can be explained from a grammaticalisation perspective: "Diessel (2000) shows that demonstratives very frequently become third-person pronouns as a result of grammaticalisation. The fact that *it* is most usually pronounced as a possibly rather indistinct [əʔ] may have assisted this process."

Green (2002) has described the use of existential *it* in AAVE in more detail. According to her, the pronoun combines with three different linking verbs or 'linkers': *be* (inflected or aspectual), *have* and *got*. For the sentence *There is some coffee in the kitchen*, for example, Green lists the following alternatives: *It's some coffee in the kitchen* (where 'it's' is pronounced /Is/), *It got some coffee in the kitchen* and *It have some coffee in the kitchen* (p. 80). Green emphasises that the phenomenon is systematic, existential sentences in this variety always presenting themselves in the form 'existential element – linker – logical subject' (p. 81). Speakers are obviously able to distinguish the different uses of '*it* + *be*': "they understand when it is being used as a pronoun and as an existential element" (p. 82).

The FRED data show that existential *it* plus *be* has not actually died out in spoken British English. Even if it is not among the most frequent dialectal features in the corpus (13 occurrences), there are enough examples to show that it is not limited to one individual speaker or geographic area. An exhaustive search for this particular phenomenon is still pending, which is why only a few examples will be presented. Some of them show existential *it is* where Standard English requires *there is*, as in (68)–(71). In most examples, however, the *it* forms part of a question tag, as seen in (72)–(75) (9 out of 13 cases; compare 3SGn_qtags.xls, CD-ROM).

In question tags, it can be difficult to decide whether *it* simply replaces *there* – which in (72)–(74) precedes it in the immediate co-text – or whether we are dealing with a generic tag. What is certain is that the question tags seen below show the same loss of referentiality as other generic tags. It seems likely, therefore, that the occurrence of existential *it* in these examples is linked to the spreading of generic *it* in question tags as discussed in the next section.

(68) But uh, in my opinion, **it's**, you know, so much chemical on the, on the ground today. (existential *it*, WS, KEN_006a.wav)

(69) {<u IntER> And she worked after she had the children?}
<u LanD1P> Ah, '**t is** only me.
{<u IntER> Oh I see, just you.}
<u LanD1P> Ah, '**t is** only me. That's all. (existential *it*, LanD1P, LAN_006f.wav)

(70) ...small little engine, but you 'd got to keep plenty of steam and water because it was six hundred and eighty foot above sea-level top, up the top -- and **it is** beautiful scenery going through. (existential *it*, IBBeBo, SAL_011)

(71) And it 's still and it 's still there in Church Street here, and **it is** about six shops all in a row... (existential *it*, Lang1P, LAN_003f.wav)

(72) Same with milk, you used to get your milk from the farms and that, there was no delivery then, you see. No paper deliveries anyway, **was it**? (existential *it* in question tag, FP, OXF_001c.wav)

(73) Well then there 's a hedge and field there and the road then, **isn't it**. (existential *it* in question tag, SRLM_GG, SOM_013)

(74) Well, there 's only Neville who 's got two, **isn't it**? (existential *it* in question tag, ICS_MW, CON_009)

(75) ...it had been a shop sometime or other, and down the bottom was a saw mill, (reg sic=**wunnit**) weren't it (/reg) (existential *it* in question tag, JG, LND_004f.wav)

6.4 Generic question tags – *innit?*

One of the best-known dialect features of spoken English is certainly the use of *innit*, a contracted form of *isn't it* frequently used in question tags. Besides *innit*, we also find contracted *wunnit*, which replaces *wasn't it*, *weren't it* or *wouldn't it*. *Innit* and *wunnit* form part of a larger group of generic question tags which can, but do not have to, be contracted.

The most basic characteristic of generic q-tags – alternatively 'independent' or 'autonomous' q-tags – is that they do not need to be in agreement with the verb and noun of the preceding sentence. Their independence makes them comparable to invariant tags in other languages. While Standard English, with its variable q-tags, differs from many other European languages, generic tags like *innit* are very similar to q-tags in German (*nicht* 'not', *nicht wahr* 'not true', *oder* 'or', *gell* or *gelt* in Swiss and Southern German), Spanish (*verdad* 'true', *no* 'no'), or French (*n'est-ce pas* 'is it not').

Functionally and formally, the genericness affects both parts of the q-tag. First of all, the verb no longer echoes the main sentence verb (e.g., *didn't* in (86)), and, at the same time, it no longer reflects its own original meaning. Generic *is* or *isn't*, for example, no longer reflect the original meaning of *be*. Furthermore, the pronoun *it* acts as an impersonal q-tag particle, showing loss of coreferentiality with the preceding NP, and even loss of referentiality altogether. In contracted expressions like *innit*, the change in meaning is accompanied by a change in the morphology and pronunciation – through lexicalisation and phonological erosion – which strengthens the perception of the tag as one unit. The fact that the same changes can be observed in *wunnit* indicates that we are dealing with a wider linguistic phenomenon which could theoretically spread to other 'V + pronoun' combinations.

In the FRED corpus, 24 instances of *innit* were found in different dialect areas (5 SE, 17 SW, 2 N), showing that the phenomenon, which is often associated with Cockney slang, is not restricted to the Southeast.

Table 6.1: 3SGn question tags

tag	generic/ non-generic (generic%)
innit/ inni'	9/ 24 (37.5%)
isn't it	13/ 138 (9.4%)
is it	3/ 20 (15.0%)
wunnit	1/ 7 (14.3%)
weren't it	1/ 11 (9.1%)
wasn't it	16/ 121 (13.2%)
was it	10/ 34 (29.4%)
wouldn't it	2/ 26 (7.7%)
didn't it	3/ 3 (100%)
TOTAL	58/ 384 (15.1%)

The figures presented in Table 6.1[52] suggest that the lexicalisation process from *isn't it* to *innit* may be attributed to frequency factors. This has also been suggested by Krug (1998: 304), who found in his data that "[b]oth *isn't* and *it* clearly dominate their respective operator and subject categories of the negated English tag paradigm. Most importantly, however, the tag question *isn't it?* alone accounts for more than one third of all negated tag questions." Krug's frequencies are surprisingly similar to the results obtained in FRED. In addition, the argument that high string frequency is an important cognitive motivator in language change (especially coalescence) can be extended to the second most frequent invariant q-tag, *wunnit*. Consider the high string frequency of *wasn't it*, which comes very close to that of *isn't it*, even more if we include the other uncontracted q-tags starting in *w*-.

In terms of agreement, *innit* is not automatically generic. In many cases, it represents a simple contraction of *isn't it*, as seen in example (76). In its generic function, however, *innit* can be used irrespective of the preceding verb ((77), (78)) and verb tense ((79), (80)), and irrespective of the preceding subject NP, as shown in (81). *Innit* is not the only generic q-tag in the corpus. Other examples include uncontracted forms like *isn't it*, *wasn't it* or *weren't it*,[53] as well as positive *is it* and *was it* (also in the examples and 3SGn_qtags.xls). Overall, generic q-tags with *be* are certainly the most common, but other auxiliaries also feature in the data, for example *wouldn't* and *didn't*. Table 6.1 shows the proportion of generic cases for each of these verbs.

52 Table 6.1 does not include invariant *no*, which was found twice in the corpus: *All sorts they used to do, no?*, CON_003, and *No compensation, no?*, NTT_013. Also excluded was *eh*, as in *So, oh, half a crown, eh?*, LND_003, or *You see the size of them, eh?*, DUR_001. For question tags in Scottish English, compare Miller (2004).

53 The singular use of *weren't* is not considered here, e.g. *Which was a big price, weren't it?*, SOM_007.

The relatively high generic proportion of uncontracted tags shows that formal contraction and genericness are not co-dependent. It appears, however, that "the load of informational content is inversely proportional to the likelihood of undergoing contraction" (Krug 1998: 310).

(76) Well if I was a bit early and I could hear him preaching the sermon I used to think miself, well there 'll be a hymn to finish up with, I used to walk around the churchyard and read the headstones (v 'laughs'), [that 's a] queer occupation **innit**? But it used to pass the time... (*innit* q-tag after verb elision, non-generic, TCA_WH, DEV_009d.wav)

(77) Well Alice used to, you know, I used to walk to Woodstock at many a time with eh, Woodstock when they had Woodstock gloves, **innit**, yes and made them. (generic *innit* 'didn't they'/ 'didn't I', FP, OXF_001e.wav)

(78) Salty Tommy we called him, **wasn't it**? Old donkey cart he used to come round with hm. (generic *wasn't it*, FP, OXF_001f.wav, trba.)

(79) {<u IntRS> Yes, when he kept the marine store, that was, hmm.} Yes, 'course that 's [that was] 'fore your time, **innit**? (generic *innit* 'wasn't it', CAVA_PV, CON_003b.wav)

(80) Well, when I started I had eighty-four, when I started. 't was eighty-four all here, **innit**. (generic *innit* 'wasn't it'/ 'wasn't there', SRLM_JA, SOM_029)

(81) That young lady that was with you the other day said that she was dead before they put her in there because she said her lungs eh {<u Int> There was no water in them. No.} There was no water in them. And had she been drowned they would have been full of water, her lungs would have been full of water, **wouldn't it**? (*wouldn't it* 'wouldn't they', TCA_RA, DEV_001d.wav)

(82) Someone or the other, they did get some of these names, **innit**? (generic *innit* 'didn't they', FP, OXF_001d.wav)

(83) They 'd be worth a fortune, you know, if you get folk wanted to buy this, **wouldn't it**? (*wouldn't it* 'wouldn't they', TrbrEGC, WIL_015)

(84) It 's a nice village I think, you know, people very friendly you know but o' course they keep, they change according to their occupation **isn't it**, they move out see and that sort of thing, you can't help that. (*isn't it* 'don't they', TrbrKM, WIL_019)

(85) Well there wasn't any buses, there wasn't any trains was there? I mean it were years and years after when trains used to run from Burton Joyce, **didn't it**? (*didn't it* 'wasn't/ weren't it', NotA80, NTT_009a.wav)

(86) ...he was a baker you see, made his own bread and cakes and all that, and same with this one along here **didn't it**? In the village they made their own bread, and that, we used to make our own bread, **didn't it**, as well. (*didn't it* 'wasn't it/ he'/ 'didn't we', FP, OXF_001a.wav)

(87) I took my dog in and eh, well it 'd be last week **wasn't it**? (*wasn't it* 'wouldn't it', A109w, NTT_015)

(88) {<u IntSH> What age, how old were you then?} I was eh twelve, **was it**? No, no, I should be about eleven I suppose. (*was it* 'wasn't I', NotA52, NTT_005b.wav)

Based on the above-mentioned typological parallels between invariant q-tags in spoken English and invariant q-tags in languages like German, Spanish or French, it is to be

expected that missing agreement between sentence and q-tag affects neither the understandability of the sentence, nor the function of the q-tag itself. The underlying reason is that q-tags are discourse-structuring devices which are void of propositional meaning. Stein (1997) has adverted to this fact in the following description:

> Finally, tags express meanings that are associated with oral language. They invite or enlist the hearer's presence and participation in endorsing the propositional content expressed. [...] In fact, to the exclusion of aspectual – and grammaticalized – meanings they often do carry discourse-structuring meanings and act as ploys to move constituents into preferred information positions. To the latter extent they are indeed propositionally empty. (Stein 1997: 43)

The fact that q-tags have no propositional meaning also explains why it is possible for speakers in spontaneous conversations to use positive–negative sequences that are considered incorrect in Standard English (*it is..., is it?*). The respective utterances remain perfectly understandable. If q-tags are propositionally empty, the Standard English requirement that a positive verb must be followed by a negative tag, and vice versa, can be regarded as a mere formality. Various empirical examples from the corpus contravene the standard rule. Interestingly, however, they all contain positive–positive sequences (17 cases), as seen in examples (89) and (90). Negative–negative sequences do not occur. A logical explanation for this restriction remains to be found.

(89) And about a year after that my brother was walking over to uh Queen's Woods over at uh Alexandra Palace, making a bit of a noise. (trunc) Pu- (/trunc), copper pulls him up, says, Oh, (gap 'indistinct'), he says, **It's** your brother who got away with that case last year, **is it**? (*is it* 'isn't it', CA, LND_001g.wav, trba.)

(90) I remembers when they first started round here, you know, they -- that was just after the, I believe **that was** just after the War, **was it**? (*was it* 'wasn't it', FP, OXF_001)

Generic q-tags, and positive q-tags which follow a positive verb, both violate the standard rules, but they retain the intrinsic function of utterance-final signals which is common to all q-tags in all languages. Just like standard q-tags, they superficially question the proposition of the preceding sentence. In doing so, they can be used to express doubt or uncertainty, elicit a response from the hearer, or for other, more subtle pragmatic functions, such as the softening or intensifying of potentially face-threatening statements (cf. Holmes 1995: 82).

In her book *Men, Women and Politeness* (1995), Holmes finds that some discourse functions can be marked by intonation: so-called epistemic modal tags expressing

uncertainty generally have rising tone (*The restaurant is this way, isn't it?*) whereas so-called facilitative tags encouraging the addressee's participation generally have falling tone (*You've been there, haven't you?*). The additional importance of pauses is illustrated by the FRED example in (91). Here, the ambiguous tag in (a) is disambiguated by the insertion of speech pauses in (b) (generic *wasn't it* instead of *didn't I*) and (c) (standard *wasn't it* after verb elision; as found in the corpus). The difference in meaning between (b) and (c) is not encoded in the sentence-final tag but in the prosody, by means of pausing and intonation (compare the soundfile).

(91) a. Then I got a job in Totnes after I met you wasn't it?
b. Then I got a job in Totnes after I met you (pause) wasn't it?
c. Then I got a job in Totnes (pause) after I met you wasn't it? (DEV_008f.wav)

Together with the findings of Holmes and the typological evidence for invariant q-tags in other European languages, several examples from FRED suggest that prosodic aspects have a higher importance in the correct processing of q-tags than formal agreement. If ambiguity arises, for example due to unclear clause boundaries, prosodic means like intonation and pausing facilitate the correct assignment. The same means, however, are also used with standard q-tags. From a functional point of view, therefore, non-standard q-tags work exactly the same way as 'correct' q-tags. Formal variation in the verb or pronoun, or in the overall morphology of the tag, has no negative effect on its interpretability or the understandability of the sentence as a whole.

The future development of generic q-tags is uncertain. It has been suggested that *innit*, the most widely studied tag in British English varieties, could one day become part of the standard grammar (cf. Krug 1998: 304). However, the invariant tag is still heavily stigmatised as a marker of 'uneducated' speech, which may prevent its spreading to more formal registers.

6.5 Summary

The two morphosyntactic categories of number and person have high priority in the correct processing of pronominal expressions. In the corpus data, this is reflected in a relatively low degree of variation, especially as compared to variation in case (chapters 8–15). Nevertheless, it could be shown that when variation occurs in these categories, it has no negative effect on pronoun attribution or referent identification.

Variation in number has different motivators, which are mostly pragmatic. The use of *us* instead of *me*, for example, carries certain connotations of modesty and closeness which can be used to make a request sound affectionate and less demanding. The use of *they* and other 3PL forms, on the other hand, offers a gender-neutral alternative to third-person singular reference. In their own plurality, 3PL forms agree with the plurality implicit in indefinite antecedents (*Nobody likes paying their taxes*), quantified antecedents (*Each house had its wash-house to themselves*) and collective referents (*they were a soft herring/ round potato/ good sea boat*).

Variation in person is restricted to impersonal, non-referential uses of *it* in the corpus. Here, too, non-standard uses could not be shown to have any negative impact on the understandability of the respective utterances. On the contrary, the phenomena discussed in sections 6.2 to 6.4 can be attributed additional pragmatic functions. This was first illustrated by the use of dummy *it* in *walk it* – one of many pleonastic pronouns expressing emotivity in informal speech.

Loss of referentiality was also observed in generic question tags. With its variable tags, Standard English sets itself apart from other European languages whereas non-standard invariant tags like *innit* show convergence with the predominant typological trend. It was argued that question tags, both standard and non-standard, are propositionally empty, and that they are, therefore, not affected by variation in either the pronoun or the verb. Even generic tags like *innit* and *wunnit* retain the intrinsic discourse functions of all question tags. The exact function of a specific tag is independent of its formal realisation. It has to be derived from the overall meaning of the sentence and the conversational context, with the additional help of prosodic cues (*they got some of these names, innit/ that was just after the War, was it/ Then I got a job in Totnes (pause) after I met you wasn't it?*). The same is true for other phenomena involving non-referential *it*, such as the use of existential *it* instead of *there* (*it's so much chemical on the ground today/ it is about six shops all in a row*).

Chapter 7
Gendered pronouns

Modern Standard English is a natural gender language (cf. Corbett 1991; Curzan 2003; Greenbaum and Quirk 1990). Male referents are referred to by masculine pronouns *he* and *him*, female referents are referred to by feminine pronouns *she* and *her*, and referents without an assigned gender are referred to by neuter *it*. The present chapter investigates occurrences which violate these rules, i.e. masculine and feminine pronoun forms with neuter reference. Pronouns of this type are generally known as *gendered pronouns*.

Natural gender systems are typologically widespread and go back as far as Indo-European (cf. Szemerényi 1996: 155–156). Tieken-Boon van Ostade (1994: 236) has observed that "grammatical gender is a concept that has to be learnt, and that is acquired considerable time after children have learnt to distinguish between the sexes." This explains why, even in languages with grammatical gender assignment, natural gender emerges as an alternative in linguistic contexts with an increased processing load. Mühlhäusler and Harré (1990: 72), for example, have noted the transition in German "from grammatical to natural marking as the distance between a pronoun and its antecedent widens." But how can we explain the opposite phenomenon, the persistence of gendered reference in a natural gender language? In pursuit of this question, the FRED data will be used to shed light on distributional tendencies in quantitative, qualitative and geographical terms. It will be shown that spontaneous conversations do not always follow the standard attribution rules, and that the observed deviations follow some regularities of their own.

The investigation will be guided by the following two questions: are gendered pronouns a frequent or rare phenomenon, and what are the determinants of variation between gendered forms and *it*? In order to answer these questions, a variety of different aspects will be taken into account, including the role of referent categories, the importance of the speaker viewpoint, the relationship between gendered pronouns and topicality, the distribution of gendered pronouns in different syntactic positions, regional differences, and differences between male and female speakers.

Every dog has his day.
(proverb)

7.1 Gendered pronouns in the literature

English gendered pronouns are usually discussed in connection with one particular phenomenon, the traditional count–mass system found in the Southwest of England. Instead of following the rules of natural gender assignment, this system uses gendered pronouns (usually masculine) to refer to count nouns or 'personal class', and *it* to refer to mass nouns or 'impersonal class' (cf. Hughes and Trudgill 1996[3]: 32; Trudgill 1999[2]: 95; Upton et al. 1994: 486). Systems of this type are typologically very rare (cf. Kortmann 2002: 202–203; Wagner 2004c: 482), but a similar development has recently been observed for Modern Dutch. Audring (2006) has pointed out the resemantisation of Dutch pronouns along a count–mass line, count nouns being increasingly referred to by masculine *hij* and *hem* while neuter *het* is being used for mass reference.[54]

The traditional English West Country phenomenon has been investigated in great detail by Wagner (2004a; 2004c), and more recently by Siemund (2008), but accounts go back as far as the 18th century (e.g., Marshall 1789; Barnes 1886; Ellis 1869–1889; Kruisinga 1905; Ihalainen 1985; Wakelin 1986; Upton et al. 1994, SED findings). In his *Provincialisms of the Vale of Glocester*, Marshall (1789: 56) described how *he* was used "almost invariably for *it*; all things inanimate being of masculine gender." Almost a century later, Halliwell (1887, Vol. I, xviii) stated that "[i]t is a common saying, that in Hampshire every thing is called *he* except a tomcat which is called *she*." And Lowsley (1888: 6) mentioned the use of 3SGn subject forms *ut*, *he* and *a*, and object forms *'e*, *'in* and *un* in the Berkshire dialect. Around the same time, Elworthy made the following observation in his *Outline of the Grammar of the Dialect of West Somerset*:

> Every class or definite noun, i.e. the name of a thing or object which has a shape of its own, whether alive or dead, is either masculine or feminine, but nearly always the former; indeed, the feminine pronouns may be taken as used only with respect to persons. For instance, in chaffering for a sow, it would be said , *Wuul, neef tez· u zuw, ee ul git au·n*, 'Well, if it is a sow, *he* will get on', i.e. get fat. [...] *It* is simply an impersonal or abstract pronoun, used to express either an action or a noun of the undefined sort, as cloth in the quantity, water, snow, air, etc. (Elworthy 1877: 32–33)

54 In addition, De Vogelaer and De Sutter (in press) have observed the development of *het* into a default pronoun "that can be used whenever grammatical gender agreement fails."

Towards the end of the 20th century, however, the situation already presented itself very differently. Ihalainen (1985: 158), for example, found that "the correct generalization today seems to be that *it* can be used for 'thing' and 'mass' referents, whereas the personal forms do not occur with 'mass' referents at all." It is of course true, as Wagner (2004a: 16, fn. 8) has pointed out, that the comparability of early accounts like Marshall or Elworthy with modern, statistically analysable datasets is by no means guaranteed. This includes the SED fieldworker notebooks, where we mostly find notes on exceptional uses. Nevertheless, a development is recognisable. In search for the traditional West Country phenomenon, the most recent studies by Wagner and Siemund reach the same conclusion, namely that the old, categorical count–mass system is disappearing:

> Yes, the traditional system is slowly dying out, witnessed by fewer and fewer masculine forms in those domains which at one time were their exclusive territory. The formerly obligatory system has developed into an optional system. (Wagner 2004a: 292; based on data from the Southwest of England and Newfoundland)

> The results yielded by my search through these sub-corpora are rather sobering. The handful of examples showing an animate pronoun used for picking out an inanimate object is a far cry from the highly regular usage of masculine pronouns for the domain of inanimates that was widespread use only a century before. (Siemund 2008: 60; based on the BNC spoken component)

7.2 Gendered pronouns in FRED

The following subsections present a qualitative and quantitative analysis of gendered reference in the FRED corpus. It will be shown that, while most cases correspond to Ihalainen's above-mentioned generalisation, gendered pronouns occasionally still have mass reference, and that the semantic field of the gendered referent plays a crucial role. It will be argued that explanations based on a count–mass distinction are insufficient for the distributional tendencies observed in the data. Rather, most gendered occurrences point towards a pragmatic system which revolves around the speaker viewpoint. This includes the referent's topicality, as well as the speaker's personal involvement and emotional attachment to the referent. The emotive aspect explains why certain groups of referents such as 'pets' have a stronger influence on gendered pronouns than other groups of

referents such as 'inanimate objects'.[55] The emotive aspect also explains why male and female speakers differ in their use of gendered pronouns for certain referents, for example 'ships'.

The referent's degree of individuation will also be taken into account. Over the last two decades, individuation scales have been attributed a prominent role in the description of gendered pronouns. I argue that individuation is not a referent property but epiphenomenal to the speaker viewpoint. The degree to which we perceive and express something as an identifiable individual entity depends on a variety of subjective factors. These factors may correlate with certain referent properties, but they form part of the speaker perspective.

Last but not least, it will be shown that gendered pronouns are a supraregional phenomenon which is by no means confined to the Southwest of England. The distribution in the data does, however, suggest a South–North continuum, the southern end of which bears resemblance but is not identical with the traditional West Country system known from the literature.

7.2.1 Quantitative distribution

Early references in the literature give the impression that gendered forms, and masculine forms in particular, were once used categorically with all count nouns, whereas more recent studies like Wagner (2004a) and Siemund (2008) suggest that the phenomenon has become optional and rather rare, especially with inanimate referents.

The question whether gendered pronouns are frequent or rare can be answered in two ways. Compared to early descriptions from the 18th and 19th centuries, gendered pronouns in present-day data seem rather infrequent. Putting the present-day results into perspective, however, we find that gendered pronouns range between middle and high compared to other non-standard phenomena in the same dataset. Tables 7.1 and 7.2 show that the average proportion of gendered forms in the corpus is larger than 60% for non-human animate referents, and larger than 9% for inanimate objects. In addition, the results show that, while masculine forms predominate as expected, a substantial amount of feminine forms also appears in the data. A restriction to masculine forms is not supported, limiting the comparability with earlier accounts.

55 According to Corbett (1991: 12–13), "careful observation in the right type of setting will reveal that the gender system of English, though certainly based on semantics, is not totally dependent on the straightforward criteria of humanness and biological sex, but may be affected by pragmatic factors."

Table 7.1: Gendered pronoun frequencies: animate referents

(m = masculine forms; f = feminine forms; n = neuter forms)

dialect area	animate human				animate non-human			
	m	f	n	TOTAL	m	f	n	TOTAL
SE	2694 79.2%	708 20.8%	0 0.0%	3402 100%	39 27.7%	5 3.5%	97 68.8%	141 100%
SW	1226 83.7%	228 15.6%	11 0.8%	1465 100%	28 28.3%	14 14.1%	57 57.6%	99 100%
MID	1514 82.0%	316 17.1%	16 0.9%	1846 100%	6 15.8%	0 0.0%	32 84.2%	38 100%
N	1739 74.9%	583 25.1%	0 0.0%	2322 100%	7 4.9%	2 1.4%	134 93.7%	143 100%
TOTAL	7173 79.4%	1835 20.3%	27 0.3%	9035 100%	80 19.0%	21 5.0%	320 76.0%	421 100%

Table 7.2: Gendered pronoun frequencies: inanimate referents

(m = masculine forms; f = feminine forms; n = neuter forms)

dialect area	inanimate count				inanimate mass			
	m	f	n	TOTAL	m	f	n	TOTAL
SE	53 2.7%	82 4.2%	1797 93.0%	1932 100%	0 0.0%	0 0.0%	534 100%	534 100%
SW	46 5.2%	1 0.1%	838 94.7%	885 100%	4 1.7%	0 0.0%	226 98.3%	230 100%
MID	42 3.3%	2 0.2%	1228 96.5%	1272 100%	0 0.0%	0 0.0%	339 100%	339 100%
N	5 0.3%	4 0.2%	1805 99.5%	1814 100%	0 0.0%	0 0.0%	290 100%	290 100%
TOTAL	146 2.5%	89 1.5%	5668 96.0%	5903 100%	4 0.3%	0 0.0%	1389 99.7%	1393 100%

Tables 7.1 and 7.2 are based on a representative text sample consisting of 10 randomly selected interviews from each dialect area (CD-ROM: Spreadsheets/Analyses/Ch7/ TextSample). Each occurrence in this sample was assigned to one of the four referent categories represented in individual blocks: 'animate human', 'animate non-human', 'inanimate count' and 'inanimate mass/ abstract'. Results for *it* were extrapolated from random samples of 100 per dialect area, excluding impersonal *it* (*it was raining*/ *it was only me*/ *it's the truth*), general *it all*, and set expressions like *damn it* (see sample400.xls).

Dead animals and parts of plants or bodies were included in the inanimate category (e.g., *smoked fish*, *tree bark*). It should also be noted that the results do not include gendered possessives, as in *keep the butter to his shape* (CON_011), or gender-neutral pronouns. In the corpus, no instances were found of either gender-neutral feminine forms or non-referential occurrences such as *there she goes*. Gender-neutral masculine forms are rare. This is probably due to the fact that the stories and events told in FRED refer to familiar individuals whose sex is usually known to the speaker. Gender-neutral masculine reference was, for example, found in the following passage: *You get a working class person [...] works himself up, and he gets to that position, he 'll do anything to keep himself there, or he 'll do anything to make himself better* (LAN_012). Details on the use of gender-neutral pronouns in the history of English can, for example, be found in Mitchell (1985) and Baron (1986).

7.2.2 Referent categories, topicality, and the speaker viewpoint

Previous studies have shown that gendered referents pertain to certain semantic categories, which means that gendered frequencies will to a large extent depend on the conversation topic. Certain referents or groups of referents have been associated with the phenomenon for a long time. In his *English Grammar* from 1640, for example, Ben Jonson classified *ships* as feminine, and *dogs* and *horses* as generally masculine. Jonson described the neuter gender as "feigned gender: whose notion conceives neither Sexe; under which are compriz'd all inanimate things, a ship excepted" (p. 57, quoted in Nevalainen and Raumolin-Brunberg 1994: 183).

More recently, it has been observed that gender choice appears to be moving away from intra-linguistic assignment rules and towards extra-linguistic conditioning factors, especially the speaker's emotional attitude towards the referent (cf. Wagner 2004a: 294). This view is clearly supported by occurrences from the FRED corpus. Here, it seems to be the viewpoint and attitude of the speaker, rather than the referent properties themselves, which favour the use of gendered reference. Even if "cats are more likely to be *shes* generically, based on the biological-semantic pattern (*dog* = neuter or +male, vs. *bitch*; *cat* = neuter or +female, vs. *tom-cat*)" (Wagner 2004a: 124), it is the individual speaker's perception that 'cats are female' which will lead him or her to use a feminine pronoun when referring to the animal. Similar to the individuation aspect, the biological-semantic classification of the referent is, therefore, in many cases epiphenomenal to the speaker's

viewpoint. I argue that the distinction between objectively verifiable referent properties and the subjective perception of these properties is subtle but crucial. While the former can be used to identify common gendered referents such as ships, it is the latter that will tip the scales in the concrete speech situation and decide whether a specific ship is *it* or *she*.

Referent categories – masculine forms

In FRED, gendered reference covers a surprisingly great variety of non-human animate and inanimate entities which are by no means restricted to the typical 'horses, dogs and ships'. The masculine categories can be summarised as follows:

- animals and plants;
- vessels and vehicles;
- machines and tools;
- buildings;
- food and drink;
- recipients;
- other objects.

Very rarely we find examples from the semantic field of 'events' (e.g. *war* in DEV_005a.wav). Most gendered forms in the conversations refer to animals: bears and bullocks, calves, colts, cows, dogs and donkeys, fish and fox, elephants, heifers and horses, lambs, mares, moles, pigs, sows, ponies, rabbits, sheep, sharks, squirrels. It is primarily masculine forms which account for the high frequency of gendered reference in the 'animate non-human' group in Table 7.1.[56] The corpus results support Wagner's findings:

> Although most grammars of modern and earlier stages of English tell us that the appropriate pronoun to use when referring to an animal is *it*, except for cases where the sex of the animal is known, actual language use could not be further removed from this prescriptive statement. (Wagner 2004a: 121)

Familiarity with the animal in question is not mandatory. Gendered forms can, however, be used to mark the one referent which represents the central topic or focus of a discourse segment, as seen in examples (92) and (93):

56 Speakers occasionally use masculine forms for non-human animates that are readily associated with female sex, such as the *sow* in example (93). However, the corpus data show no regular masculinisation of anaphoric pronouns. Masculine references to female referents are the exception rather than the rule. Compare Curzan (2003: 93) for similar results in Old English data.

(92) I can also remember ehr the bear that used to be (v ʻshort laughter') kept at the Queen's (v ʻlaughter') Head, which used to sit (v ʻshort laughter') used to be chained to a tree at the back and they used to bring ʻ**im** out sometimes and chain ʻ**im** to the horse trough in the in the front of the Queen's Head and eh my old uncle used to give **him** (v ʻlaughter') pints of beer, to drink. **He** used to sit up on there, and eh **he** got a bit nasty, **he** got a bit spiteful, they had to have **him** destroyed. The that was when the Queen's Head was run by the Billows. [...] Well, **he** -- I know, my uncle was, had a drop too much to drink, and he was teasing ʻ**im** with a bottle of beer when **he** (unclear) cleavered (/unclear) his arm and he had he had to go round to the church to old Dr Hatch, who was at Chestnut Cottage, where Chestnut Cottage is now, an' he I think he had fourteen stitches in his arm through ʻim. In the finish **he** got so naughty, they had to they had to put **him** down. (e ʻclattering in the background') [...] Tremendous big bear **he** was though. (gendered reference, non-human animate ʻbear', CP, MDX_002b.wav)

(93) If you had a sow with a litter, that would have three meals a day, quite often. **He** **'**d be fed in the morning then after while **he** 'd be left out, and then **he** 'd have a little bit more food and left into the young ones again, and out again, and that. (gendered reference, non-human animate ʻsow', CorJLN, CON_011)

Among the inanimate referents of *he* and *him* (also *himself* outside the text sample), we find many vessels and vehicles (58%) such as barges, boats, coaches, engines, planes, ships, submarines, tractors, trains and vans. There is also a large variety of other referents inside and outside the text sample, including all sorts of machines and tools (like reapers, sewing machines, weaving looms, ploughs, drills, irons, ovens, wooden churns, and even the tape recorder used in the interview), buildings (farm houses, cottages, churches, mills, pubs, sheds, shops), recipients (bags, bottles, boxes, barrels, buckets, mugs, pots, pails, a washing tub), and various other inanimate objects, such as an accordion, a ditch, a generating plant, clock, hill, hay stack, ladder, lamp post, letter, pram, button, painting, pond, rope, quilt, stove, teeth, bricks and walls, the moon, an apple, seen in example (94), a kiln, seen in (95), and even a toilet, seen in (96).

Similar to the animals mentioned above, most of these inanimate objects are the central topic of the respective discourse segment. Thus, masculine reference can even be used if the referent is assigned inanimate status by other expressions in the same sentence. In example (97), for instance, the *rake* is first referred to as a *terrible dangerous thing*, but becomes *he* as soon as the speaker starts to describe it in detail. This passage is a good example for the distinction stressed above, between objectively verifiable referent properties and the speaker's subjective perception of the referent. From the first perspective, the rake is an inanimate tool; from the second perspective the rake is dangerous. The shift in perspective is, however, not to be confused with personalisation.

(94) Well, like if you picked one [an apple] and cooked it early **he** isn't same as when **he** 's been picked and kept, is **he**? And that. {<u IntML> Yes. Exactly.} So that 's what we

did do. We made (reg sic=un) **him** (/reg) late. (gendered reference, inanimate 'apple', SRLM_GG, SOM_013)

(95) ...but the back end was like I told you before, got hotter than the front, so he could tell by looking at them and looking in the kiln **himself** and he could tell when the back end was finished. (gendered reference, inanimate 'kiln', IBED, SAL_016)

(96) The toilet was right down the bottom of the garden and that was still there when we took over the garden. {<u IntKS> I think there used to be two mud huts there at one time, I don't know. On the old tithe map I think there was two mud huts there then.} Was there? This old toilet, **he** was still there when we were doing the garden. (gendered reference, inanimate 'toilet', SRLM_GP, SOM_027)

(97) They used to be -- they call **it** a turnover, a arrish rake. And **it** was a terrible dangerous thing. **He** had spikes each end, and you went along and when you had enough on **it**, you just lift the handles up, that the prongs would stick in the ground and turn over. Then you 'd go on again until **it** was full again. But it was just the prongs with the slides and **he** was terrible dangerous, dangerous thing. (gendered reference, inanimate 'rake', CorJLN, CON_011)

(98) Well anyhow (unclear) last year (/unclear) **he** tumbled and tossed and then, **he** come in, and they took off, went down, a hold of this barrel and they started to shake '**im**, and the old man said, he said, I don't know, he said, Full of salt water or what, he says, but **it** 's full, anyhow. Wasn't a very big barrel, just a little ordinary barrel. (gendered reference, inanimate 'barrel', CAVA_WW, CON_006)

(99) But the best corn would come this side and when that chute did fill up his bag there so as you could tie **un** nicely, he would switch over, turn the chute over in that bag and while that bag was filling he would pull **him** back and tie **him** and put **him** back in a nice row and hang up another empty one there. (gendered reference, inanimate 'bag', SRLM_WB, SOM_010)

A different mechanism, which has not received much attention in the literature, manifests itself in several interviews. I call it the 'animation of the topical referent'. In the two examples in (100) and (101), the gendered forms accentuate the agent status of the referent (a coach in one example, a plant in the other), making it appear more vivid and animated than neuter *it* would allow for. In both passages, the use of gendered reference makes the described event livelier, more colourful and, ultimately, more story-like. A very similar effect will be described for O-form determiners in gerund-participles in chapter 10.

(100) But this one occasion, I don't know what happened, but it is so that **he** went straight through, this coach did, down the main line, and they sent the shunting engine out of Wellington goods yard to follow **him** down the line and got on to **him** and catch **him**, which they did by means of getting the coal pricker over a lamp bracket and gradually brought **him** to a stand like that and brought **him** back to Wellington with, a shunting engine. (gendered reference with animation effect, 'coach', IBBeBo, SAL_011)

(101) Well you cut them off with a spitter see, about three inches long, the root tap root, and **he** 'd never grow no more **he** 'd just shoot out some more roots out each side, but **he** 'd never grow in **himself** again... (gendered reference with animation effect, 'teasel' (plant), SRLM_AB, SOM_018)

Last but not least, we need to mention mass nouns. Most factors which seem to further the use of gendered pronouns for count referents do not seem to apply to this category. Mass referents are usually attributed a very low degree of individuation, and speakers are less likely to feel emotional attachment. Nevertheless, masculine reference for mass nouns can be observed in the data (no examples with abstract referents, cf. Wagner 2004a: 292–293). Within the selected text sample, occurrences are rare and only appear in a few interviews from the SW area (see Table 7.2), but they provide evidence against the categorical exclusion of gendered reference for non-count referents (contrary, e.g., to Elworthy 1886: 328, or Trudgill 1999[2]: 95).

All mass referents in the corpus belong to the semantic field of food and drink, such as juice, milk, butter, flour, salt, or cow feed (called "kek"). The FRED examples are comparable to masculine references recorded in the SED for mass nouns like gravy, porridge and broth. Some examples allow different interpretations: a specific sort of liquid, for example, may be perceived as a mass referent, or it may be perceived as a count referent if it is contained in some recipient. In example (103), for instance, the referent could be classified as mass noun (rum) or count noun (this specific drink) – a distinction which is grammatically relevant in other languages.[57] In the present study, this distinction is not made but it should be kept in mind.

(102) And then **it** [the butter] was put on the, whether he used to put the salt in then, and then, of course, **it** was ready to go on. In them days, nearly every dairy had the big slate benches all around. [...] And that used to be put on that, and beaten up from there. And then **he** would keep good shape. If it 's summer time and warm weather, there wasn't any fridge or anything, that it used to be awkward to keep the butter to **his** shape. **He** would, you know, go soft, but the slate would retain the coolness for a lot longer. (mixed reference, mass noun 'butter', CorJLN, CON_011)

(103) I said, You 'd better have a drop of hot water and a bit of sugar in (reg sic=em) **him** (/reg) [the rum]. He said, I think I will. [...] He used t' say, It isn't too much trouble? I said, Trouble, no, ain't no trouble at all putting hot water and a bit of sugar in (reg sic=em) **him** (/reg) and the old man used to sit down, he 's happy as a king. (gendered reference, mass noun 'rum', CAVA_WW, CON_006)

(104) ...you know kek for cows like. 'Cause you see we do feed **en** in uh, (unclear) muggen (/unclear) baler there, right there, and **it** 's tipped in, see, and then **it** 's put in to the cows with a shovel, you know what I mean. [...] 'Cause if **he** 's all mixed up, then you got to bag **en** up to tip in, you know, but this here kek do come and uh... (mixed reference, mass noun 'cow feed', CAVA_WJB, CON_005d.wav)

(105) {<u IntCMH> Oh? Well like 'cause they couldn't get the flour, I suppose.} That 's right. Had to wait forty-eight hours before you, before anybody could have **un**, barring seafaring folk... (gendered reference, mass noun 'flour', SRLM_WT, SOM_028)

57 Compare Kimball (1991: 451) for Koasati, a Native American language spoken in Louisiana.

(106) And they used to say, Put a bit of salt in the butter -- well, we did put some salt -- put a bit of salt in the butter, and the (unclear) maynards (/unclear) can taste **en** down underground, you know, under bread and butter. (gendered reference, mass noun 'salt', CAVA_WJB, CON_005)

Referent categories – feminine forms

In the data at hand, gendered reference is by no means restricted to masculine forms, even if gendered *she* and *her* are less frequent than their masculine counterparts (compare Tables 7.1 and 7.2). This result is in conflict with earlier sources that have suggested such a restriction for Modern English (e.g., Ihalainen 1985: 155).

In addition, the referent range of feminine examples is wider than usually expected, even if it is more restricted than for masculine forms.[58] Similar to gendered *he* and *him*, gendered *she* and *her* refer to both non-human animates and inanimate objects, encompassing the following categories:

- animals (cows, dogs, horses in general, mares in particular);
- vessels and vehicles (mainly boats, ships, submarines, engines, trains, bikes);
- other objects (e.g., beam, cross piece, road, envelope, stove).

Unlike *he* and *him*, all gendered *she* and *her* in the data refer to countable entities, never to mass nouns. Examples are given below, including the story of Fan the pony in (107).

Regarding the well-known use of feminine reference for ships and other vessels, I agree with Wales (2002) and Wagner (2004a) that 'personification' is not an adequate label – the label being too general, and the phenomenon itself being too frequent. The respective occurrences can, however, be interpreted as a "subtle form of gender symbolism" (Wales 2002: 333) which is typically used by male speakers with a maritime background. Since different types of vessels are also referred to by masculine forms, gendered reference is to a certain extent unpredictable, as seen in (108). In this particular example, the speaker uses feminine reference for the one large ship which has its own name, distinguishing it from the other, smaller vessels in the same passage.

The fame of ships and their female symbolism, of course, goes far beyond the linguistic literature. This revealing text was recently found on a poster in a maritime shop in Covent Garden, London:

58 In Wagner's data, feminine reference is restricted to vehicles and farming machines. The *her* referents mentioned in Upton et al. (1994: 487) are more varied. Gendered *she* is also mentioned in the SED fieldworker notes under 35,2: *you see how she (the cowshed) is thatched*.

Why is a ship called she?... because there is always a great deal of bustle around her; there is usually a gang of men about, she has a waist and stays; it takes a lot of paint to keep her good looking; it is not the initial expense that breaks you, it is the upkeep; she can be all decked out; it takes an experienced man to handle her correctly; and without a man at the helm, she is absolutely uncontrollable. She shows her topsides, hides her bottom and, when coming into port, always heads for the buoys.

Here are some examples from the corpus:

(107) And, this same bloke Gregory, this one that 's alive up Pinner Hill, when his father died, he had **the pony**. And eh he used to go out at night, (v 'whispers') and used to get as drunk as a newt. And he used to get over to The Ship at Eastcote. And when he come out of The Ship at Eastcote, they used to just open the trap, and they used to put him in the trap, and used to tie the reins, on, right, on the horse's collar, [...] and lead the pony out into the middle of the road, and say, Go on, **Fan**, home. And **she** used to go up that road, they reckoned, like the wind. And if there was anything coming down the road, **she** used to pull out, pull up dead, and draw right into the side until they gone by, and as soon as they gone by, away **she** 'd go again. **She** used to go straight across Potter Street, [...] up the hill, see, turn round, come into the farm -- actually about two or three o'clock in the morning -- right up to the back door, start neighing and hollering like hell, **she** would. [...] And she used to, Mrs Gregory used to come down then, see? There was Jim drunk as a newt in the back, they used to have to take him indoors and unharness **Fan** and (pause) fantastic (trunc) hor- (/trunc)... (gendered reference, non-human animate 'pony', CP, MDX_002a.wav)

(108) No, well, they had a brig first -- **he** got square sails on two masts. But if you had a (trunc) b- (/trunc) , if you had a (trunc) bri- (/trunc), a, a barque, **he** (e 'clock chimes') got square sails on two masts -- if **he** was three master -- and four-and-a-half sails on the mizzen mast. But if was a four-mast barque **he** 'd have square sails on three masts. But a (trunc) sh- (/trunc), forward ship had square sails on all masts. Even the five-master what went ashore before the first war, (trunc) i- (/trunc), just up to the side of Dover, **she** been in collision, the Prussian. **She** was a five-master, but square sails on all masts, that 's a ship. (mixed gendered reference, different vessels, KentSPR, KEN_008)

(109) I liked the ol' Strive. I was in **her**, **she** was a good ol' boat. (gendered reference, inanimate 'boat', ELF, SFK_004)

(110) We come across from Penarth, like I say, we went through the gap, then come hard over and, mind, **she** droved **herself** over, though **she** was double reefed and that... (gendered reference, inanimate 'ship', SRLM_WT, SOM_028)

(111) ...and I saw this cow, and I spoke to **her**, went over and made a fuss of **her**,... (gendered reference, non-human animate 'cow', TCA_GB, DEV_002)

(112) ...and the driver said to me, he said, Thank God I got **her** round Coalbrookdale. (gendered reference, inanimate 'train', IBErDi, SAL_020)

7.2.3 Graphic representation on the Animacy Hierarchy

Gendered pronouns are frequently perceived as a signal for the speaker's emotional involvement or attachment to the referent in question. However, the lack of a direct means for evaluating the speaker's point of view presents a major difficulty in the quantification and interpretation of gendered forms. Descriptions of gendered reference, therefore, usually revert to other, objectively verifiable factors. Siemund (2008), for example, defines the scope of gendered forms, as compared to *it*, on the Animacy Hierarchy (Croft 1990) and the Continuum of Individuation (Sasse 1993). Siemund's findings indicate that the difference between non-standard varieties and Standard English regarding the use of masculine and feminine pronouns "concerns the location of the cut-off point" in these hierarchies (2008: 142).

I have argued above that individuation is not really a referent property. Rather, the "degree to which we see something as a clearly delimited and identifiable individual entity" (Yamamoto 1999: 3) depends on the speaker's perspective. If this perspective is not directly accessible, the degree of individuation of each referent is very difficult to determine. Aspects like humanness and animacy facilitate a more reliable classification (cf. Croft 1990; Kuno 1987).[59]

At this point, the results presented in Tables 7.1 and 7.2 come back into play. Systematic variation was observed within each pronominal gender category (masculine, feminine, neuter) and each animacy category (human animate, non-human animate, inanimate). In Figure 7.1, these results are projected onto the well-known Animacy Hierarchy by Croft (1990), visualising the referential overlap between gendered pronouns and *it*. In accordance with earlier accounts, masculine and feminine forms encroach on the hierarchy from left to right, with linear decline towards the right. At the same time, *it* shows linear decline to the left, predominating in the inanimate class and only rarely referring to humans (in the data only neonates). In addition, the Continuum of Individuation, as presented by Sasse (1993: 659), was also included in Figure 7.1 in order to illustrate the overall correlation between animacy and individuation.

In contradiction to earlier accounts,[60] the use of gendered pronouns, specifically

59 Awareness is acknowledged that the concepts of animacy and humanness are subjective to some extent. Language users may of course regard a specific object as animate based on their cultural or religious background, for instance 'the water' or 'the moon', or personify an object based on their own personal relationship.

60 "The use of animate pronouns for the entire continuum is not known, except for some Creoles (Jamaican creole *im*; cf. Patrick 2004: 428). *He* and *she* are normally not found for reference to abstracts and non-individuated matter." (Siemund 2008: 4)

masculine forms, can be verified for the entire continuum including referents of the inanimate mass category (examples above; no abstract referents).

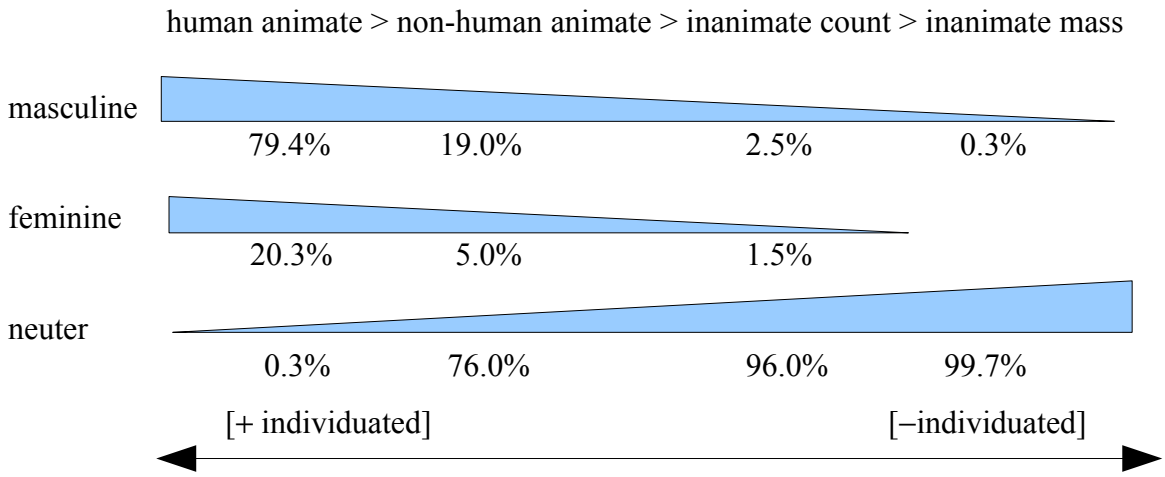


Figure 7.1: Gendered pronouns vs. *it* on the Animacy Hierarchy

7.2.4 Switches between gendered and non-gendered forms

It occasionally happens that speakers use gendered forms and *it* for the same referent in a conversation. Such switches can even occur within the same paragraph or sentence, as seen in (113)–(115). In some examples, the two variants appear to be in free variation (apparently unmotivated switches), whereas examples like (97) show how a change in focus, or a change in the speaker's perception of the referent, can provoke a switch to gendered reference. Switches of this kind can be called *functional switches*.

(113) If **it** was well built, **he** would keep fairly dry. (mixed reference, inanimate 'hayrick', CorJLN, CON_011)

(114) D' you know we 've got one calf before **it** was twelve month old and **he** won't turn without the dummy. (mixed reference, non-human animate 'calf', SRLM_GG, SOM_013)

(115) Oh yes, I remember when **he** was an old pub then, cor wasn't **it** hellish low, yeh you had to go down a step first to go in, yeh. (mixed reference, inanimate 'pub', TCA_FK, DEV_008e.wav)

One subtype of functional switches are what I call *demarcation switches*, i.e. switches between gendered reference and *it* which are used to demarcate different referents in the discourse. In (108), for example, the alternation between masculine and feminine gendered

forms sets apart different ships. Remember that in this particular example, feminine reference was used for the one large ship which had its own name, while other vessels were masculine. A similar effect can be observed in (116) and (117).

(116) And the first horse, I bought a horse, or father did for me, I gave eighty-five pound for **it**. I bought **it** in September, and by Christmas **it** has lockjaw and died. And then I had to buy another one. But in the meantime, a friend of mine, down Ludgvan, he seen me. I used to buy feeding stuff off his brother, and he said, I got a horse home, Gilbert, he said, that 'll tide you over for a few months. **He** 's old, but **he** 'll work. And I was very grateful for that man for that. He helped me out a lot. And after a few months, I got over the shock. I went up to Mithian, and I bought a horse for fifty pounds. **She** wasn't a perfect horse, but **she** could do the job. And that kept me going. (switching between gendered and non-gendered reference, different horses, ICS_EW, CON_009)

(117) And well, we pulled this thing, **it** [horse1] leapt, and **it** was quite a thing, because **it** was fairly strong at that age. And we brought **it** in, and we hadn't had any experience much, of breaking in. But, you know, **it** came in all right. With the other, we had an old mare, was a big Shire one, that was a bit slow really, but we put **her** [mare] in beside of that one, and **he** [horse1] couldn't do much. You know, **he** had to go on with **her**, because **she** 'd drag **him** along too, and then **she** settled down to do **her** part. (switching between gendered and non-gendered reference, different horses, CorJLN, CON_011)

In juxtaposition, gendered pronouns outweigh *it* in speaker closeness, since they naturally indicate greater familiarity or intimacy in the speaker–referent relationship. In (116), for example, the speaker uses *it* for the first horse in the first few sentences – an animal he only had for a short time. The second horse in the middle of the paragraph, probably a stallion or gelding, is referred to as *he*; and the third horse, probably a mare, is referred to as *she*. Similarly, the speaker in (117) describes the untamed wild horse (*thing*) as *it*, in contrast to his *old mare* which is referred to as *she*. Once the untamed horse joins the mare and slowly becomes domesticated, *it* becomes *he*. The different referential forms in this example not only set apart the two referents – the old mare and the new horse – but also upgrade the latter from *it* to *he*.[61] Similar examples leave no doubt that emotive aspects play an important role in the choice of gendered reference.

7.2.5 In-text distribution: semantic priming and semantic differentiation

In addition to the possible determinants of gendered reference described above – the semantic properties of the referent, the speaker perspective, and functional switches – it is important to consider the influence of priming effects. In their investigation of complex noun phrases, for example, Cleland and Pickering (2003) found that production priming is

61 Referents can be upgraded from *it* to *he*/ *she*, or downgraded from *he*/ *she* to *it*. According to Wagner (2004a: 124), this "intimate pattern" can also involve switches between masculine and feminine forms.

stronger if the head nouns in the prime and target match.[62] Accordingly, production priming can be expected to be strongest in pronominal expressions if the referent of the prime and target match. This explains why we find some texts in the corpus which have no gendered pronouns at all whereas others feature numerous occurrences, often accumulated in certain text passages. Two such examples are the story of the bear in (92) and the story of the pony in (107). In spontaneous conversation, turn-taking can of course interrupt a chain of gendered references and lead to pronoun switches, as shown in (118).

(118) <u WflsWGP> Now, oh, you ain't got **en** on, have ye?
{<u IntWH> Yeah. Ah, I can take **it** off again.}
(referring to the recorded conversation, WIL_010c.wav)

A different finding from Cleland and Pickering (2003) is not supported, namely that production priming of complex NPs is stronger when the prime and target head nouns are semantically related. According to the authors, it can be assumed that the production of a sentence *the sheep that is red* (instead of *the red sheep*) increases the likelihood of producing *the goat that is red* (instead of *the red goat*; cf. Cleland and Pickering 2003: 218). Examples from FRED show that this mechanism rarely applies to gendered pronouns. On the contrary, text passages like (108), (116) and (117) show how demarcation switches can help distinguish between different semantically related referents.

Taking into account, (i) the use of demarcation switches, (ii) the complexity of relationships between the speaker and different referents in the same text, (iii) possible shifts in topicality and perspective during the interview, as well as (iv) the individual speaker's general predisposition to use gendered forms, and (v) a certain amount of free variation, the issue of semantic priming becomes exceedingly complex. Overall, sequences of gendered forms appear to be strongest where the prime and target referents are identical, but even then a variety of factors may impede the priming effect.

7.2.6 Syntactic positions

In most examples, the use of gendered reference appears to be intuitively functional or pragmatic, but a look at structural factors yields some interesting results. Most importantly, it can be shown that the syntactic position of the pronoun has an influence on gender

62 In his analysis of persistence in genitive choice, however, Szmrecsanyi (2006: 106) found that "for a given slot in which the same possessor/ possessum phrase is repeated, that genitive type will be chosen which establishes old-before-new order. Ergo, the findings suggest that the whole possessor or possessum phrase [...] is relevant with regard to information status, but a single head noun [...] is not."

choice. This can best be seen if we compare the two referent categories with the largest overlap between gendered forms and *it*, namely non-human animate referents and count objects (see Figure 7.1). The numeric results are shown in Table 7.3.

Table 7.3: Non-human references in text sample, breakdown by syntactic function

syntactic function	animate non-human			inanimate count		
	gendered forms (m+f)	*it*	TOTAL	gendered forms (m+f)	*it*	TOTAL
subject	65 (33.2%)	131 (66.8%)	196 (100%)	137 (6.5%)	1978 (93.5%)	2115 (100%)
object	33 (20.1%)	131 (79.9%)	164 (100%)	73 (3.6%)	1933 (96.4%)	2006 (100%)

In both categories – animate non-human and inanimate count – the numbers show a significantly higher proportion of gendered forms in subject function as compared to object function ($p < .05$), and a subject–object ratio of approximately 2:1. From a wider perspective, these results link up with the general tendency of traditional non-standard phenomena to survive in the most frequent grammatical relations (cf. Krug 2003: 18). Should it prove true that gendered pronouns are taking the usual path of morphological change, future stages of English may witness their disappearance, starting in object positions.

7.2.7 Gendered pronouns and speaker sex

Gendered pronouns are a widespread phenomenon. Among the 48 speakers in the text sample, 26 used gendered forms at least occasionally (21 used gendered *he*, 20 *him*, 10 *she*, 10 *her*). In the whole dataset, 72 speakers were found to use gendered *him*, the pronoun with the highest absolute number of gendered occurrences (385; compare him_allgendered.xls). This is over a third of all speakers. Gendered reference can hence be described as a generalised feature (see 5.2.3).

An aspect not addressed so far is the linguistic behaviour of male speakers as compared to female speakers. If we assume that gendered reference is largely influenced by emotive aspects and the speaker viewpoint, and if we also assume that men and women differ in their perception of and attachment to certain referents, as well as in the way they

express this attachment, differences are to be expected. Table 7.4 shows a breakdown of gendered pronouns by speaker sex, including 11 female and 36 male speakers (1 unknown; neuter values extrapolated from same random sample as above).

Table 7.4: Gendered pronouns and speaker sex

referent category	non-human animate (N=421)	male speakers	masculine reference	71 (25.4%)
			feminine reference	20 (7.2%)
			neuter reference	188 (67.4%)
		female speakers	masculine reference	9 (6.3%)
			feminine reference	1 (0.7%)
			neuter reference	132 (93.0%)
	inanimate count (N=5903)	male speakers	masculine reference	145 (2.7%)
			feminine reference	89 (1.7%)
			neuter reference	5147 (95.7%)
		female speakers	masculine reference	1 (0.2%)
			feminine reference	0 (0.0%)
			neuter reference	521 (99.8%)

The empirical results can be summarised as follows:

1. the male speakers used gendered reference more frequently than the female speakers;
2. gendered masculine forms were generally preferred over gendered feminine forms by both sexes (compare the results in 7.2.1 and Figure 7.1);
3. the male speakers also preferred masculine forms over feminine forms for inanimate referents, while female speakers used basically no gendered reference in this category; the use of gendered pronouns for inanimate referents is hence almost exclusively attributable to male speakers.

It is also important to note that no correlation was found between speaker sex and pronominal gender of the type suggested by Mathiot and Roberts (1979) for American English, i.e. no preference of male speakers for feminine forms, or female speakers for masculine forms.

The difference between male and female speakers in Table 7.4 is very pronounced in both referent categories, but it is strongest for inanimate count objects ('non-human animate': 7.24 σ, $p = 3*10^{-4}$; 'inanimate count': 12.32 σ, $p = 7*10^{-10}$). Considering that the

female speakers in the text sample used basically no gendered forms for inanimate referents, future analyses of sex-specific linguistic behaviour should take into account the referent category as an additional determinant.

An explanation for the observed differences is once more found in the speaker viewpoint and the speaker's attachment to the referent in question. There is no doubt that, on average, the male speakers in the corpus talk about objects like vehicles, ships, machines and tools more frequently and in more detail than the female speakers. If we, therefore, assume that male speakers feel, on average, more familiar with these objects, the occurrence of gendered forms comes at no surprise. It also stands to reason that a speaker with a marine background (among the traditional dialect speakers in FRED only men) will refer to boats and ships as *she*, using so-called symbolic gender or gender of affection. Taking into account these considerations, we can argue that, even if speaker sex and viewpoint occasionally correlate, the two aspects need to be treated separately.

7.2.8 Areal distribution

The areal distribution of gendered pronouns observed in this study stands in opposition to the frequent assumption that gendered pronouns are an exclusively Southwestern phenomenon. At the same time, however, the results indicate why occurrences in other parts of the country are occasionally overlooked (e.g., Trudgill 1999[2]; Upton et al. 1994): gendered pronouns are significantly more frequent in the South than in the Midlands and North.[63] The observed frequencies mark a South–North continuum which is most noticeable among the more frequent masculine forms.

Figure 7.2 shows the relative frequencies of gendered forms in the 'non-human animate' category. Masculine forms are in blue, feminine forms in orange, and all gendered forms in green (masculine + feminine), with exponential trendlines in the respective colours. Figure 7.3 shows the corresponding distribution for the 'inanimate count' category.

63 The different case forms behave very similar in this respect; form-specific properties (e.g., phonological aspects) can hence be ruled out as a major determinant. Note that the SED questionnaires and incidental material show gendered pronouns mainly in the SW, but also, for example, in Yorkshire. For the geographic distribution of gendered proportions among the different Southwestern counties compare Wagner (2004a).

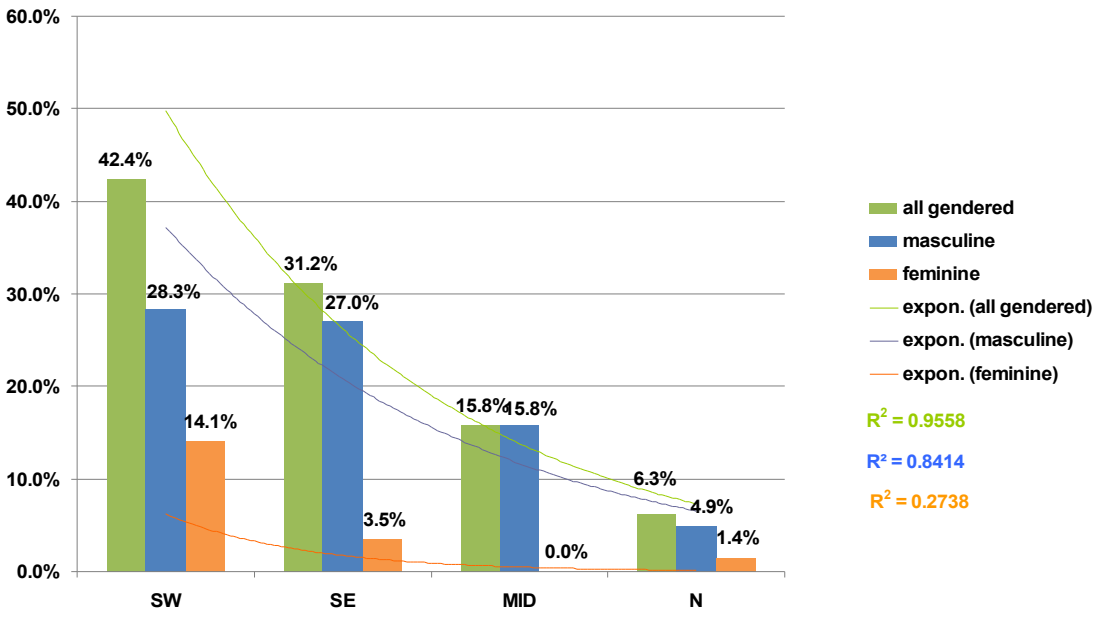


Figure 7.2: Geographic continuum: gendered pronouns, 'non-human animate' category

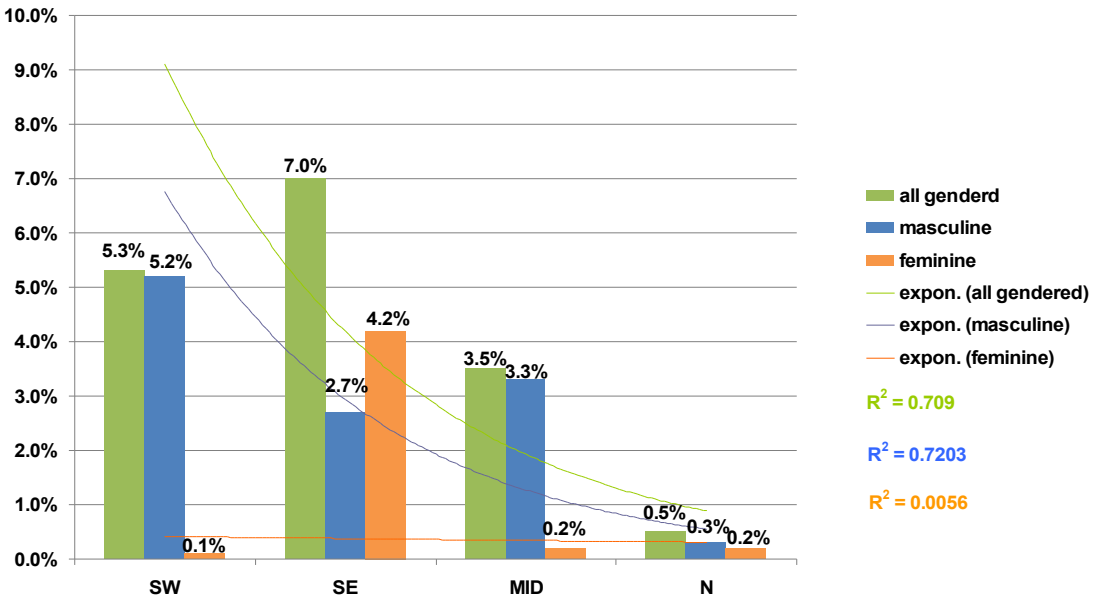


Figure 7.3: Geographic continuum: gendered pronouns, 'inanimate count' category

In both referent categories, we get significant differences in frequencies between the SW and N areas, and between the SE and N areas, for both masculine forms and gendered forms in general. The only noticeable exception within this South–North continuum is the relatively high proportion of feminine forms used with inanimate count referents in the SE

area. A closer look at the data reveals that the exceptionally high frequency of 4.2% (as compared to 0.1% and 0.2% in the rest of the country) is attributable to two speakers who use feminine reference throughout their stories for different ships, boats and engines. Together, speaker JRS (male, SFK_011, 44 cases) and speaker EF1 (male, SFK_003, 27 cases) are responsible for 71 out of 82 *she* and *her* with inanimate count reference in the SE area. In terms of speaker–case ratio, these two speakers represent outliers who strongly bias the observed frequencies. The overall South–North cline, however, remains visible.

7.3 Summary

In this chapter, a cross-regional analysis was presented of gendered pronouns in England. Compared to other non-standard features in the same data, gendered pronouns represent a medium to high frequency phenomenon. Compared to earlier accounts, however, a significant reduction in frequency seems to have taken place. The well-known traditional West Country system based on count–mass distinctions is no longer in place. Instead, the results describe a South–North continuum, where the highest gendered proportions are attributable to masculine forms with non-human animate referents.

The corpus results show a general predominance of masculine forms over feminine forms. Unlike in previous studies, empirical evidence was found for masculine pronouns with mass reference. Yet, the number of examples is limited, and it can be suspected that masculine reference will not persist in this referent category much longer.

It was argued that gendered pronouns encode pragmatic language use in close connection with the speaker viewpoint. In a variety of different functions, gendered reference is used for the topicalisation and animation of focal referents, as well as the demarcation of different referents in the conversation by functionally motivated switches between gendered forms and *it*.

Gendered pronouns are one of various linguistic phenomena where emotive aspects play a crucial role in pronoun choice. Gendered forms can be used to express the speaker's personal involvement or attachment to a specific referent. This is also reflected in differing preferences between the male and female speakers in the corpus: the well-known gender symbolism of vessels, for example, is typically found in male speakers with a marine background. The results obtained in this study have serious implications for predictions about sex-related language use. Most importantly, they call for a specification of generalised claims, such as the alleged avoidance of vernacular features in female speech.

In the literature, aspects like individuation and animacy have been used to describe the scope of gendered reference in English. I have tried to illustrate the subtle but important distinction between objectively verifiable referent properties and the subjective attribution of referent properties. Especially individuation – defined as the degree to which we see something as a clearly delimited and identifiable individual entity – is epiphenomenal to the speaker viewpoint, and can therefore not be attributed determinant status. However, it needs to be acknowledged that in data like FRED the speaker viewpoint is not directly accessible, since no additional inquiries are possible. Animacy, in this respect, provides the most objective means of classification. In Figure 7.1, the distribution of gendered pronouns vs. *it* was projected onto the Animacy Hierarchy. The biggest overlap is recognisable in the 'non-human animate' category.

Last but not least, gendered pronouns appear to be following the general tendency of traditional non-standard phenomena to survive in the most frequent syntactic roles. This is indicated by the significantly lower proportion of gendered forms in object function, as compared to subject function. Based on the empirical results, it is to be expected that, if gendered pronouns were to disappear from the English language, they would first disappear from the inanimate referent categories (i.e. right to left on the Animacy Hierarchy), and the less prominent syntactic positions (i.e. first object, then subject).

Chapter 8
Pronoun exchange

When it comes to variation in pronoun use, case is certainly the most widely discussed grammatical category. In spontaneous spoken English, the morphological realisation of syntactic relations allows for substantial variation in many different syntactic functions. The aim of the following chapters is to describe the different non-standard uses found in the data at hand, starting with the well-known phenomenon of pronoun exchange.

8.1 Traditional definition

Pronoun exchange (PE) is probably the best-known vernacular phenomenon regarding variation in case. Although it is not language-specific, the phenomenon also known as pronoun substitution (Trudgill 2004a: 147) has certainly made its mark in the linguistics of English.[64] From a typological perspective, "pronoun exchange in dialects of English contradicts the prototypical case-marking of subjects and objects in 'accusative' languages" (Kortmann 1999: 6). The acceptability of PE in everyday conversation is readily explainable by the fact that speakers of English are used to a language that has already lost its morphological case in nouns, and which heavily relies on word order for the identification of syntactic roles (cf. Koktova 1999).

Pronoun exchange was first documented in the dialectological literature of the 18th and 19th centuries. In its traditional definition, the term describes the exchangeability of subject and object forms in subject and object functions. Historically, the use of *me*, *him*, *her*, etc. in subject function (*her's not in*) is known since Old English whereas the use of *I*, *he*, *she*, etc. in object function (*she told I*) emerged during the Early Modern English period (compare 3.1). Early definitions of PE can, for example, be found in Elworthy (1877: 35), Kruisinga (1905: 35) and Wright (1905: 271).

The great attention given to irregular uses of subject and object case forms is probably due to the fact that the distinction between subject and object forms and functions is

64 Azevedo (1989: 863–864), for example, has reported the use of S-form objects in Brazilian Portuguese, both in the 'substandard' varieties and the colloquial speech of 'educated' speakers. For example, *Ela chamou eu* 'she called I' (Standard Brasilian Portuguese *chamou-me*), or *Eu chamei ela mas ela não respondeu* 'I called she but she did not answer' (StBP *chamei-a*). Unlike English PE, the Portuguese phenomenon appears to be uni-directional.

generally considered an essential part of the language. In the early 1920s, Edward Sapir stated:

> Surely the distinction between subjective *I* and objective *me*, between subjective *he* and objective *him*, and correspondingly for other personal pronouns, belongs to the very core of the language. We can throw *whom* to the dogs, somehow make shift to do without an *its*, but to level *I* and *me* to a single case – would that not be to un-English our language beyond recognition? There is no drift towards such horrors as *Me see him* or *I see he*. (Sapir 1921: 177)

Geographically, PE is often associated with dialects of the Midlands and the Southwest of England (see Map D3, Appendix; cf. Burchfield 1994; Hughes and Trudgill 1996; Wagner 2004b). In one of the earliest accounts, for example, Halliwell (1887) mentioned occurrences of PE in the Midland counties of Worcestershire, Warwickshire and Gloucestershire. The SED findings identify two main areas for subject *her* in the Southwest and western Midlands, as shown in Maps D4 and D5 (Appendix). The phenomenon has, however, also been observed elsewhere, including East Anglia (cf. Trudgill 2004a, only S-form objects), the North of England (cf. Beal 1993; 2004), Irish English (cf. Filppula 1999), and Newfoundland English (cf. Wagner 2002).

Despite the many references in the literature, descriptions of PE are often anecdotal and restricted to individual examples from specific dialect areas. According to Trudgill (2004a: 147), the exchange of different case forms in English "has not yet been subjected to any definitive analysis." The present study fills this gap with a systematic empirical investigation which takes into account the cross-regional distribution of PE occurrences and the different PE sub-features.

An additional incentive for this particular analysis is the fact that, side by side with the literature on PE, there is an ongoing debate on so-called 'unbound', 'untriggered' or 'independent' reflexives which occur instead of personal pronouns in certain syntactic and pragmatic functions. Given that the three case forms – S-, O- and self-forms – have a long history of co-existence and exchangeability (compare 3.1), it is surprising that studies on either phenomenon have not elaborated on a possible connection. The present study considers both phenomena, the exchangeability of S- and O-forms as well as the exchangeability of personal pronouns and self-forms, where they occur in the same syntactic functions. The following analyses break free from the traditional definition of PE.

The resulting picture shows a more complex phenomenon which includes both personal pronouns and self-forms in a variety of syntactic contexts, leading to a modified definition of PE in 8.6.

No matter where I roam
I will return to my English rose
for no bonds can ever tempt me from she
(The Jam "English Rose")

8.2 Subject forms in object function

The phenomenon known as pronoun exchange consists of different sub-features which will be discussed separately in the following sections. We start by taking a look at subject forms in object function. The most common explanation for this feature is emphasis. It has been argued that S-forms tend to be used in object function when the respective pronoun carries additional stress. Wakelin (1986: 34), for example, mentions the use of subject case forms as emphatic grammatical objects in the Southwest of England,[65] and Trudgill (2004a: 147) notes that "it seems possible that what happens is that the Standard English subject pronouns occur as objects when the pronoun is emphasised, and object pronouns as subjects when the pronoun is not emphasised." In an early account of the Berkshire dialect, Lowsley (1888: 6) lists *she* as an emphatic object in opposition to non-emphatic *her*. Lowsley also includes two syntactic rules which state that "active verbs govern the nominative case", as in *they love we* or *he hates they* (p. 13), and that "[p]repositions sometimes govern the nominative case", as in *from they as hate you expect malice* or *from he as is cunning expect deceit* (p. 14).

Other studies express a different view, finding that non-standard S-forms are not restricted to emphatic uses. Wagner (2002: 25), for example, observes that in her SW data "there is no reason whatsoever to claim an 'emphatic' status for PE forms. Neither do speakers put particular stress on the forms, nor is there anything in the immediate environment that would cause a need for emphasis." Similarly, Ihalainen observes that most nominative objects attested in his recordings are "not 'emphatic' at all" (1991: 107) and that "at this stage, no more can be safely said about nominative objects other than that they are by no means restricted to emphatic contexts" (1985: 160). In the present study, this view is supported by different occurrences where S-form objects can, but do not need

65 Based on 27 extracts from tape-recorded interviews from Cornwall, Devon, Somerset, Dorset, Wiltshire, Bristol, Avon, West Hampshire; approx. 13,000 words.

to, function as emphatic pronouns (compare the soundfiles for different examples mentioned below).

The first quantifiable result regarding non-standard S-forms in the FRED corpus is that all of them – *I*, *he*, *she*, *we*, *they*[66] – are occasionally used as simplex, non-coordinated objects, as seen in the function matrix in Table C1 (Appendix). Furthermore, all S-forms also occur in prepositional complement function, another syntactic context where Standard English requires object case (also shown in the matrix). The areal distribution of non-standard frequencies in the two syntactic functions is very similar, non-standard S-forms clustering in the SW (more details below). In addition, S-forms occasionally appear in for–to and ECM constructions, showing that the grammatical range of PE features is wider than usually described.

8.2.1 Object and prepositional complement function

Table 8.1 shows the S-form proportions for all pronominal objects and prepositional complements in the corpus (limited significance where the overall number of occurrences is too small). The overall most frequent S-form object in the data is *he*. With 41 occurrences, *he* covers 2.1% of all 3SGm pronominal objects.

Two aspects are particularly striking if we compare the two syntactic functions. First of all, the range of non-standard frequencies in non-coordinated cases is quite narrow, reflecting the behavioural similarity of the different pronouns *I*, *he*, etc.: S-form objects range between 0.4% and 2.1%, S-form prepositional complements between 0.2% and 1.7%.[67] Secondly, a comparison of the two functions shows similar median and average values which are slightly higher for prepositional complement function (1.3) as compared to object function (1.0).

The distributional similarities of non-standard S-forms in the two functions support the inclusion of 'prepositional complement' in the definition of PE. Similar to object function,

66 *You* and *it* are not included in the analyses due to the missing case distinction. Also note that, according to Cardinaletti and Starke (1994), *it* falls into a universal category of pronouns with [–human] reference which can not be coordinated (**It and the other one are nice*). *It* is therefore excluded from the analysis of both PE and the influence of coordination on PE.

67 Second person cases are not included because of the missing case distinction in *you*. Among the 38 occurrences of *thou* there are two objects: *he said, Look here, Frank, if I take thou to court, thou will have to pay the value of the cow* (SAL_013), and *my mum 'd say, Oh, What 's eh, what 's he give thou?* (LND_003). Note that all *thou* in the corpus occur in quotations from the bible or quotations of direct speech produced at some point in the past. Occasionally, *thou* occurs where a speaker produces lists of dialect words from a specific region, e.g. *lookst tha was lookest thou, and seest tha was seest thou, and harkest tha was harkedst thou* (WES_005). In one case, the speaker used *thou* to address the interviewer: *Then her 'd shuffle herself, thou know'st, stand up straight* (SAL_038).

prepositional complement function is another syntactic context where actual production data violate the Standard English requirement for the exclusive use of O-forms.[68]

Table 8.1: S-form proportions (%) in object and prepositional complement function (including all POS 20–23 and POS 40–42; not including POS 62a, POS 81 and POS 70–73; compare Table B3)

syntactic function	*I* (of all 1SG)	*he* (of all 3SGm)	*she* (of all 3SGf)	*we* (of all 1PL)	*they* (of all 3PL)	TOT. %
object	**1.3** (25/ 1915)	**2.1** (41/ 1949)	**0.4** (3/ 754)	**1.0** (9/ 945)	**0.5** (29/ 5406)	**1.0** (107/10969)
object, coord. initial	**0.0** (0/ 9)	**0.0** (0/ 1)	**0.0** (0/ 0)	**0.0** (0/ 1)	**0.0** (0/ 1)	**0.0** (0/ 12)
object, coord. final	**50.0** (2/ 4)	**0.0** (0/ 0)	**0.0** (0/ 1)	**0.0** (0/ 1)	**0.0** (0/ 3)	**22.2** (2/ 9)
object, coord. middle	**0.0** (0/ 3)	**0.0** (0/ 0)	**0.0** (0/ 0)	**0.0** (0/ 0)	**0.0** (0/ 0)	**0.0** (0/ 3)
prep.complement	**1.1** (14/ 1221)	**1.7** (16/ 941)	**0.2** (1/ 425)	**1.5** (10/ 665)	**1.3** (18/ 1418)	**1.3** (59/ 4670)
prep.comp., coord.initial	**6.7** (1/ 15)	**0.0** (0/ 2)	**0.0** (0/ 1)	**0.0** (0/ 1)	**0.0** (0/ 0)	**5.3** (1/ 19)
prep.comp., coord. final	**33.3** (2/ 6)	**0.0** (0/ 2)	**0.0** (0/ 0)	**0.0** (0/ 0)	**66.7** (2/ 3)	**36.4** (4/ 11)
TOTAL	**1.4** (44/ 3173)	**2.0** (57/ 2895)	**0.3** (4/ 1181)	**1.2** (19/ 1613)	**0.7** (49/ 6831)	**1.1** (173/15693)

(119) ...so he told **I** he ’d gid **I** the sack, **I** and my father (simplex and coord. object *I*, WflsFP, WIL_009)

(120) And uh, and then I sawed the, I sawed the timber round what were a fit to keep the (trunc) s- (/trunc) to draw down, and I made two hundred and fifty pound of that, so that brought **I** back to three thousand. (object *I*, WflsWGP, WIL_010b.wav)

(121) ...but he never interfered **with I**, but anybody else who came down here, he ’d go for. (prep. compl. *I*, SRLM_AE, SOM_020)

(122) There was ten years **between my sister and I** see, she died last year, she was eighty-nine. (coord. prep. compl. *I*, final position, WIL_023, TrbrAS)

(123) We snapped **he** off like a damn carrot! (object *he*, referent ‘anchor’, SRLM_WT, SOM_028)

(124) So he said, I could do **with he** for a fortnight. (prep. compl. *he*, ICS_EW, CON_009)

(125) I did give **she** a ‘and and she did give **I** a ‘and and we did ‘elp one another. (*she* and *I* objects, WflsWH, WIL_011)

68 S-form objects are, for example, documented in the SED (e.g. Question VIII.7.5 ‘burglars steal they’); S-forms in prepositional complement function are, for example, documented in the CLAE (Viereck 1991; map M15 ‘with me’).

(126) That were, she were two year old. Never had no fault at all **with she**. (prep. compl. *she*, referent 'horse', SRLM_RF, SOM_005)

(127) Work didn't frighten **we**, we knew we had to do it, we had to get on with it. (object *we*, SRLM_R, SOM_005c.wav)

(128) And it 's still the Grasses' house, you know, **with we**. (prep. compl. *we*, CAVA_WJB, CON_005a.wav)

(129) Of course the Plymouth buses then was the DMT's Devon Motor Transport that was the green buses I remember **they**, you can't remember **they** I suppose now. (object *they*, TCA_FK, DEV_008b.wav)

(130) All sorts of long vases, I can show you some **of they**, I got photographs here. (prep. compl. *they*, SRLM_CA, SOM_009)

8.2.2 For–to and ECM constructions

Subject forms also appear occasionally in for–to and ECM constructions, as seen in the examples below (coded POS 47 and 48 in the spreadsheets, compare Table B3). In these two syntactic environments, standard object case is assigned to the logical subject of the embedded clause by the main verb in sentences like *he would* ***wait for me*** *to push the bike* or *she* ***wanted us*** *to have it*. A different explanation proposed by Joseph Emonds is missing case assignment which leads to the use of default case, i.e. object case. According to Emonds (1986: 127), "the subject NP in the for–to construction is attached via a local, language-specific transformational rule to the introductory subordinating conjunction" leaving "the subject position inside S 'empty' in surface."

Neither of the two constructions is usually mentioned in connection with PE, but the observable variation in case is similar to that in object or prepositional complement function. Due to the small number of ECM constructions in the corpus,[69] S-form occurrences are extremely rare. Further investigation is required for a more precise description. In for–to constructions, the number of occurrences is only slightly higher, but a comparison of the measured total frequency against 0 suggests that the detected S-forms can not be dismissed as 'errors' (2.31 σ, $p = .02$; used by 5 different speakers).

(131) ...he said, Don't suppose you would like for **I** to draw your ship, cap'n? (*I* in for–to construction, SRLM_WT, SOM_028)

(132) ...and when we been going up hill, he 'd go up the hill and wait for **I** to push the bike up, ah. (*I* in for–to construction, TrbrGR, WIL_001)

69 Excluding main verbs *ask* and *persuade*.

Table 8.2: S-form proportions (%) in for–to and ECM constructions

	I (of all 1SG)	*he* (of all 3SGm)	*she* (of all 3SGf)	we (of all 1PL)	*they* (of all 3PL)	TOT. %	σ	*p*
for–to	**11.8** (2/ 17)	**0.0** (0/ 5)	**0.0** (0/ 6)	**7.7** (1/ 13)	**5.0** (2/ 40)	**6.2** (5/ 81)	**2.31**	**.02**
ECM	**1.9** (1/ 52)	**0.0** (0/ 0)	**0.0** (0/ 12)	**100.0** (1/ 1)	**0.0** (0/ 9)	**2.7** (2/ 74)	**1.43**	**.15**

(133) He 'd done it so many times and wait for **we** t' come. (*we* in for–to construction, WflsWH, WIL_011)

(134) She used to bring old family stockings for **we girls** to darn and woe betide anybody who didn't do them properly. (qualified *we* in for–to construction, IBAnBa, SAL_029)

(135) ...then you take them ou', pu' them on the mules for **they** to spin the bobbins, for the looms. (*they* in for–to construction, TrbrDC, WIL_018)

(136) There 's no fire in there, the hot water, the boiler, the boiler 's in the basement [...] all domestic uses, in hospitals and all schools and all that. That was the hot-water system see for (pause) **they** to use, see. (*they* in for–to construction, TrbrCS, WIL_005)

(137) I said, they want **I** to go to Charlton. (*I* in ECM construction, WflsEF, WIL_012)

(138) ...she never recovered from it and ehr she wanted **we** to have 'is and give us so much money to have ours. (*we* in ECM construction, SRLM_RF, SOM_005g.wav)

8.2.3 Areal distribution

The areal distribution of S-form objects and prepositional complements in FRED is visualised in Figures 8.1 and 8.2, based on relative weighted frequencies.[70] Both sub-features cluster heavily in the counties of Cornwall, Devon, Somerset and Wiltshire, representing predominantly Southwestern features (compare the 'area' and 'county' columns in the spreadsheets).[71] The non-standard frequencies are not correlated with the overall number of pronominal objects and prepositional complements in the individual dialect areas.

In the following, it will be shown that the observed distribution of non-standard S-forms is exceptional compared to other sub-features of PE. In for–to and ECM

70 Weighted frequencies were used to facilitate a comparison between the four dialect areas. The weighted frequencies are based on the relative frequencies for each area. For subject *me*, for example, we get frequencies of 0.24 SE, 0.07 SW, 0.17 MID, 0.18 N; taken *100 this renders values of 24 SE, 7 SW, 17 MID, 18 N (Total 66), and weighted percentages of 36.4% SE, 10.6% SW, 25.8% MID, 27.3% N. Compare Dahl (2001: 1458): "the study of social and geographic variation in synchronic linguistics has to take uniformity between dialects as the null hypothesis."

71 The distribution in FRED is, for example, consistent with results from CLAE, Map M15 'Will you come with me', where prepositional complement *I* only appears in the South/ Soutwest. Also consider the distribution of 1PL *us* and *we* in WM 178, shown in Map D6 (Appendix).

constructions, S-forms also appear more frequently in the SW but more data are needed to validate the results.[72]

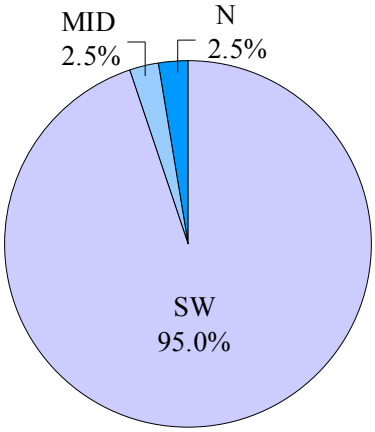


Figure 8.1: Areal distribution: S-form objects (*I*, *he*, *she*, *we*, *they*)

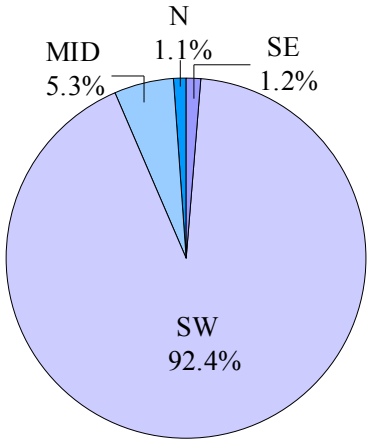


Figure 8.2: Areal distribution: S-form prepositional complements (*I*, *he*, *she*, *we*, *they*)

8.3 Object forms in subject function

The use of object forms in subject function, especially in coordinated NPs such as *him and me* or *us and them*, has been described as "arguably widespread enough among educated speakers in [present-day English] to be called standard" (Denison 1998: 109). O-form subjects feature, for example, in several SED questions, and they were also noted in the incidental material of the survey.[73]

Typologically, English is not the only language where O-forms appear in the subject domain. In spoken Swedish, for example, *dem* and *dom* 'them' have almost completely replaced *de* 'they' (cf. Kjellmer 1986: 446, fn 4). The phenomenon is, however, not to be confused with the regular use of so-called quirky subjects in Icelandic, i.e. subjects with lexically selected case (see 3.3; cf. Sigurðsson 1992; Fanselow 2002; Woolford 2006).[74]

Historically, O-form subjects have been known for a long time, early occurrences being found in Old English examples such as (1) above. In the 2SG, the diffusion of object

72 For–to: SE 0/ 10, SW 5/ 26, MID 0/ 22, N 0/ 18; ECM: SE 0/ 13, SW 2/ 23, MID 0/ 16, N 0/ 20.

73 Questions IX.7.2. *is her married/ is them (two) wedded*; IX.7.7. *her is*, *them is*; IX.7.9. *her is*; IX.7.10. *her's not*; IX.7.3. *aren't them*; IX.7.4.+ 5. *isn't them*; IX.7.6. *wasn't/ weren't them*. Incid. mat.: 15,2 VIII.5. *didn't them*; 15,4 VI.7.4 *them's acome*; 37,1 IX.7.2. *does thee think them married?* Compare the different maps in LSV 94; LAE M70, M74; CLAE M16; compare Chambers and Trudgill (1998[2]: 130/ 133–134/ 136).

74 In Icelandic, certain verbs specify that their subjects are to be in a case other than NOM. The verb *drift*, e.g., requires subjects in ACC case, whereas *recover* requires subjects in DAT case (compare 3.3.)

you into the domain of subject *ye* came to completion during Early Modern English (see 3.1.3), and although 2SG *th-* forms continue to exist in some varieties, they, too, show a similar development.[75] In present-day English, the use of subject *me* in coordination ranges among the three most prominent non-standard pronoun features listed in the *Handbook of Varieties of English* (Kortmann and Szmrecsanyi 2004: 1198), next to demonstrative *them* and distinct 2PL forms.

While O-form subjects form part of the traditional definition of PE, the use of O-forms after linking verb *be* is not always mentioned. Due to their extremely high frequency in Modern English sentences like *it is me* or *it was him*, O-forms in this syntactic function are usually considered standard. Quirk et al. (1972: 210), for example, state that "although the prescriptive grammar tradition stipulates the subjective case form, the objective case form is normally felt to be the natural one, particularly in informal style." (also cf. Erdmann 1978; Harris 1981).

Historically, the diffusion of O-forms in subject complement position appears to have resulted from the reanalysis of impersonal *it*, as described by Riley and Parker (1998: 38)[76] for the following period:

500–1000:	He it is	*he* is subject, i.e. 'He is it' = 'He is the one'
1300–1400:	It is he	word order change, but *he* is still subject
1400–1500:	It is he	*it* is reanalyzed as subject, so nominative case apparently follows *be*
1500:	It is I	nominative case following *be* extends to first person
1600:	It is me	objective case follows *be* in analogy with other verbs

75 According to Stratton (1949: 148), the singular *th*-forms were still used by Quakers in the mid 20th century, but in Quaker English "they have replaced the older correct nominative *thou* by the objective *thee*." The SED identified a subject *thou* area in the North of England, a subject *thee* area in the western Midlands and the Southwest, and a small subject *ye* area in the Northeast. See Wakelin (1986: 33) for the use of *thyself/ theeself* in southwestern dialects.

76 Based on Traugott (1972: 129). According to Kiparsky (2002: 31), personal pronouns in predicative position are not attested for Old English (e.g., **Se cuma wæs ic* 'The guest was I'). According to Rissanen (1999: 262), O-forms started to emerge in subject complement positions towards the end of the 16th century after a development from *I it am* via *It am I* to *It is I*: "In the last-mentioned type, the form of the copula [*is*] reveals that the impersonal *it* has become the subject, with *I* as its complement." According to Jespersen (1949: 262–264), the chronological order of development of different pronouns could be phonetically conditioned: objective *me*, which is similar to subjective *he*, *she*, *we*, gets used in subject contexts where *him*, *her*, *us* are not found. Ihalainen (1994: 315, fn. 16) contradicts this view, claiming that "person overrides any phonetic influence". According to Ihalainen's data, 1SG and 2SG pronouns were the first to appear after copula verbs, followed by 3PL and 3SG in the 17th, and 1PL in the 18th century. A French origin can probably be excluded, since the phenomenon emerged "at a time in which the contact with French was not intimate enough to affect the syntactic structure of English" (Rissanen 1999: 262).

According to this brief outline, O-forms started to replace S-forms after *be* around 1600, following the general tendency of post-verbal pronouns to be in the objective case. It appears that *It is I* was already being perceived as stilted and unduly formal in 18th century informal correspondence (cf. Tieken-Boon van Ostade 1994: 233).[77] Nevertheless, advocates of Standard English have supported the use of *It is I* well into the 20th century, even when aware that at least part of the English-speaking population feels uncomfortable using it. In his *Guide to Correct English*, for example, Stratton (1949: 146) holds that, although to some people "these nominative pronoun forms after verbs seem strange", they should "become comfortable with them. Learn to say over the phone 'Is that *you*, Ethel?' 'Yes, this is *I*.' 'Is Mrs. Martin there?' 'This is *she*.' "

Stratton's prescriptive attitude contrasts with more recent descriptive grammars such as the *Cambridge Grammar of the English Language* (Huddleston and Pullum 2002). Geoffrey Pullum, one of the authors, took a humorous stand towards subject complement *I* in the following interview with the University of California, Santa Cruz:

> The forms with nominative pronouns sound ridiculously stuffy today. In present-day English, the copular verb takes accusative pronoun complements and so does 'than.' My advice would be this: If someone knocks at your door, and you say 'Who's there?' and what you hear in response is 'It is I,' don't let them in. It's no one you want to know. (http://www.ucsc.edu/currents/01-02/04-15/rules.html)[78]

The underlying assumption that personal traits are revealed in linguistic behaviour is not new. One recent example from fiction, where the use of subject complement *I* has similar connotations to those expressed by Pullum, is found in the last Harry Potter novel. Here, Severus Snape reveals himself as "It is I" – an apparently suitable answer for such an obscure character.[79]

The discrepancy between different opinions concerning grammatical correctness can, of course, cause uncertainty among language users and possibly lead to hypercorrection. The mechanism has been aptly described by Vallins as follows:

> Probably most people who mistakenly prefer *I* to *me* have at the back of their mind the glow of satisfaction they had, especially in school days, when they thought (just in time) to say 'It is I' instead of 'It's me.' (Vallins 1952: 27)

77 Emonds' assumption that the use of prescriptive *It is I* "is not part of a dialect spoken (and hence acquired) as a native language by any natural language speech community" (1986: 93) would imply that the historical developments up to 1600 represent 'unnatural' developments, or changes from above.

78 According to Quirk et al. (1985), different responses to *Who's there?* are *It's I* (formal) or *It's me./ Me.* (informal).

79 Joanne K. Rowling (2007) *Harry Potter and the Deathly Hallows*. London: Bloomsbury. p. 480.

A similar explanation has been offered by Lass:

> The pedagogical emphasis on the quite un-English nominative after be [...] has led to a common hypercorrection: the use of nominatives after prepositions (as for my wife and I, and so on). The origin is simple: people have been taught that me is 'bad', so they avoid it except where they can't possibly (nobody yet says *give it to I). (Lass 1987: 152)

8.3.1 Subject and subject complement function

All O-forms in the corpus – *me*, *him*, *her*, *us*, *them*[80] – appear in subject and subject complement function. Various examples will be shown at the end of this section, and more can be found in the pronoun spreadsheets. Table 8.3 lists all tensed occurrences, with limited significance where the overall number of occurrences is too small. No-verb utterances and question tags are discussed separately below.

Table 8.3: O-form proportions (%) in subject and subject complement function

syntactic function	*me* (of all 1SG)	*him* (of all 3SGm)	*her* (of all 3SGf)	*us* (of all 1PL)	*them* (of all 3PL)	TOT. %
subject	**0.02** (6/ 38472)	**0.03** (6/ 18640)	**2.1** (123/ 5913)	**0.6** (107/ 16859)	**0.2** (51/ 21722)	**0.3** (293/101606)
subject, coord. initial	**78.9** (45/ 57)	**50.0** (11/ 22)	**55.6** (5/ 9)	**0.0** (0/ 2)	**50.0** (1/ 2)	**67.4** (62/92)
subject, coord. final	**20.8** (11/ 53)	**85.7** (6/ 7)	**100.0** (4/ 4)	**0.0** (0/ 1)	**57.1** (4/ 7)	**34.7** (25/72)
subject, coord. middle	**0.0** (0/ 2)	**0.0** (0/ 0)	**0.0** (0/ 0)	**0.0** (0/ 0)	**0.0** (0/ 0)	**0.0** (0/2)
subject complement	**85.1** (40/ 47)	**80.0** (16/ 20)	**83.3** (5/ 6)	**70.0** (7/ 10)	**100.0** (5/ 5)	**83.0** (73/88)
subj.comp., coord.initial	**78.6** (22/ 28)	**100.0** (3/ 3)	**100.0** (1/ 1)	**100.0** (2/ 2)	**100.0** (1/ 1)	**82.9** (29/35)
subj.comp., coord.final	**14.3** (1/ 7)	**50.0** (1/ 2)	**100.0** (1/ 1)	**0.0** (0/ 0)	**100.0** (1/ 1)	**36.4** (4/11)
subj.comp., coord. mid.	**62.5** (5/ 8)	**100.0** (1/ 1)	**0.0** (0/ 0)	**0.0** (0/ 0)	**0.0** (0/ 0)	**66.7** (6/9)

80 Some studies have ruled out the use of specific O-forms in subject function. According to Ihalainen (1991), for example, *him* and *us* subjects do not appear in the Somerset dialect.

Table 8.3 includes all subjects and subject complements in the corpus (POS 10–13 and POS 30–33; compare Table B3). Not included are no-verb utterances (POS 14–16), question tags (POS 64) and disjunctive pronouns (POS 70–73). Also not included is subject *thee* (compare 2SGPL_you.xls). Among those *thee* examples which do not form part of a biblical citation, prayer, poem or direct speech produced at some point in the past, there are 43 simplex subjects, as opposed to 24 objects and 6 prepositional complements. *Thee* is obviously no longer exclusively objective.[81] Examples for *thee* in different functions are given below.

The obtained results show that O-forms are not very frequently used as simplex, non-coordinated subjects (mostly < 1%), but they contribute substantially to the number of coordinated subjects in the corpus. In coordination, each O-form accounts for at least 50% of all pronominal occurrences in its category (except *us*, due to the small overall number of 1PL cases). There is no doubt that coordination has a positive influence on non-standard case forms. At the same time, however, the results also show that coordination does not block standard case invariably,[82] giving no support to the assumption that " 'double bound' prevents constituents from being related by a transformational rule, so pronouns in conjoined subjects cannot be subjective in form." (Emonds 1986: 99, tree structure 14).

Looking at the figures in more detail, another interesting result is that coordinated *me* in final position has a relatively low frequency compared to other pronouns. Examples from the corpus indicate that speakers tend to prefer *and I* over *and me*, whereas in initial position *me* is largely preferred over *I*. One possible explanation for this tendency is that *and I* is treated like a frozen expression (compare 17.2; cf. Householder 1987; Redfern 1994; Honey 1995). However, the strong preference for *I* in second-conjunct position does not constitute an "X *and I* constraint" as suggested by Grano (2006: 36).

(139) I once dropped in a muck, **me** did there. (subject *me*, LanD1P, LAN_006)

(140) And many a time **me and** the boys have picked him up, sitting on the doorstep somewhere... (coord. subject *me*, initial, JG, LND_004)

(141) ...and **him what** owned that wood yard he used to own the cottage... (subject *him* followed by relative clause, NotA80, NTT_009)

(142) **Him and** I ain't been fishing for these last six weeks. (coord. subject *him*, initial, and final *I*, MdxCG, MDX_001)

81 Compare Wakelin (1986: 33), for the use of *thou* subjects, and *thee* subjects and objects, in the SW of England. The SED reports the use of *thee* as a stressed nominative in Somerset, Wiltshire, Cornwall, Devon and Dorset, the use of *thou* as an unstressed objective in Wiltshire, and the use of *thee* as an unstressed objective in the southwestern counties, except Cornwall.

82 Unless we attribute all S-forms in coordination to learned 'Prestige Usage' (cf. Emonds 1986).

(143) **Her** says, I can get two loaves of bread, **her** says, for sixpence halfpenny. (subject *her*, WigWHJ, LAN_020)

(144) ...well **us** used to be shoved out there Saturday afternoons and go pictures and when **us** come out of there first place **us** went was to the Island because the pictures we saw was cowboys... (subject *us*, TCA_FK, DEV_008)

(145) And he ʼd buy mi mother two love books, and **them** were a penny apiece. (subject *them*, LanC5P, LAN_005)

(146) I said, I ʼll have an echo-sounder. But Abey **and them** had wireless, which was better really... (coord. subject *them*, final, GCS, SFK_010)

(147) And, course when you turn them out while **thee**, while you want to see that there (reg sic=idn) isn't (/reg) no production there at all. (subject *thee* > *you*, SRLM_PT, SOM_007)

(148) {<u Int> How deep would that shaft have been?} About ten to twenty yards, course I suppose they ʼd had different gantries in, ʼcos, see, when **thee** gets the coal, **thee** puts a gantry in... (subject *thee*, IBJP, SAL_038, trba.)

(149) ... ʼcos if these hills wun steep, **thee** ʼd got to get **thee** feet in one o' these steps, and if **thee** missed one, **thee** could'st have the tub on top on **thee** if **thee** wasna strong enough to ʼowd it. (*thee* in different syntactic functions, IBJP, SAL_038)

(150) Well, if **thee** said'st anythin' wrong he ʼd soon tell **thee...** (subject and object *thee* in same sentence, IBJP, SAL_038)

(151) In the little pits -- they wouldna listen to **thee**. (prep. compl. *thee*, IBJP, SAL_038)

In subject complement function, the results show a clear preference for O-forms which ties in nicely with the historical development described above. The Modern English data in FRED show how rare S-forms have become in this syntactic function, making the prescriptivist call for S-forms untenable. As Labov (1996: 84) stated in his ʼprinciple of validity': "When the use of language is shown to be more consistent than introspective judgments, a valid description of the language will agree with that use rather than with intuitions."

The absolute number of occurrences in Table 8.3 may be on the low side for some pronouns, but the results nevertheless show that O-form subject complements are highly frequent in both coordinated and simplex NPs, leaving little space for variation. It is interesting to see that S-forms, despite their preferential treatment in prescriptive grammars, are an exception. Their occasional appearance nevertheless contradicts claims which expect the invariable use of O-forms to the right of an inflected verb.[83]

(152) And there was **me** in mi cut-down khaki trousers... (subj. compl. *me*, DurNB, DUR_002)

83 Emonds' rule 11 for ʼNormal Usage', for example, requires pronouns to the left of Infl, but not to the right of Infl, to have subject form: "Pronoun–Inflection → [Pronoun Subject]–Inflection" (cf. Emonds 1986: 97).

(153) Any-rate, uh, it was **me and** my two brothers... (coord. subj. compl. *me*, initial, CA, LND_001)

(154) ...and there was **me and him**, and we were at this guard meeting... (coord. subj. compl. *me*, initial, and *him*, final, WesBR, WES_008)

(155) And there ʼs **her** (v ʼinterviewer clears throat') **and** Mr Simpson, they ʼd be drinking beer all the time... (coord. subj. compl. *her*, initial, CA, LND_001)

(156) It wasn't **us**, we was in such a place we didn't bother. (subj. compl. *us*, WigMB, LAN_017)

8.3.2 Coordination

The general relaxation of case assignment in coordinate structures has, for example, been observed by Parker, Riley and Meyer (1988; 1990). The authors suggest that because government, and thus case assignment, is blocked by the NP dominating the coordinate structure, any form of personal pronoun can occur. This includes so-called untriggered reflexives, i.e. independent self-forms, which "in coordinate constructions are not anaphors at all but, rather, alternative forms of personal pronouns." (Parker, Riley and Meyer 1990: 55).

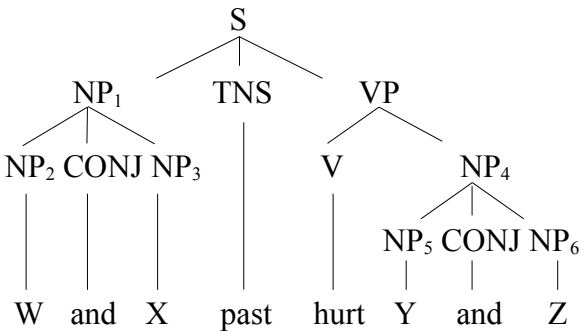


Figure 8.3: Tree structure for untriggered reflexives in co-ordinated NPs

Interestingly, case assignment is not the only mechanism to relax in coordination. In her study on Belfast English, for instance, Henry (1995: 18) found that coordination has an influence on both pronoun choice and verbal agreement. Henry's results show that "singular concord is impossible if the subject is a simple personal pronoun," but "pronouns which are part of a co-ordination can have a singular verb, provided that they are not nominative." In Henry's study (p. 23–24), the case and verb combinations shown in

sentences (157a)–(157c) are hence acceptable whereas the coordinated subject forms in (157d) can not be followed by a singular verb.

(157) a. He and I are going.
b. Him and me are going.
c. Him and me is going
d. *He and I is going.

In the future, this apparent correlation between coordination, O-form subjects and singular concord would certainly be worth investigating in more detail and other varieties of English.

8.3.3 Question tags

In the literature, examples used to illustrate O-form subjects usually consist of tensed S–V structures. In the corpus, non-standard O-forms also appear in a variety of other structures, including no-verb statements and questions, question tags and disjunctive NPs (for the latter see 12.5). In order to detect any systematic differences between these structures, they were analysed separately. The results show that O-form frequencies do, indeed, vary significantly depending on the syntactic context.

In question tags, S-forms are the clear default option (512 cases, 95%), while O-forms are only used occasionally (27 cases, 5%; excluding *you*; no self-forms).[84] The non-standard proportion here is higher than for simplex subjects in tensed clauses, but significantly lower than in subject complements or coordinated subject NPs. The fact that not all O-forms appear in this context may be coincidental. Examples were found for *him*, *us* and *them*; should there exist a lexical constraint, it is not clear at this stage what that constraint should be.

(158) ...the first lad as I seen down on this green was in was in with pony, and he -- used to have a pony and run round, you know, **didn't him?** (*him* as question tag subject, FP, OXF_001)

(159) <u SRLM_HW> [...] We had acres of eggs. <u W> Wheeling and dealing, **weren't us?** (v 'laughter') (*us* as question tag subject, SOM_030)

(160) ...everybody seemed happy **didn't 'em** (mname) Fred (/mname)? (*them* as question tag subject, BF, OXF_001b.wav, trba.)

84 Question tags with *do*, *can*, *should*, *would*, *have*, *be*, and corresponding negative and inflected forms, incl. *ain't*, *won't*, and variants like *wun't*, *din't*, *hent*. No instances were found of third person enclitic O-form subjects of the type *Don' er?* 'Doesn't she/ he?' (cf. Ihalainen 1994: 216).

(161) ... ʻcourse they had to get these carrots for a lot of troops, **hadn't ʻem?** (*them* as question tag subject, WflsWGP, WIL_010)

8.3.4 No-verb utterances(1)

The distribution of S- and O-forms in question tags differs very clearly from no-verb utterances such as (162)–(176).[85] While the pronouns in question tags are mostly S-forms accompanied by occasional O-forms, no-verb utterances show the opposite tendency. In the corpus, we find O-forms in single-pronoun questions (POS 14),

(162) {What, what politics did you take up, when you..?} **Me?** Oh, I, I didn't used to trouble about anything much. (*me* in no-verb question, KentTAD, KEN_005b.wav)

in no-verb answers with simplex or coordinated pronominal NPs (POS 15 and 16),[86]

(163) {How many of you were working there?} Ehm **me and** a man, a horseman **and me**.[...] {Did the farmer's wife do all the cooking and --} Oh yes, **her and** t' girl, yes, we used to bake in a a brick oven. (coord. *me* and *her* in no-verb answers, WesBR, WES_008d.wav)

(164) {Would you yourself hunt them back or did you get somebody?} No, **yourself,** after you buy them it 's you it 's up to. (subject *yourself* in no-verb answer, SRLM_HW, SOM_030)

(165) {Who would milk the cows?} Oh, **I and** the wife (v ʻlaughter'). (coord. *I* in no-verb answer, SRLM_Y, SOM_031)

and in no-verb utterances with *just*, *only*, *not* or *too* which express some sort of contrast (also POS 15 and 16; no-verb objects are not considered in the analysis but are included in the examples):

(166) {So it it was just you and her, actually?} Yeah. **Just me and her**. {Yes.} Yes. {You do, you do (unclear) all (/unclear) -- } Yes, oh yes, **just me and her**. (no-verb answer with *just*, LanW1P, LAN_004e.wav)

(167) {Is that is that what her duties were, like looking after the house and looking after the kids and...} Yes, yes. She had nothing to do with the pub, **just us**. (no-verb object *us* with *just*, YksHS, YKS_006b.wav)

(168) {Were there some other lads living in the farm too?} No **only me**, and this here labourer. (no-verb answer with *only*, NotA71, NTT_008b.wav)

(169) I don't know who your mother is! I says, You do, says, The woman with the long plaits. Oh, says, **Her**. I says, That 's right. She says, Is she a fortune teller? I says, No, **not her**, that 's my mother. She had such long plaits she would sit on. (no-verb subject *her*, NotA8, NTT_013f.wav, trba.)

85 Excluding direct-speech and turn-taking markers such as *And she, Well*,... (LAN_002).

86 No-verb answers such as *Who did that? – Her*. have long been noticed as one of different disjunctive O-form environments (e.g., Wright 1905).

(170) {So, sorry, you, you said you had, when you moved Campbell Road, you had nothing. I think you said your mum pushed you up there in a, in a bassinet?} **Not me**. {Ahah.} ʻCause I was big enough to walk, but the babies, you know, **not me**. (no-verb object *me* with *not*, JG, LND_004d.wav)

(171) Yes, **me too**. (no-verb answer with *too*, TrbrD1, WIL_005)

In some texts, we find O-form subjects in no-verb statements such as (172), or independent self-forms in picture pronouns such as (173).[87] Finally, pronoun variation is observed in pronouns which are followed by locative adjuncts, as seen in (174) and (175):

(172) In fact I camped at Ayton with Norman when he was a young lad of fourteen years old. **Him and** a load more of us in tents. (no-verb coord. subject *him*, YksWF, YKS_001a.wav)

(173) {Can you tell me anything about W.M.S.52/9 please? [photograph]} Now, this is a group of drivers in their winter uniform I don't know many on but the fitter in the centre was Jack Smith the driver behind Bill Roscoe. {He is third on the left on the back row.} [...] Harry Butler which in time became the Chief Inspector. {Back row first right.} **Myself**. {Middle row first right. That was in the thirties. Can you tell me anything about W.M.S.52/10 please? [another photograph]} Now this is the summer uniform. And this was taken in front of the depot. Lionel Lee he was a driver. {First left.} **Myself at** the back... (*myself* picture pronouns, no verb, WigWO, LAN_023)

(174) We we was we was slitted, so was eh Wall End, you know, and **him at** eh over at eh, where 's that eh Grasmere eh he was slitted both ears. (no-verb subject *him* with locative adjunct, WesBR, WES_008c.wav, trba.)

(175) {And who 's the driver on the other engine, was it --} Mr Fred Moores, the one that 's on that there little traction, **he on** that one there. (no-verb subject *he* with locative adjunct, picture pronoun, SRLM_RM, SOM_014b.wav)

(176) We had three teachers, Miss Routledge the head, Miss Thomas the second, and Miss Thackeray the third one who taught the small ones. **Myself** to start with. (no-verb object *myself*, WesDT, WES_016)

Overall, O-form subjects are the default in no-verb contexts. They are used in 78.3% of all pronominal cases whereas S-forms only cover 4.3%. In addition, self-forms account for 17.4% of all no-verb pronouns (compare Table 8.4; self-form examples in (164), (173) and (176)). In combination, O-forms and self-forms in no-verb utterances account for the highest non-standard proportion of all syntactic functions analysed in this study, with

87 Picture pronouns are not discussed separately in this study. Without video documentation, it can be difficult to determine the situational context as an analysis parameter. Picture NPs are a phenomenon where the situational context plays a decisive role in pronoun resolution, due to the simultaneous presence of speaker, hearer and referent (e.g. if the picture lies before the conversation participants), but also paralinguistic means such as pointing (at the referent in the picture). The term was probably first used by Florence Warshawsky (1965) *Reflexivization*, mimeographed, MIT. For discourse-analytic approaches to picture pronouns see Cantrall (1973: 46–47; 1974: 22) and Kuno (1987: 166). For an experimental study of third-person picture NPs see Goldwater and Runner (2006). Compare Stern (2004: 276–277) for an explanation based on the *Role Conflict Hypothesis*: "the referent of a pronoun used for a logophoric message plays at least two roles in an utterance: the role for which the referent is mentioned, and as the cognizer of the situation. Zribi-Hertz calls this role the Subject of Consciousness (SC)."

97.4% in simplex cases and 93.5% in coordination. With respect to the individual pronouns, *him* is the only 3SGm form in no-verb utterances except for two examples with *he*; *her* the only 3SGf form; *us* the only 1PL form; and *them* the only 3PL form. In 1SG cases, the speakers used either *me* or *myself*, the only exception being the one example with coordinated *I* in (165). Based on the underlying speaker distribution pattern (see 5.2), O-form subjects in no-verb utterances classify as a generalised feature: SN = 32, SP = 15.2%, MCR = 1.7.

Table 8.4: Case form distribution in no-verb statements and questions[88]

no-verb	simplex				coordinated			
	S-form	O-form	self-form	TOTAL	S-form	O-form	self-form	TOTAL
1SG	0	21	7	28	1	19	5	25
3SGm	1	3	0	4	1	1	0	2
3SGf	0	4	0	4	0	4	0	4
1PL	0	1	0	1	0	0	0	0
3PL	0	1	0	1	0	0	0	0
TOTAL	1	30	7	38	2	24	5	31

A comparison of no-verb utterances and tensed clauses provides additional insights on the influence of verb absence and coordination on pronoun choice. While the difference in frequencies between coordinated no-verb vs. tensed cases is not significant (1.2 σ, $p = .21$), it is highly significant for non-coordinated cases (5.9 σ, $p = 3*10^{-9}$).[89] The influence of verb absence is thus only noticeable in the absence of coordination (a similar result will be described for disjunctives in 12.5). From a wider perspective, this has far-reaching implications for a theory of case assignment in so-called default case environments, indicating that case in coordination is primarily influenced by coordination, irrespective of the presence or absence of the verb.

88 Including comparative constructions with verb elision, as in *No, him I was going to have, not her* (LAN_004), or *She 's told me that, not him* (YKS_011). Excluding *you* (1 case: *Jack, Nell, and you* (LAN_002)) and *yourself* (see example 164); *all them*; and demonstrative or qualified *them*, as in *Just them that were no use* (WES_011).

89 Simplex: no-verb 30 out of 38, 78.9%, vs. tensed 366 out of 101694, 0.48%. Coordinated: no-verb 24 out of 31, 77.4%, vs. tensed 126 out of 221, 57%.

8.3.5 Relative clauses and clefts

The last syntactic context analysed with regard to subject and object case variation in this section are pronouns which are followed by a relative clause. This context, too, has long been discussed in the literature. Wright's *English Dialect Grammar* (1905), for example, already included the use of O-forms in cleft sentences such as *it was her that did it*, and the use of O-forms before relative pronouns, as in *him that did that ought to...* . Among the more recent studies, Beal (2004: 119), for instance, has mentioned the use of third-person non-standard subjects in emphatic utterances such as *You know, her that's always late* in the Newcastle Electronic Corpus of Tyneside English.[90]

In the FRED corpus, speakers generally prefer O-forms before relative pronouns (see relativeclauses.xls).[91] In non-cleft occurrences, there is only 1 S-form, shown in (177), as compared to 13 O-forms (*him*, *her*, *us*), 8 of which have subject function. In cleft sentences, we get a similar distribution: 2 S-forms (*he*) as compared to 10 O-forms (*me*, *him*, *her*, *us*). No areal patterns could be detected due to the small number of occurrences.

(177) But, **we that** thought we were going to have to ... (subject *we* before relative clause, SRLM_JA, SOM_029)

(178) The only ones as needed water was the fire beater, **him as** used have to look after t' furnace, he needed his water... (subject *him* before relative clause, WigWHJ, LAN_020)

(179) Of course **us what** could see a bit when it was time to go in, we used to get uh anybody with your hand... (subject *us* before relative clause, NotA58, NTT_006)

(180) And **it was he who** also kept the Park in front of the Castle... (*he* in cleft sentence, WesCB, WES_010)

(181) ...she 's supposed to be living with Gertie, **that 's her that** was forewoman at eh Standfast... (*her* in cleft sentence, H3L, LAN_010)

(182) ...all the family have worked in pubs, **there is only me that** hasn't worked in pubs. (*me* in cleft sentence, DurEL, DUR_003)

8.3.6 Context hierarchy

Based on the O-form frequencies in the different subject functions analysed above, a graded context hierarchy can be established. This hierarchy goes from the highest frequencies in subject complement function (left) to the lowest frequencies in non-coordinated subjects (right). The empirically derived order lends itself to future comparisons with other datasets.

90 Information on the NECTE corpus is available on http://www.ncl.ac.uk/necte/.

91 Including relative pronouns *as, that, what, which, who, whom*. Excluding prayers and bible citations, as in *He who helpe us rise from our warm bed in the morn'/ It's He who guides us and all the creatures* (WIL_009). Excluding *they* and *them* in demonstrative function.

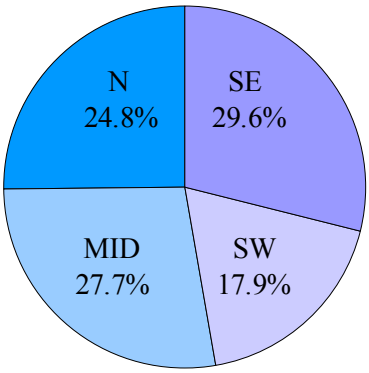
(183) Context hierarchy of non-standard object forms

subject complement > no-verb > coordinated subject > question tag > simplex subject

8.3.7 Areal distribution

Object forms are the supraregional default in subject complement function and no-verb utterances. In the latter context, the following numbers of O-forms were found in the four dialect areas: 10 SE (100%; O-forms in all no-verb cases), 8 SW (50%; plus 3 S-forms and 5 *myself*), 8 MID (88.9%; plus 1 *myself*), 28 N (82.4%, plus 6 *myself*).

In subject complement function, O-forms clearly dominate over S-forms in all dialect areas, representing a showcase general feature of spoken English. Their areal distribution is visualised in Figure 8.4, based on relative weighted frequencies (compare subjectcomplements.xls). The additional information on the right-hand side shows the number of standard deviations between the non-standard frequencies of different dialect areas, where $\Delta_{(SE, SW)}$ simply is the difference between the proportion of O-forms in subject complement function in the Southeast as compared to O-forms in subject complement function in the Southwest. At a minimum confidence level of $p < .05$, the differences in frequency between the four areas are not significant.



$\Delta_{(SE, SW)} = 3.5\ \sigma \quad (p = .08)$
$\Delta_{(SW, MID)} = 2.9\ \sigma \quad (p = .14)$
$\Delta_{(SE, MID)} = 0.5\ \sigma \quad (p = .79)$
$\Delta_{(SW, N)} = 2.1\ \sigma \quad (p = .29)$
$\Delta_{(SE, N)} = 1.3\ \sigma \quad (p = .50)$
$\Delta_{(MID, N)} = 0.8\ \sigma \quad (p = .68)$

Figure 8.4: Areal distribution: O-form subject complements (*me*, *him*, *her*, *us*, *them*)

Unlike the areal distribution of S-form objects shown in Figures 8.1 and 8.2, O-form subjects are distributed evenly across the four dialect areas, representing a supraregional phenomenon.[92] Figure 8.5 shows the individual charts for all O-forms, as well as the overall distribution of O-form subjects across the four areas in the bottom right graph. The values are based on relative weighted frequencies for all O-form proportions in subject function, since the absolute number of pronominal subjects varies for each pronoun and dialect area.

The small graphs illustrate the general supraregionality of the individual pronouns. It can, for example, be seen that the distribution of *her* in FRED roughly tallies with the corresponding SED findings. In both datasets, *her* subjects are characteristic of the SW and MID areas. However, *her* in FRED is not confined to these areas, whereas the SED findings suggest its use west of a line running from Portsmouth to south Derbyshire (SED question IX.7.7.3 'she is').[93]

A clear predominance of SW cases could only be confirmed for *us*. The distribution is, however, attributable to an underlying speaker distribution pattern with two outliers. Two of the SW speakers have exceptionally high O-form proportions in an almost 50–50 distribution of S- and O-form subjects: speaker TCA_EA used 69 *we* vs. 63 *us*, and speaker TCA_FK used 24 *we* vs. 21 *us*. If we take these two speakers out of the equation, the overall proportion of *us* in the SW drops from over 77% to under 1%, and the differences in frequencies between the SW and the other areas become statistically insignificant. The influence of these two outliers should also be kept in mind regarding the overall distribution depicted in the bottom right graph.

The most surprising result in this particular analysis is perhaps the areal distribution of O-form subjects in question tags. As was observed above, this syntactic context has default S-forms, but O-forms are used occasionally. Now, a comparison of the four dialect areas shows that all O-forms in question tags appear in data from the SW. There are 27 O-form

92 The supraregional distribution of O-form subjects in FRED contradicts, for instance, Trudgill (2004a: 148), who stated that "[t]he evidence of [Charles Benham's *Essex Ballads*] and of the SED records suggests that in southern East Anglia the phenomenon was more restricted than in the southwest. The southwestern usage of *him, her, us* as subjects does not seem to have been a possibility; we witness merely the use of *he, she, we, they* as objects."

93 The results presented in this study give no indication that the occurrence of *her* subjects is historically motivated. Wagner (2002), for example, found that *her* and *us* were the two most frequently used O-forms in her data. According to Wagner (ibid., fn. 21), "[a]t least for the feminine form, this can very well be historically motivated. While regions under strong Scandinavian influence adopted the new form *she* quite readily to have two clearly distinct forms for the subject and oblique cases, the Southwest could have kept the old *h*-form, modern PE thus representing a kind of relict form/ historical retention/ conservatism." The idea that the Southwest of England, as a conservative area, may still be preserving a preference for the older *h*-form (OE *hēo*) does, however, not tally with the fact that all occurrences of object *she* in FRED were found in this particular area.

cases in the SW, which equals 12.8% of all SW cases. Even the most common O-form, *them* (18 cases, over 17.6% of all 3PL cases in the SW), appears exclusively in this geographic area. Hence, the distribution of non-standard *them* in question tags is very different from the distribution of non-standard *them* in other syntactic contexts, for instance in tensed S–V structures where it is distributed quite evenly across the country.

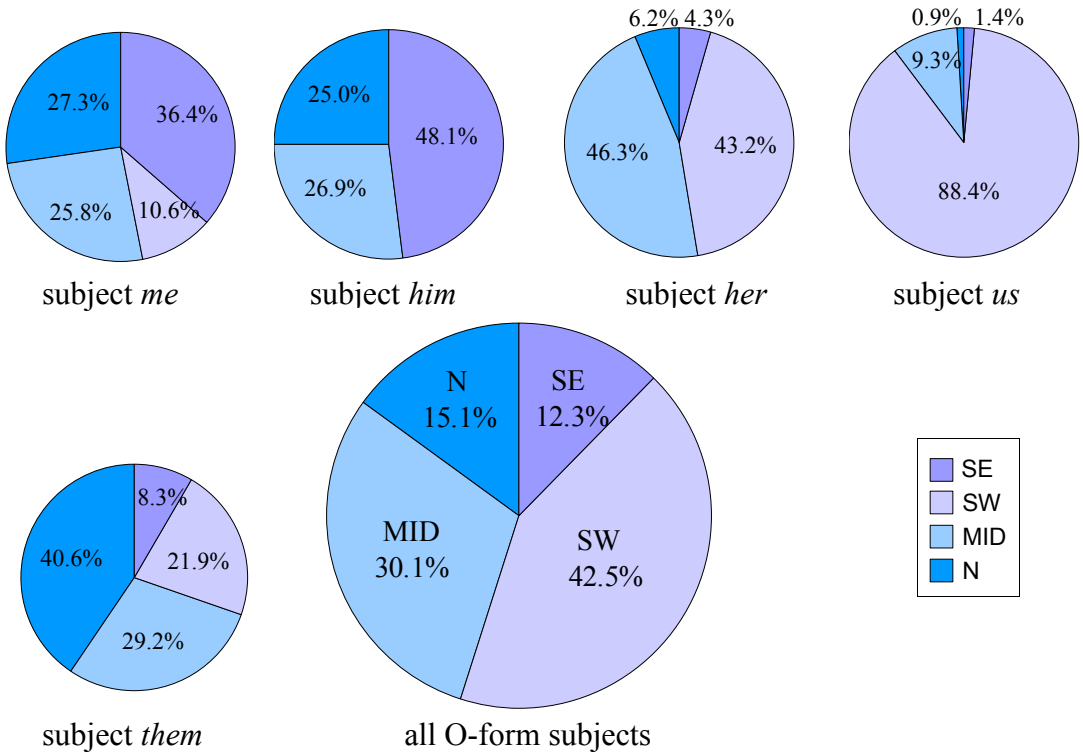


Figure 8.5: Areal distribution: O-form subjects (*me*, *him*, *her*, *us*, *them*)

8.4 Personal pronouns in reflexive function

As mentioned in 8.1, PE is traditionally described as a binary feature involving subject and object case forms. However, occurrences of personal pronouns in reflexive function, on the one hand, and the use of self-forms in different subject and object functions, on the other hand, show that the 'exchange' takes place between three, not two, pronominal categories (see Figure 8.7).

The previous chapters have shown that the exchange between S-forms and O-forms is bi-directional. In the following, we will see that this also applies to O-forms and self-forms. With S-forms, on the contrary, the exchange seems to be uni-directional, given that no occurrences were found of reflexive *I*, *he*, *she*, *we* or *they*. The situation in FRED is

reminiscent of earlier stages of English, where reflexivity could be expressed by O-forms but not S-forms (compare 3.1). Typologically, the situation in FRED tallies with the fact that identical reflexive and object forms are attested in various languages[94] whereas "no nominative anaphors" appears to be "a pretty robust tendency" (Kiparsky 2008: 31).

(184) Mind you I was a bit on the safe side, I put a rope round **me** just, to tension up... (reflexive *me*, YksWF, YKS_001)

(185) Well, Boss lived in the houses at the back. He used to use this towel the one week and the next week he had the tail of a shirt for drying **him** on. (reflexive *him*, IBFB, SAL_013)

(186) No, he said, That isn't me. Now he couldn't recognise **him** [his own voice], he wouldn't have it! (reflexive *him*, CAVA_WW, CON_006)

(187) But Harold, he was a first bowler, and **he got him** a trial with fourteen. (reflexive *him*, LanEG, LAN_012)

(188) ...she 'd, she 'd got hold -- to try to save **her**, she 'd, she 'd got hold of this rail and it broke. (reflexive *her*, LanD1P, LAN_006)

Among the O-form reflexives found in the corpus, there are 6 examples with *me*, 3 *him*, 2 *her* and *you*, 1 *us* and *them* (see Table 8.5; reflexives.xls). Although the numbers seem low, O-forms account for almost 4% of all reflexively used pronouns in the data. In about half of the non-standard examples, the pronoun follows a verb which is intransitive in Standard English, such as *wash*, *change* or *turn* (compare ch. 12.2). Occurrences with so-called benefactive pronouns, as in *I was having me a drink*, are not included in Table 8.5 but will be discussed separately in 12.3.

Table 8.5: O-form reflexives: absolute and relative frequencies

O-form reflexives		frequency	%
me	(of all 1SG reflexives)	6/ 116	5.2
you	(of all 2SG/ PL reflexives)	2/ 87	2.3
him	(of all 3SGm reflexives)	3/ 56	5.4
her	(of all 3SGf reflexives)	2/ 29	6.9
it[95]	(of all 3SGn reflexives)	0/ 16	0.0
us	(of all 1PL reflexives)	1/ 46	2.2
them	(of all 3PL reflexives)	1/ 53	1.9
TOTAL (of all reflexives)		15/ 403	3.7

94 For example Frisian. According to Kiparsky (2002: 203), "[m]any languages lack reflexive pronouns entirely and simply use personal pronouns in their place."

95 No reflexive *it* were found in the texts selected for the investigation of gendered pronouns. Additional word searches confirmed their overall absence from the data.

In terms of areal distribution, there are not enough O-form reflexives for any definite conclusions, but at least one example was found in each dialect area (1 SE, 3 SW, 6 MID, 5 N).

This is ourselves, under pressure.
(Queen and David Bowie "Under Pressure")

8.5 Independent self-forms

Self-forms like *myself*, *yourself*, and so forth, represent the third type of variants involved in pronoun exchange. The current section focuses on their use in different subject and object functions, including subject and prepositional complement positions, as well as no-verb utterances. Self-forms which fill one of these functions are neither reflexive nor should they be confused with standard intensifiers (compare 3.3, (37)). Different terms have been proposed in the literature, including 'absolute', 'unbound' or 'untriggered reflexives'. We will here avoid the term 'reflexive' altogether and refer to these occurrences as *independent self-forms*.[96]

In the literature, independent self-forms have been described as characteristic of Irish English (cf. Filppula 1999, 2004; Bliss 1979; Harris 1993), Scottish English (Miller 1993), and American English vernaculars (Evans and Evans 1957; Heacock 2008[2]). This study describes their distribution in England, where they represent a general spoken feature.[97]

The Modern English examples from FRED are historically linked to earlier stages of the language, as was shown in 3.1. It appears that, since Old English times, writers have never stopped using independent self-forms. They even appear in texts from the 18th century, a time when the ground was laid for many normative rules still respected today. Tieken-Boon van Ostade (1994: 220) has described the common use of 'non-reflexive -*self* pronominals' in the personal letters of various men and women of the time, both in subject and object functions, and especially in coordination. Similarly, Filppula (2004: 93) has observed for Irish English that "although this feature is mainly found in vernacular and

96 Compare Kiss (2001), for 'exempt anaphora', i.e. anaphora which are exempt from binding principle A like personal pronouns. Compare Huddleston and Pullum (2002: 1484–1485), for 'override reflexives'.

97 See Hernández (2002), for a context hierarchy of independent self-forms based on the BNC and NITCS corpora. Another study by the author (MA thesis, unpublished) showed that independent self-forms were more frequent in spoken English than in literature or press texts, more frequent in corpora from the 1990s (FLOB, Frown) than 1960s (LOB, Brown), and more frequent in British (LOB, FLOB) than American press texts (Brown, Frown). An analysis of the spoken component of the BNC corpus and a supplementary questionnaire especially designed for the study showed that independent self-forms were used by different age-groups, their frequency rising parallel to the speaker age.

colloquial styles, occurrences can be spotted even in 'educated' varieties, including written language."

In their independent uses, self-forms are best understood in semantic or pragmatic terms, since they elude explanations of rule-based sentence grammar. Among the different explanations proposed in the literature are emphasis and contrast (cf. Kuno 1987; Zribi-Hertz 1989), and politeness and modesty (cf. Filppula 1999; Parker, Riley and Meyer 1988; Baker 1995).[98] Another possible explanation seems to be that speakers who feel uncertain about the correct choice of pronoun case, for example in coordinated NPs such as *me and him*, choose self-forms as a neutral alternative (cf. Emonds 1986: 119).

In discourse grammar there is a shared conviction that "a grammatical theory of English reflexive pronouns cannot be complete without a discourse component" (Zribi-Hertz 1989: 703). Most explanations evolve around the concepts of speaker viewpoint (cf. Cantrall 1974: 94, 'theory of incorporated identification'), empathy (Kuno 1987: 29, "syntactic manifestations of the speaker's camera angles"), and logophoricity (Milroy and Milroy 1993: 147, "the reference draws on the shared knowledge of the speaker and hearer"). Filppula (2004) mentions the 'topic reading', a related concept which is applicable to most FRED examples, especially when an emphatic interpretation is not indicated by stress in the recording: "an absolute reflexive is often used with reference to that person or those persons who constitute the 'topic' of the conversation in some way or another" (p. 93).

Most independent self-forms, including most occurrences in FRED, conform to one of these explanations. In this study, independent self-forms are, therefore, not regarded as mere shortenings of underlying 'personal pronoun + self-form' combinations (*you yourself* → *yourself*), nor do they automatically correspond in meaning to complex intensified NPs (*you yourself* = *yourself*).[99]

Table 8.6 shows the absolute and relative frequencies of independent self-forms in different syntactic functions in the corpus. This overview helps detect patterns of context-specific behaviour, as well as individual high non-standard frequencies which are not

98 Also compare Kuno's 'modesty principle' (1987: 233). Note that a classification of independent self-forms within word field or lexical field theory proves difficult. According to Sankoff (1978: 24), the rule of thumb is that "a word can be used to express any one of its most general meanings [...] as well as any more specific meanings [...] where 'more specific' indicates that additional features are conjoined. A word may not be used for a meaning more general than one of its prime implicants." The most general meaning of independent self-forms is simple referentiality (non-reflexive and non-anaphoric).

99 I argue against the view that self-forms in argument positions are really intensifiers attached to covert or incorporated pronominal heads (cf. König and Siemund 2000; Baker 1995). The problem of such a definition lies in utterances where the independent self-form does not function as an intensifier or contrastive expression (this problem has been recognised by König and Siemund).

apparent from the overall total obtained for each pronoun. The results will be interpreted on the following pages. Note that no independent self-forms were found in ECM or for–to constructions, although their use seems intuitively possible.

Table 8.6: Self-form proportions (%) in subject and object functions

syntactic function	*myself* (of 1SG)	*yourself/-selves* (of 2SG/ 2PL)*	*himself* (of 3SGm)	*herself* (of 3SGf)	*ourselves* (of 1PL)	*themselves* (of 3PL)	TOT. %
subject	0.0 (0/38472)	0.0 (0/29977)	0.0 (0/18640)	0.0 (0/5913)	0.0 (0/16859)	0.0 (0/21722)	0.0 (0/131583)
subject, coord.	6.3 (7/112)	0.0 (0/4)	3.4 (1/29)	0.0 (0/13)	33.3 (1/3)	0.0 (0/9)	**5.3** (9/170)
subject complement	2.1 (1/47)	0.0 (0/4)	0.0 (0/20)	0.0 (0/6)	0.0 (0/10)	0.0 (0/5)	**1.1** (1/92)
subj. compl., coord.	25.6 (11/43)	0.0 (0/0)	16.7 (1/6)	0.0 (0/2)	0.0 (0/2)	0.0 (0/2)	**21.8** (12/55)
no-verb subject	25.0 (7/28)	100.0 (1/1)	0.0 (0/4)	0.0 (0/4)	0.0 (0/1)	0.0 (0/1)	**20.5** (8/39)
no-verb subj., coord.	20.0 (5/25)	0.0 (0/0)	0.0 (0/2)	0.0 (0/4)	0.0 (0/0)	0.0 (0/0)	**16.1** (5/31)
object	0.1 (1/1915)	0.04 (1/2278)	0.0 (0/1949)	0.1 (1/754)	0.1 (1/945)	0.02 (1/5406)	0.04 (5/13247)
object, coord.	6.3 (1/16)	0.0 (0/0)	0.0 (0/1)	0.0 (0/1)	0.0 (0/2)	0.0 (0/4)	**4.2** (1/24)
prepositional compl.	1.1 (13/1210)	0.2 (2/821)	0.1 (1/941)	1.0 (4/407)	0.4 (2/565)	0.1 (2/1418)	**0.4** (24/5362)
prep. compl., coord.	0.0 (0/21)	0.0 (0/0)	0.0 (0/4)	0.0 (0/1)	0.0 (0/1)	0.0 (0/3)	0.0 (0/30)
TOTAL	0.1 (46/41848)	0.01 (4/33085)	0.01 (3/21590)	0.07 (5/7097)	0.02 (4/18387)	0.01 (3/28569)	0.04 (65/150576)

*extrapolated from sample, see 4.2

8.5.1 Subject function

The exchange between self-forms and S-forms has probably always been uni-directional. In the data, self-forms appear in both subject and subject complement roles, although their use seems restricted to coordinated NPs. The complete absence of simplex cases tallies with observations found, for example, in the OED, where the use of non-coordinated independent self-forms is referred to as archaic or poetic. The latest OED examples date from the 19th century. They include the poetic use of *myself* after *be*, as in *What am myself* (year 1864); subject *himself*, archaic, as in *The dagger which himself Gave Edith* (1864); subject *herself* in Welsh and Gaelic (also to ridicule these varieties), as in *Herself would... seat her down upon some linden's root* (1814); subject *ourselves*, as in *Ourself learnt this*

craft of healing. *Were you sick, ourself would tend upon you.* (1847); and *themselves*, archaic or poetic, as in *People's timorousness shows how insecurely grounded themselves are.* (1853). Even the *Dictionary of Contemporary American Usage* (Evans and Evans 1957) – according to which self-forms are established beyond question in speech and in literature in absolute constructions and after linking verb *be* – deems their use as simplex subjects artificial. Going even further, Saha (1987: 232), in his study of reflexivisation strategies in American English, claims that "a reflexive [i.e. self-form] can never be the subject of a tensed verb."

In other varieties such as Irish English, independent self-forms are more readily accepted in subject function. This is, for example, reflected in their regular use in Irish poetry, as seen in the *Songs of the Glens of Antrim* by Moira O'Neill in (189).[100]

(189) Och, when we lived in ould Glenann, Meself could lift a song! (*A Song of Glennan*)

Meself began the night ye went, An' hasn't done it yet... (*Forgetting*)

Herself 'ud take the rush-dip an' light it for us all [...]
Himself 'ud put his pipe down, an' say the good word more... (*Grace for Light*)

Let us take a look at the distribution in FRED. As mentioned above, independent self-form subjects only appear in coordination (see Table 8.6). Even if examples are rare, there are some observable tendencies. First of all, *myself* seems more likely to occur as a coordinated subject than other self-forms (7 *myself*, 1 *himself*, 1 *ourselves*).[101] Secondly, the component structure of the coordinated NPs shows that all self-form subjects occur in non-initial position.[102] In terms of geographical distribution, the number of cases is too small to disclose any patterns, but examples were found in all four dialect areas.

(190) I expect Mr Boobyer **and myself** have been in every rhyne there is in the moor that belongs to the Drainage Board. (coord. subj. *myself*, final, SRLM_CK, SOM_004)

(191) No, my younger brother **and myself** was his favourites. (coord. subj. *myself*, final, CA, LND_001)

(192) ...this Fred (trunc) mis- (/trunc) **and miself** mi brother and Art we used to get sixpence on a Saturday, and we used to walk to the trams. (coord. subj. *myself*, middle, NotA99, NTT_011)

(193) I know as a boy I 've seen them around here, my granny **and himself** would get a big chunk of beef... (coord. subj. *himself*, final, CAVA_WW, CON_006)

100Moira O'Neill (1933) *Collected Poems of Moira O'Neill*, Edinburgh/ London: William Blackwood & Sons. http://xtf.lib.virginia.edu/xtf/view?docId=chadwyck_ep/uvaGenText/tei/chep_3.2381.xml.

101No manifest correlation with the overall number of pronominal subjects.

102It is unclear how these tendencies relate to Kuno's Word Order Empathy Hierarchy, which says that "it is easier for the speaker to empathize with the referent of a left-hand NP in a coordinate structure than with that of a right-hand NP." (1987: 232).

(194) ...there was a place put on one side walled in where all the neighbours **and ourselves** used to put all our ashes... (coord. subject *ourselves*, final, IBCB, SAL_005)

Emphasis seems to be a possible explanation for occurrences like (190)–(194), but in a slightly different sense than suggested in the literature. I want to propose the following idea. It is commonly accepted that conjunction naturally emphasises the difference between two or more entities involved, whereas plurality emphasises the similarities between them. Hence:

(195) a. John, Bill and Tom have gone home. (conjunction: different boys)

b. The boys have gone home. (plurality: all boys)

These two examples were taken from Bhat (2004: 95), according to whom "[o]ne cannot use conjunction if no difference is indicated […] whereas one cannot use plurality if differences cannot be disregarded." If conjunction, by its very nature, emphasises the differences between conjunct elements, and if the use of self-forms is generally associated with emphasis, this explains the speaker's predilection for self-forms in coordination. It also explains why, in many cases, the self-form itself carries no extra stress; it is a variant that simply fits into the natural semantics of conjunction.

8.5.2 Third person cases

Special attention has been given to third person independent self-forms in the literature. Levinson (1997), for example, has deemed them unacceptable, similar to Jespersen before him. An explanation is given in Levinson's 'performative hypothesis', which states that:

> every sentence has as its highest clause in deep or underlying syntactic structure a clause of the form ['I (hereby) V_p you (that) S'] – i.e. a structure that corresponds to the overt prefix in the explicit performative, whether or not it is an overt or explicit performative in the surface structure. (Levinson 1997: 247)

Thus, sentence (196a) is judged acceptable, since it derives from (196c), whereas (196b) is judged unacceptable.

(196) a. Solar energy was invented by God and myself.

b. *Solar energy was invented by God and herself.

c. [I say to you that] solar energy was invented by God and myself.

Levinson is not alone with his argumentation. It is common practice in the literature to distinguish first/ second person pronouns from third person pronouns, based on the fact that only the former denote speech act participant roles.[103] Some linguists even exclude third person from person altogether, regarding it as the absence of person (cf. Harley and Ritter 2002b: 26; Richards 2008: 140).[104] Another alleged difference between first/ second and third person is the stronger need for reference disambiguation in third person pronouns, meaning the need to distinguish co-reference in sentences such as *John*$_i$ *admires himself*$_i$, from disjoint reference in sentences such as *John*$_i$ *admires him*$_j$. According to König and Siemund (1997: 102), this distinction is a highly likely motivation for a distinctive reflexive paradigm.

As logical as all of these arguments sound, the data in this study show that the claimed differences between first/ second and third person pronouns are not necessarily reflected in the degree of variation which the different pronouns allow for in actual conversations (see Table 8.6).[105] Rather, the overall results support the following statement by Newmeyer:

> The amount of formal ambiguity that one finds in language is enormous and 'usefulness' is such a vague concept that it seems inherently undesirable to base an explanation on it. In any event, it is worth asking how much ambiguity is reduced by a 3rd person reflexive anyway. (Newmeyer 2004: 535)

8.5.3 Subject complement function

More frequent than independent self-forms in subject function are independent self-forms in subject complement function (Table 8.6). In this function, too, almost all occurrences were found in coordinated NPs,[106] almost all have *myself* (12 *myself*, 1 *himself*), and most of them are non-initial. In quantitative terms, the difference in frequency between simplex

103 Compare Forchheimer (1953: 5–6): "Whoever does not act a rôle in the conversation either as speaker or as addressed remains in the great pool of the impersonal, referred to as 'third person.'." Also compare König and Siemund (1997: 105): "As a result of the special role speaker and hearer have in a verbal interaction they are more likely to be chosen as centre than a non-speaker or non-hearer." The typology of reflexives (cf. Faltz 1985; Geniušiene 1987; Schachter 1985) shows that languages with first and second person reflexives will also have third person reflexives, which is in keeping with Huang's person hierarchy for anaphora.

104 A different view is held by Boas (1911) who does not distinguish between speech act participants and non-participants but between speaker (first person) and non-speakers (second and third person).

105 Note that the development from variant to invariant reflexive markers can currently be observed in Brazilian Portuguese (*se* replacing *me*, *te*, *nos* and *vos*; cf. Azevedo 1989: 865).

106 A coordinated structure is even recognisable in the only simplex example in (197), joining the self-form and the preceding NP *my brother Jack*.

and coordinated cases is statistically significant (χ^2: 3.65 σ, p = .0003). Similar to self-form subjects, examples were found in all four dialect areas.

(197) My brother Jack, he was the eldest. And there was **miself**, so,... (subj. compl. *myself*, NblAR, NBL_007)

(198) Used to be **miself and** my other brother and mi (trunc) s- (/trunc) oldest sister. (subj. compl. *myself*, initial, WS, KEN_006)

(199) ...we have the cowman and two two more men and then there 's John and Michael **and miself and** a girl. (coord. subj. compl. *myself*, middle, A109, NTT_015)

(200) Let me see, there was Mr Scott isn't it, suppose there was about eight men in the accounts and then there was two other girls **and miself**. (coord. subj. compl. *myself*, final, TrbrKM, WIL_019)

(201) {<u IntPL> ...did people help your father, I mean did he have workers, with him?} Well, there was my older brother at that early time, mi older brother, and one man they used to employ, **and himself**, and ehr, the family used to do the rest of it, you see, long family. (coord. subj. compl. *himself*, final, SRLM_B, SOM_003)

8.5.4 No-verb utterances(2)

Most frequently, independent self-forms appear as subjects in no-verb statements and questions (one object, see (206)). In the corpus, no-verb utterances alone account for over 20% of all independent self-forms. Although the numbers are not very high, absence of V appears to play a significant role. Similar to the previous functions, most occurrences have *myself* (12 *myself*, 1 *yourself*; *me* being the default 1SG variant), and their use is supraregional (SW, MID, N).

Among the corpus examples are several picture pronouns, a phenomenon which can not be discussed separately (see fn. 87). One example, where the speaker was asked to comment on different photographs during the interview, was shown in (173).

(202) {When you started, what age did you leave school then?} **Myself**, fourteen. (subject *myself* in no-verb answer, WesAQ, WES_003)

(203) {Did you have brothers and sisters?} Yes, one brother **and then myself and** three sisters, younger than myself. (subj. compl. *myself* in no-verb listing, WesBE, WES_005)

(204) {And they heard that you bred pigeons?} Yes, oh, George Coleman **and myself**, George used to do the sending off. (coord. subj. *myself* in no-verb answer, TCA_FP, DEV_005)

(205) {Would you yourself hunt them back or did you get somebody?} No, **yourself**, after you buy them it 's you it 's up to. (subj. *yourself* in no-verb answer, SRLM_HW, SOM_030)

(206) We had three teachers, Miss Routledge the head, Miss Thomas the second, and Miss Thackeray the third one who taught the small ones. **Myself** to start with. (no-verb obj. *myself*, WesDT, WES_016)

8.5.5 Object function

In the literature, the use of self-forms in object function is less frequently mentioned, and therefore less frequently condemned, than in subject function. In coordination, especially in final position, self-form objects are usually even considered acceptable (cf. OED; Evans and Evans 1957).

Independent self-form objects are rare in the corpus. We find 2 cases with *myself* and 1 case each with *yourself*, *herself*, *ourselves* and *themselves*. Although the number of examples is small, it suffices to confirm the status of self-form objects as possible linguistic variants, especially because the different occurrences were produced by different speakers. A null hypothesis defined as self-forms never being used as independent objects, can be rejected at a significance of 2.45 σ (p = .014). Examples occur in all dialect areas.

(207) Well, alas believe **meself** the wrongs we ’ve done. (object *myself*, NotA8, NTT_013a.wav, trba.)

(208) Anyway, he had an accident in France and was killed and then that left Gunstone and the engineer, **myself** plus o’ course the top people. (coord. obj. *myself*, TrbrEGC, WIL_015)

(209) And there ’s a story to this. Might interest **yourself**. (obj. *yourself*, LanEG, LAN_012)

(210) If there ’d ha’ been any sea at all he ’d ha’ drownded **herself**. (obj. *herself*, referent ‘ship’, HED, SFK_005)

(211) ...she did keep **ourselves**, and she did do all that in addition to her ordinary, you know... (obj. *ourselves*, SRLM_WB, SOM_010)

(212) ...when we got amalgamated [...] I think what they were more concerned perhaps was selling the buildings and the land, then I could see that Mr Ponting was trying to wheedle, you know put **themselves** in and so that ’s one thing that happened... (obj. *themselves*, TrbrEGC, WIL_015; ambiguous: *themselves* refers either to the firm of which Mr Ponting is the manager or the firm that bought it in the end)

8.5.6 Prepositional complement function

Independent self-forms are most frequent in absolute terms in prepositional complement function. Compared to the large overall number of pronominal prepositional complements, however, they represent a relatively rare phenomenon (0.4%, see Table 8.6). The Modern English examples from the corpus have historical precursors in earlier stages of the language, as seen in examples (8) and (9) in section 3.1.

The analysis only includes occurrences which do not comply with the broad definition of reflexive given in 1.2 (pronouns coreferential with the preceding subject and bearing a

beneficiary or recipient thematic relation).[107] In the pronoun spreadsheets, non-reflexive prepositional complements were coded as POS 40 (simplex), POS 41 (coordinated, initial position) and POS 42 (coordinated, final position). Overall, 24 occurrences were found in the data. Despite this relatively small number, a null hypothesis defined as self-forms never being used in prepositional complement function can be rejected at a significance of 2.4 σ (p = .02). Similar to the other functions described above, we find various *myself* (13), but also *himself* (1), *herself* (4), *yourselves* (2), *ourselves* (2) and *themselves* (2). In prepositional complement function, too, independent self-forms appear throughout the country.

Perhaps surprisingly – given the general rejection of non-coordinated independent self-forms in the literature – all examples in the corpus are non-coordinated. Once more, emphasis and contrast present possible explanations. Especially the fact that self-forms appear to be favoured in comparative or contrastive utterances, indicates that they can be used to emphasise a specific verb argument as opposed to other arguments in the same sentence. This is, of course, most obvious after prepositions like *as*, *besides, but, like* and *than* (compare ch. 9.2). One of the few non-contrastive examples in the corpus is shown in (214).

(213) Which was very nice of him ʻcourse he ’s **like miself** now, he must be getting on in years. (prep. compl. *myself* after *like*, TCA_FP, DEV_005)

(214) ...and I side-stepped the manager because he was a bit apprehensive **about miself** and asked if I could, you know, get the job. (prep. compl. *myself*, TrbrEGC, WIL_015)

(215) Yes, one brother and then myself and three sisters, younger **than myself**. (prep. compl. *myself* after *than*, WesBE, WES_005)

(216) ...and she used to have to go in there at six o’clock and nobody else in **but herself**. (prep. compl. *herself*, NblJB, NBL_006)

(217) Yes it was kept **amongst ourselves**... (prep. compl. *ourselves*, IBSH, SAL_029)

(218) She make a bit of butter sometimes, she ’d save enough make a pound or a couple, you know, just **for ourselves**. (prep. compl. *ourselves*, SRLM_WB, SOM_010)

(219) They ’d tell the Lord everything that was wrong, not only **with theirselves**, not only with the church,... (prep. compl. *themselves*, IBME, SAL_023)

107 Also excluding prepositional reflexives and reflexive emphatics, e.g. *If I can work hard for other people, I might as well work for myself* (CON_009); *[he had] a pound to look after himself* (NBL_007); *I had to find out for miself what was going on*’ (KEN_005).

8.6 Summary and revised definition

Pronoun exchange is a multi-faceted phenomenon consisting of different sub-features with differing distributional tendencies. Overall, it appears that variation in case does not impinge on the functional interpretability of either personal pronouns or self-forms, as long as they maintain an identifiable position within the sentence structure. Nevertheless, the distribution of formal variants is influenced by a variety of different factors. First and foremost, the degree of variation observed between S-forms and O-forms, on the one hand, and S-/ O-forms and self-forms, on the other hand, varies in different syntactic functions (see the Context hierarchy of non-standard object forms in (183)). A highly influential factor is verb absence: the highest non-standard proportions of all syntactic functions analysed in this study were found in no-verb utterances. The influence of verb absence becomes clear in the absence of coordination.

Coordination can be attributed special importance in pronoun variation. In coordinated NPs, the non-standard frequencies are demonstrably higher than in simplex NPs. This applies to O-form subjects, S-form objects and prepositional complements, and, in particular, self-form subjects and subject complements. The pronoun's position within the coordinated structure can be decisive, too, as indicated by the preference for independent self-forms in non-initial position, and the contrast between *me and* and *and I*. Two pronouns with particularly high non-standard frequencies in coordinated NPs are *me* and *myself*.

In addition to syntactic factors, pronoun variation can be triggered by pragmatic and discourse-structuring factors. Common explanations for independent self-forms, in particular, are emphasis and contrast. Considering the many different influential factors which make PE a complex phenomenon, it is difficult to sustain any general statements, both in terms of explanations and in terms of distributional tendencies.

The different sub-features all appear throughout the country, i.e. in the four dialect areas specified in the England component of FRED (Figures 8.1, 8.2, 8.4, 8.5). This makes PE a supraregional phenomenon of spoken English, in contrast to descriptions found in earlier accounts. A classification of PE as a SW phenomenon is only justified for non-standard S-forms, but even here examples can be found in other parts of the country. The particular areal pattern of non-standard S-forms, as opposed to the more level distribution of non-standard O-forms and self-forms, explains why the former sub-feature is generally perceived as more dialectal and, therefore, given more attention in the literature. In the present study, however, the detailed analysis of different pronominal variants and syntactic

functions showed that characteristic properties of one sub-feature of PE do not necessarily apply to the phenomenon as a whole.

After a careful revaluation of the different PE sub-features, I propose a revised definition of PE which extends the traditional binary system visualised in Figure 8.6 to the ternary system shown in Figure 8.7. This revised definition includes self-forms as a third type of variant which interacts with both S- and O-forms. It extends the functional range of PE into further syntactic contexts, including no-verb utterances, for–to and ECM constructions, and reflexive occurrences. Furthermore, the revised definition acknowledges the supraregional status of PE as a general feature of spoken English.

(220) Revised definition of pronoun exchange:

Pronoun exchange is a multi-faceted phenomenon which includes different non-standard uses of subject, object and self-forms, in subject, object and reflexive functions. Variation can be observed between S-forms and O-forms, as well as between S-/ O-forms and self-forms. The exchange is bi-directional between all variants except between S-forms and self-forms, due to the absence of S-form reflexives. Pronoun variation can be triggered by syntactic factors such as coordination and verb absence, as well as pragmatic and discourse-structuring factors such as viewpoint, topicality, emphasis and contrast. While individual sub-features may be regionally restricted, PE as a whole represents a supraregional phenomenon of spoken English.

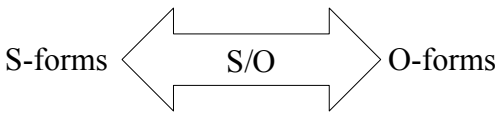


Figure 8.6: Binary PE system, traditional definition

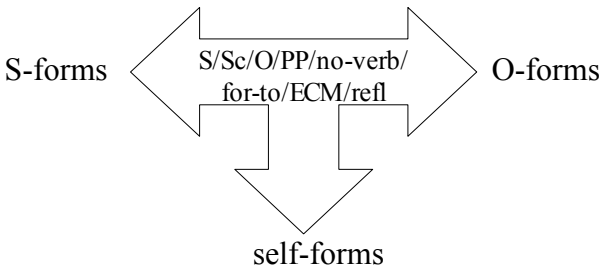


Figure 8.7: Ternary PE system based on FRED

Chapter 9
Case variation in prepositional phrases

In the previous chapter on pronoun exchange, prepositional phrases (PPs) were described as one of various syntactic environments where variation can be observed between S-forms, O-forms and self-forms. This chapter looks at pronominal PPs in more detail, giving special attention to factors which are known to favour non-standard case. After a brief general description summarising the findings for S-forms and independent self-forms presented in 8.2 and 8.5, the focus will be on comparative PPs with *as*, *like* and *than*, as well as so-called snake sentences.

A secret, kept from all the rest,
between yourself and me.
(Lewis Carroll *Alice's Adventures in Wonderland*, ch. 12)

9.1 Subject and self-forms in pronominal PPs

As might be expected, the vast majority of complements in pronominal PPs are O-forms, especially in non-coordinated occurrences. The two other morphological variants, S- and self-forms, show some interesting differences. Historically, to begin with, the use of S-forms in prepositional complement function emerged during the Early Modern English period, representing a more recent development than the use of independent self-forms in the same function since Old English (compare 3.1).

In the FRED corpus, the use of S-forms after prepositions is most noticeable in the Southwestern counties and affects all pronouns to a similar extent (Table 8.1, Figure 8.2), whereas the use of self-forms shows no regional clustering and is largely restricted to *myself* (the self-form which is also most likely to occur in other non-standard functions; Table 8.6). Case variation in pronominal PPs is a low-frequency phenomenon. Nevertheless, this particular syntactic context appears to favour variation: in non-coordinated pronouns, S-form proportions are slightly higher after prepositions than in object function, and self-forms appear more frequently after prepositions than in object or subject function. Here are some examples repeated from chapter 8:

(121, rep.) ...but he never interfered **with I**, but anybody else who came down here, he 'd go for. (prep. compl. *I*, SRLM_AE, SOM_020)

(122, rep.) There was ten years **between my sister and I** see, she died last year, she was eighty-nine. (coord. prep. compl. *I*, TrbrAS, WIL_023)

(214, rep.) ...and I side-stepped the manager because he was a bit apprehensive **about miself** and asked if I could, you know, get the job. (prep. compl. *myself*, TrbrEGC, WIL_015)

(219, rep) They 'd tell the Lord everything that was wrong, not only **with theirselves**, not only with the church,... (prep. compl. *themselves*, IBME, SAL_023)

As was mentioned in 8.2, the use of S-forms in prepositional complement function is often associated with emphasis, for example in sentences like *Give it to he, not they* (Trudgill 1999[2]: 97), or it is described as restricted to coordination, as in *between you and I*.[108] Even if counter-examples do not abound in the data, evidence from FRED contradicts any categorical descriptions. Regarding the role of specific prepositions, the corpus examples also fail to support the common perception that "the most common manifestation of this error [i.e. S-forms in prepositional complement function] contains the preposition *between*" (Riley and Parker 1998: 41). While this specific point still needs to be investigated in more detail, the present analysis did not return more instances with *between* (2x *between X and I*) than, for example, *for* (2x for *X and they*). The possibility for speakers to vacillate between different pronominal variants, even within the same sentence or paragraph, is illustrated by the consecutive use of *between you* and *between yourselves* in this example:

(221) And when you 'd done that perhaps you 'd have three in the night to clean **between you**, and you 'd settle the work up **between yourselves** using cleaning oil (prep. compl. *you* and *yourselves* in same sentence; SAL_011)

9.2 'He's as tall as me' – *as*, *like*, *than*: prepositions or conjunctions?

Grammarians generally agree that prepositional complements should be encoded by object forms (*come with me*/ *give it to him*). Opinions vary, however, with regard to comparative constructions involving expressions like *as*, *like* and *than*, since these expressions are sometimes classified as prepositions and sometimes as conjunctions. In the first classification, the pronoun following the preposition is a complement within the PP. It hence requires object case as shown in (222a). In the alternative classification, the pronoun

108 Compare Riley and Parker (1998: 41), who state that "the prohibited nominative case occurs after a preposition only in coordinate constructions." Contrary to Riley's and Parker's categorical statement, other surveys have shown that subject case is occasionally used in non-coordinated sentences such as *come with me/ I*, especially by speakers from the SW (cf. CLAE map M15).

following the conjunction is the subject of a VP with verb ellipsis, requiring subject case as shown in (222b).[109] For the purpose of this study, comparative expressions like *as*, *like* and *than* are treated as prepositions.

(222) a. You are taller [PP than [NP him]].
b. You are taller than [VP he (is)].

The dispute dates back as far as the 18th century, when comparative PPs were already one of the contexts discussed with regard to pronoun case (cf. Leonard 1929: 263–264). According to Jespersen (1949: 227–228), "the existence of such two-sided words as *but*, etc, is one of the primary causes of mistakes of *me* for *I* or vice versa, and careless uses of the cases generally." Jespersen (ibid., p. 231) also mentions that "the conjunctions *as* and *than*, used in comparisons, give rise to similar phenomena [...] the feeling for the correct use of the cases is here easily obscured, and *he* is used where the rules of grammar would lead us to expect *him* and inversely." Part of the problem is the speakers' uncertainty with regard to pronoun choice, caused by contradictory rules and recommendations. Even today, the natural use of *me* in sentences like (222a)[110] is still branded incorrect in prescriptive grammars and educational institutions. At the same time, the use of *I* without a verb to follow continues to be promoted, leading speakers who feel uneasy about this rule to "add a superfluous verb more frequently than people of other nations in such sentences as 'He is older than I am.' " (Jespersen 1949: 264).

For many decades, speakers striving to use 'good English' have been confronted with conflicting views. In his guide to *Good English*, for example, Vallins (1952: 73) stresses that *than* is a conjunction, not a preposition. He attributes the popularity of sentence-final *me* after comparative expressions to the fact that *me*, as he sees it, carries stress better than *I*. Since stress in English usually falls on the last syllable of a sentence, Vallins argues, speakers tend to put sentence-final pronouns in the objective case. In a different *Guide to Correct English*, Stratton (1949: 160) insists that "*[b]etween* is a preposition; pronouns after prepositions must be in the objective case." Stratton therefore concludes that "no matter who says it 'between you and *I*' is always wrong. So are 'between him and *I*,' 'between them and *I*,' 'between them and *we*.' Only correct are 'between you and *me*,' 'between him and *me*,' 'between them and *me*,' 'between them and *us*.' " (ibid.).

109 The fact that "there is no independent evidence that true conjunctions can be case assigners" (Schütze 2001: 213) has so far been ignored in this classification.

110 Compare Swan (1992: 78): "When *as* and *than* are followed by personal pronouns, both subject and object forms are possible. In informal English, object forms are much more common. Subject forms are more often used in a formal style (for instance, in careful writing), and some people consider them more 'correct'."

Even if the present study is not interested in the 'correctness' or 'well-formedness' of linguistic utterances, it is necessary to note just how much confusion such controversial statements can cause, and that extra-linguistic factors can influence the speech production to a considerable extent (cf. Wales 1996: 194; Harris 1981: 18). As an additional consequence, the same uncertainty may be responsible for the appearance of self-forms as an avoidance strategy, i.e. in order to avoid having to choose between subject and object case. According to Jespersen (1949: 172), "[s]ometimes the self-pronoun stands by itself; this is especially found in groups with *and*, and after *as*, *like* and *than*."

A different explanation for the appearance of self-forms after comparative expressions like *as* and *than* is emphasis, or rather contrast. Emphatic pronouns can be used to direct the hearer's attention towards one specific referent in the discourse, or they can be used to stress a particular difference between two or more referents. Zribi-Hertz (1989: 699), for instance, found in her studies that emphatic uses may trigger an alternation between pronouns and anaphors (self-forms) in violation of binding principle A (compare fn. 4 above). This can happen in multiple-foci structures or closed sets where a certain NP expresses a dominant role as opposed to other NPs. Zribi-Hertz distinguishes between three types of structures: conjunctive (i.e. coordinated), disjunctive, and comparative, as shown in the following three examples:

(223) a. John believes that letter was sent to both him and Mary/ Mary and himself. (conjunctive)

b. John believes that letter was sent to no one but him/ himself. (disjunctive)

c. John thinks that Mary is taller than him/ himself. (comparative)

In the FRED corpus, more than half of all independent self-forms in pronominal PPs form part of comparisons. They either appear after obviously comparative expressions like *as*, *besides*, *but*, *like* or *than*, or in comparative sentences after prepositions like *about*, *for*, *of* or *with*. Before we look at the quantitative distribution of the different case forms in the corpus, consider the following examples containing different comparative expressions (an example with *just* was shown in (218)):

(224) Well mi brother and I, Fred, he was older **than I**, we bought the farmhouse and the buildings... (prep. compl. *I* after *than*, NotA99, NTT_011a.wav)

(225) I mean mi sisters knew, they were about four years older **than me**, both of them... (prep. compl. *me* after *than*, YksHS, YKS_006)

(226) Oh they was **like me** they couldn't leave school quick enough could they? (prep. compl. *me* after *like*, A109, NTT_015)

(227) Oh dear, she were very particular in all **like miself**. (prep. compl. *myself* after *like*, LanW1P, LAN_004)

(228) Well I was younger (reg sic='n) **than** (/reg) **he**, but that don't mean naught. (prep. compl. *he* after *than*, CAVA_WJB, CON_005b.wav, trba.)

(229) ...because I was earnin' more **than 'im**, you see. (prep. compl. *him* after *than*, TrbrDC, WIL_018)

(230) I might be in the same boat **as him**. (prep. compl. *him* after *as*, CA, LND_001)

(231) {<u IntSF> So your father would have had his team of two helpers, two under-gardeners} <u WesCU> The two gardeners, much, nearly the same age **as himself**, I would think,... (prep. compl. *himself* after *as*, WES_012)

(232) ...somehow or the other my father found out, what, she was going out with the chap in the house. And to crown the matters, he was eight years older **than her**. (prep. compl. *her* after *than*, FM, LND_003d.wav, trba.)

(233) ...and she used to have to go in there at six o'clock and nobody else in **but herself**. (prep. compl. *herself* after *but*, NblJB, NBL_006)

(234) ...my mother would take me with her, now she had, she had as I said nine of family **besides herself**... (prep. compl. *herself* after *besides*, YksML, YKS_007)

(235) ...small places like you didn't have a lot cattle to sell. No not **like we** here one, last year, year before we sold forty up Camborne in one day. (prep. compl. *we* after *like*, CAVA_WJB, CON_005c.wav, trba.)

(236) Well, one of them didn't marry a Catholic **like us**, so she got married in the registrar office... (prep. compl. *us* after *like*, LanH1P, LAN_009b.wav)

(237) Mother and Dad, they used to play draughts quite a lot. He was very keen on draughts, and Mother was quite good. She could play quite well, too. And that used to be uh, and I used to play, but, you know, I couldn't play as well **as they**. But uh that used to put through a lot of the evenings. (prep. compl. *they* after *as*, CorJLN, CON_011a.wav)

(238) Same as you in your job, you know, same as Prime Minister and that, if Mrs Thatcher gets Prime Minister, she gets same pay **as them**. (prep. compl. *them* after *as*, LanEG, LAN_012a.wav, trba.)

(239) I 've had more holidays likes, you know, **like they** on these buses going out. (prep. compl. *they* after *like*, SRLM_RF, SOM_005, trba.)

Compared to other prepositions in the corpus, the distribution of pronoun case after *as*, *like* and *than* indicates that comparative contexts favour variation in case. Table 9.1 shows a breakdown of all pronouns after *as*, *like* and *than* by pronoun case, as well as all non-comparative occurrences. The totals in the last two rows show a comparatively higher proportion of non-object case forms in the first group. In addition, a closer look at the individual pronouns contradicts those linguistic authorities which assume a fundamental behavioural difference between *me* and third-person pronouns *him* or *her*. *Merriam Webster's Online Dictionary*, for example, states in the dictionary entry for 'than

(preposition)' that "*me* is more common after the preposition [*than*] than the third-person objective pronouns." No such tendency is supported by the FRED data.[111]

Table 9.1: Case form distribution after *as*, *like* and *than* (non-coordinated occurrences)

pronoun case after...	S-forms	O-forms	self-forms	TOTAL
as	3 (7.7%)	34 (87.2%)	2 (5.1%)	39 (100%)
like	3 (9.1%)	24 (72.7%)	6 (18.2%)	33 (100%)
than	6 (4.8%)	118 (94.4%)	1 (0.8%)	125 (100%)
TOTAL {*as*, *like*, *than*}	12 (6.1%)	176 (89.3%)	9 (4.6%)	197 (100%)
TOTAL, non-comparative	46 (1.1%)	4318 (98.7%)	10 (0.2%)	4374 (100%)

From a wider perspective, comparative PPs match the default case environments described by Schütze (2001). These are syntactic environments which are characterised by a general predominance of O-forms (in Table 9.1: 89.3%), accompanied by a relatively high proportion of other case forms, i.e. S-forms (here 6.1%) and independent self-forms (4.6%). Similar to other elliptical constructions, it could be argued that pronoun case in sentences with understood predicates is not determined by the syntax, and that speakers therefore choose default case as seen in most examples in the corpus. Schütze himself, however, describes sentences with understood predicates as "one further construction that superficially looks like an elliptical default environment" (2001: 212, fn. 8), treating comparative expressions like *as* or *than* as regular prepositions.

In the generative account of Emonds (1986: 97), subjects with understood predicates are not immediate constituents of a sentence containing an inflected verbal element, i.e. they are not governed by Infl, thus permitting the use of object case. The sentence structure is exemplified in Figure 9.1.

111 For 3SG pronouns after *like*, see the 'Constraint on NP Like X-self' in Kuno (1987: 123–124): "The NP like x-self pattern in a subordinate clause is acceptable if it is in a logophoric complement clause [i.e., pertaining to the speaker or hearer] and if the antecedent of the reflexive is [+ logo–1] [i.e., first person speaker or experiencer] with respect to the logophoric verb that takes the complement clause. It is acceptable, awkward, or marginal, subject to idiolectal variation, if the antecedent is [+ logo–2] [i.e., second person, addressee]. Otherwise it is unacceptable." According to this constraint, a sentence sequence like *Mary didn't hear from John that physicists like himself were a godsend. Jane heard it from him*. is acceptable, *himself* being coindexed with a speaker antecedent, whereas the first sentence in this sequence is not acceptable if it is followed by *She heard it from Bill*. Note that Kuno's constraint could not be tested in this study due to the lack of comparable subordinate clauses.

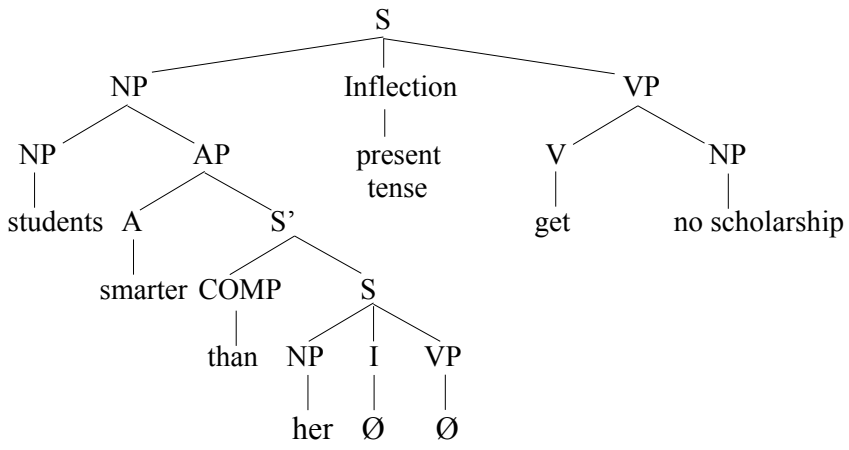


Figure 9.1: Tree structure for subjects with understood predicates (Emonds 1986: 99)

Based on the empirical results from FRED, *as*, *like* and *than* favour the use of non-object case more than other prepositions. However, the difference is effectively not very strong, given the speakers' natural predilection for O-forms in both environments. It, therefore, makes sense to argue that *as*, *like* and *than* are generally treated like prepositions, and that additional variation in case is induced by emphasis and contrast, but also uncertainty and hypercorrection. These are the same factors which explain non-standard occurrences in various other elliptical environments, including disjunctive pronouns (see 12.5), no-verb utterances (8.3.4, 8.5.4), and coordination (8.3.2, 12.5.3).

9.3 Snake sentences

A different environment which appears to be conducive to variation in case are locative PPs. Here, variation mainly concerns the use of O-forms and self-forms after prepositions indicating a location relative to the subject, such as *near* or *behind*. An extensive list of such prepositions is found in Jespersen (1949: 165–167), who includes *above*, *beneath*, *over*, *about*, *around*, *before*, *in front of*, *behind*, *after*, *within*, *upon* and *with*, directional *to*, *towards* and *for*, as well as reciprocal *between* and *among*. In the following, constructions with these prepositions will be referred to as *snake sentences*, a term which owes its popularity to examples like *Mary saw a snake near her* or *Bill found a snake near him* (e.g., Koktova 1999: 252; Huang 2000: 23; Haspelmath 2008: 55).

In the literature, snake sentences are usually discussed in connection with anaphor binding. Compare the examples in (240), where the sentence subject is coreferential with

the pronoun inside the locative PP. The acceptability of prepositional complements which are not marked for coreference in similar examples tallies with the typological tendency that direct objects are more likely to be specially marked than adjuncts or indirect objects (cf. Van Gelderen 2000b). English, in this respect, behaves similar to other languages with optional coreferentiality-marking, such as Standard Dutch (see (242)), but different from languages where the use of a reflexive form is required. One such language is Standard German, as seen in (241). According to Faltz (1985: 100), languages like German have a Strict Clause Condition (SC) stating that "a reflexivization rule must apply whenever the distance between the [coreferential] NPs is smaller than a clause (i.e. whenever the NPs are in the same clause)." Hence, the German coreferential pronoun in (241) is an overt reflexive (*sich*).

(240) a. Mary$_i$ saw a snake near her$_i$.

b. Mary$_i$ saw a snake near herself$_i$.

c. Mary$_i$ saw a snake [which was] near her$_i$.

(241) Sie$_i$ sah eine Schlange neben sich$_i$/ *ihr$_i$.
She saw a snake besides REFL/ DAT

(242) Zij zag een slang naast zich/ haar.
She saw a snake next to REFL/ OBJ

According to Faltz, snake sentences in English

> may have arisen through a confusion between the interpretation of *near him* as a locative in the main clause and the interpretation of it as a reduced relative clause on the direct object. In the second interpretation (which is still possible) the coreferent nonreflexive pronoun really *is* in a different clause, hence it is to be expected because of the Clause Mate condition. (Faltz 1985: 102)

This Clause Mate Condition (CM) states that a reflexivisation rule can only apply if the two coreferential NPs are in the same clause, which in many cases coincides with the Governing Category. Reflexivisation is, therefore, prevented if we assume that the locative PP is a reduced relative clause of the type shown in (240c) (similar, in a way, to the elliptical construction in (222b)). Unfortunately, the CM condition makes sense in some examples but is difficult to implement in others, as can be seen in (243) and (244) (**They placed the guns which were in front of them.*/ **John hid the book which was behind him*).

(243) a. They$_i$ placed the guns in front of them$_i$.

b. They$_i$ placed the guns in front of themselves$_i$.

(244) a. John$_i$ hid the book behind him$_i$.

b. John$_i$ hid the book behind himself$_i$.

Based on the two conditions identified by Faltz, the observed variability in English could be attributed to rivalling interpretations: the subject and anaphor are in the same clause (*she saw a snake near herself*), vs. the subject and anaphor are in different clauses (*she saw a snake [which was near her]*). Either way, the problem remains that an underlying relative-clause structure can only be argued for in some examples.

9.3.1 Discourse perspective

An entirely different analytical perspective focuses on the discourse function of morphological variants in sentences like (240), (243) and (244). William Cantrall (1973; 1974) has argued that pronoun case can be used to indicate a change in the speaker viewpoint. It could, for example, be argued that *Mary* is described as seeing the *snake* from an outside perspective in (240a), but from an inside perspective in (240b). A subtle change in meaning is also illustrated in (243a) and (243b) (cf. Cantrall 1974). Here, it could be argued that the act of placing the gun in the first sentence is described from the viewpoint of somebody who does not belong to the group referred to by *they*, as opposed to somebody who forms part of *they* in the second sentence. In Cantrall's words, the antecedents in (240b) and (243b) are "being asserted to be involved in the recognition of the co-reference" (1973: 46–47).[112]

Subsequent studies have proposed approaches similar to Cantrall's key-concept of viewpoint. Kuno (1987), for example, bases his explanations on the concept of empathy, which he defines as "the speaker's identification, which may vary in degree, with a person/ thing that participates in the event or state that he describes in a sentence" (p. 206). For Zribi-Hertz (1989), it is the subject of consciousness, not the syntactic subject, which has the strongest influence on pronoun choice. This semantic property is assigned "to a referent whose thoughts or feelings, optionally expressed in speech, are conveyed by a portion of the discourse" (p. 711).[113] In sentence (244), for example, the difference in meaning is

112 In this argumentation, of course, the antecedent has to be animate and physically present in the event.

113 Subjectivity has also been recognised to play a crucial role in the semantics of substitutional relations, for example regarding relative clause restrictiveness and verbal aspect (cf. Bache 1985).

attributed to the fact that *John* is the subject of consciousness in (244b), but not in (244a). While sentence (244b) can be understood as the book being very close to John, (244a) simply states that the book is behind John. Thus, "it is the structural properties of pronouns that are, in a sense, derived from their discourse properties" (ibid., p. 705).

Discourse grammar has changed the perception and understanding of snake sentences by breaking free from syntactic justifications solely based on the antecedent–anaphor relation. Instead, the focus has shifted to the extra-linguistic referent and the speaker–referent relation. Based on discussions of self-forms in snake PPs, inter alia, the binding domain of anaphora has been extended indefinitely "by assuming that NPs can be bound by NPs in any Ss provided that a certain 'command' relationship is met and that a set of semantico-syntactic conditions is fulfilled" (Kuno 1987: 74).[114]

9.3.2 Self-directed vs. other-directed verbs

An additional aspect in the discussion of snake sentences is the directedness of the verb, i.e. whether a certain verb is interpreted by default as self-directed or other-directed. This distinction probably goes back to Jespersen (1949).

König and Siemund (1997: 103) have shown that the distinction between self-directedness and other-directedness can be used to explain the morphological case marking in typical snake sentences. Considering the meaning of *put* in sentences (245) and (246), for example, it can be argued that *to put something somewhere* usually involves an other-directed change of location, and that self-directedness, therefore, had to be marked in (245). In (246), on the other hand, the action described by the verb is naturally self-directed. No re-directing is required and the anaphor can stay unmarked.

(245) Mary put the book behind herself.

(246) Mary put all problems behind her.

Similarly, Kiparsky (2008: 42) assumes that morphological variants in sentences like (247) and (248) mark a structural difference. In his view, the pronoun needs to be marked as a

114 Conditions such as the so-called 'Empathy Constraint on Reflexives' (Kuno 1987: 158): self-forms which are not the direct object of the verb need to be interpreted as "produced from the camera angle of the referent of the reflexive if the sentence is in the tense and aspect that requires an explicit camera angle with respect to the event described in the sentence." For example: *??Someone$_i$ talked to Mary about himself$_i$*; *?Someone$_i$ was talking to Mary about himself$_i$*; *Anyone$_i$ can talk to Mary about themselves$_i$*.

reflexive if the PP is a locative argument as in the first sentence, and "if and only if a referential expression can be substituted for it" (e.g., *John aimed the gun at Bill*). The pronoun stays unmarked if the PP is part of the predicate, as seen in the second sentence (compare **John brought the gun with Bill*).

(247) John$_i$ aimed the gun at *him$_i$/ himself$_i$.

(248) John$_i$ brought the gun with him$_i$/ *himself$_i$.

9.3.3 Corpus results

The corpus results corroborate the outlined approaches only in part, and only where the coreferential pronouns are unmarked. In accordance with Cantrall's concept of viewpoint, for example, third-person referents who are not involved in the recognition of the co-reference are always encoded by object forms. This can be seen in sentences like (249)–(251), where the respective referents and their actions are described from an outside perspective.

(249) ...oh Christ aye, I seen them many a time sleeping under that bloody bridge, with papers all **over them**, and sacking and one thing and another, like that. (O-form in locative PP, NotA8, NTT_013g.wav)

(250) ...and halfway through the match it started to rain and mi father always got a tarpaulin sheet and they were sitting in the back with a tarpaulin sheet **over them**. (O-form in locative PP, IBCE, SAL_026)

(251) And Fred he 'd they 'd always got a mechanic **round them**... (O-form in locative PP, A15, NTT_014)

Other examples like (252)–(257) correspond to the descriptions by Kiparsky and by König and Siemund, showing naturally self-directed verbs and instances where the PP is part of the predicate with unmarked forms. Also note the colloquial use of overt anaphora with verbs like *look behind* and *leave behind*.

(252) And from there onwards he never **looked behind him**, because he sold it to an American. (WesDX, WES_017)

(253) <u WesZs> And the girls ran out when this bang came. <u WesZb> And they carried their knitting **with them**. And they **carried** their knitting **with them**, and **left** the wool **behind them**, and it was all a network of wool. (WES_019)

(254) You, and if you did go down, you always used to **take** the candle down **with ya**, see? (JG, LND_004e.wav)

(255) I remember her coming from Little Ness and **bringing** her gardeners **with her**. There was two or three gardeners she had, and a chauffeur. She **brought** a forester **with her**; his name actually was Forester... (IBWB, SAL_004)

(256) Well, we **had** a cow dog **with us**, called Sharpy... (SRLM_WB, SOM_011)

(257) Some might **take** ʻOld Dr Watson's Tonic Stout' **with them**, that we used to make. (WflsFP, WIL_009)

Not in accordance with the outlined approaches are those utterances where we would perhaps expect to find a self-form anaphor, but where coreferentiality and self-directedness are not overtly marked. Consider the role of viewpoint and empathy first. Events or actions in which the speaker himself or herself has a central role, for example as the agent or benefactive of the verb, naturally describe an inside perspective and the highest possible degree of speaker–referent identification. In examples (258)–(263), however, none of the speakers marked the coreferential pronoun as reflexive, despite the inside perspective and high degree of identification in each utterance. This is also the case with third-person examples like (264).

(258) You had to, you used to do your own safety work, do your own timber. [...] you used to throw so much coal off the [coal] face, take that bit of coal off, and **you used to tim-timber up in front of you**, and you used to go a bit further in and take a bit more coal out, then timber again**...** (DurML, DUR_001d.wav)

(259) **I had crowds around me**, I couldn't half belt them. (DurEL, DUR_003)

(260) Mind you I was a bit on the safe side, **I put a rope round me** just, to tension up... (YksWF, YKS_001)

(261) Well, **you** just picked them up and **put them in front of you**, and went round. (CorJLN, CON_011)

(262) When we came out, **you couldn't see your hand in front of ʻee** ʻcause ʻt was so dark. (SRLM_HG, SOM_032)

(263) I always had to go ʻround from school to get a jug of it, and **we shared it between us**, like. (CA, LND_001)

(264) And the eel comes up and bites at the eels, at the worms, and then **they 've got their big tank beside them** that they just flicks it out and the eel drops off into the bath. (SRLM_CK, SOM_004a.wav)

(265) ...and all at once these bullocks started, went mad or somewhat, and **they got me between them** and squashed mi stomach. (SRLM_HG, SOM_032)

Reflexive marking is also missing from some locative PPs which require it according to the above-mentioned arguments by Kiparsky. In (260), for example, reflexive *me* could be substituted for a different referential expression, as in *I put a rope around him/ John*. The verb *put* is typically other-directed (compare (245)). The only difference between the FRED example in (260) and Kiparsky's example in (247) lies in the person: first-person

me in (260) can only refer to the speaker, whereas the absence of a reflexivity marker in (247) could potentially cause ambiguity. In Kiparsky's sentence, the reflexivity marker therefore also functions as a disambiguation device.

Probably the most interesting point in this discussion is the observable gap between, on the one hand, examples given in the literature to illustrate the need for reflexive marking, and, on the other hand, the rare need for disambiguation in actual conversations. In the FRED interviews reflexive marking in potentially ambiguous sentences is much less common than one would intuitively expect. While the distinction between self- and other-directed verbs presents a logical solution for otherwise ambiguous cases (*she*$_i$ *put the book behind her*$_{i/j}$), it appears that its use in spontaneous conversation is limited, especially if coreferentiality can be inferred from the wider context. Instead, speakers appear to follow a very simple, least effort principle which I propose to call the 'Avoid Ambiguity' Principle. This principle implies that no special case marking is required in sentences where no disambiguation is required, including sentences where relations of coreferentiality can be inferred from the context.[115] Although, at present, the existence of such a principle still rests on conjecture, a tentative formulation is presented in (266). Special tests will be needed to ascertain that such a principle really exists and in order to measure its influence on linguistic choices.

(266) 'Avoid Ambiguity' Principle
Use unambiguous case if disambiguation is required;
if no disambiguation is required mark case as usual.

9.4 Summary

In pronominal PPs, variation in case can be triggered by a variety of factors. Emphasis, for instance, suggests itself as an explanation for the use of non-object forms after comparative expressions like *as*, *like* and *than*, but also in other PPs where one referent is delimited against other entities in the conversation. At the same time, external factors should not be underestimated. This concerns, in particular, the speakers' uncertainty regarding some syntactic constructions. The ongoing debate on 'correct' case assignment after comparative

115 Compare the description of German reflexivisation strategies by Faltz (1985: 118): "In the case of languages like German, the reflexive is not used when it does not contribute information that cannot be carried by ordinary pronouns. [...] We might describe a setup like this as being 'functionally streamlined'; use a reflexive pronoun only when an ordinary pronoun will not do."

as, *like* and *than*, for example, is reflected in many contradictory rules and recommendations in official grammars and educational institutions. A certain degree of hypercorrection can, therefore, be expected in sentences containing these expressions.

The debate also concerns so-called snake sentences. It has been argued that the morphological distinction between personal pronouns and self-forms in such sentences is used to mark underlying structural differences or differences in viewpoint. Based on the corpus results, however, it appears that speakers rarely make use of this option in spontaneous speech. The findings fully confirm Faltz's statement that "English speakers agree that the nonreflexive pronouns are perfectly acceptable even when coreference with the subject is intended" (1985: 101). In the FRED interviews, O-forms are found in all contexts where, according to theories proposed in the literature, one would at least expect a strong predominance of self-form reflexives. The absence of overt reflexivity marking does not appear to cause any misunderstandings, even in supposedly ambiguous cases.

It is, however, important to note that most examples used in the literature to illustrate links between pronoun case and viewpoint (and related concepts), or between pronoun case and underlying structural differences, are taken from formal discourse, often narratives. Such examples differ fundamentally from the data at hand, the latter being characterised by spontaneity and a face-to-face setting. It can, therefore, be argued that the tendencies observed in this study do not so much contradict the existing theories but rather reflect the morphosyntactic variability and diversity of different registers and media. For the interpretation of examples from the corpus this means that the observed variation is not, or not entirely, predictable by the available theories. In the contrastive and locative PPs in our data, viewpoint, empathy, and even coreferentiality, are rarely deducible from pronoun case. Ambiguity is usually absorbed by contextual information and logical connections between the sentence components (e.g., cause and effect). Overall, the results show a general preference for O-forms, irrespective of the semantic and structural differences that could potentially be marked by pronoun case.

Chapter 10
Gerunds and participles

This chapter is concerned with pronominal modifier variation in participial constructions.[116] It has been noted in the literature that the use of possessive modifiers is in decline, even in 'pronoun + V-*ing*' constructions which clearly function as nominals (gerunds as 'nouns in disguise') and even in the more formal registers of English.

In the following discussion, the focus will first be on the distinction between participles and gerunds from a structural and cognitive perspective (10.1). Based on the distribution of different case forms in the corpus (10.2), we will then pinpoint the underlying semantics of 'pronoun + V-*ing*' constructions (10.3). Special attention will be given to the use of O-form modifiers: in alternation with possessive modifiers (*excuse me/ my asking your age*) and in alternation with S-forms (*it's no good me/ I taking you*; *I can never remember us/ we having a pig*). We will also take a look at the use of 'O-form + V-*ing*' instead of *to*-infinitives (*what do you want me doing*), bare infinitives (*we used to watch them putting it into the big boats*), and for–to constructions (*I used to wait for them coming out*). The use of O-form modifiers instead of the more formal zero option in adverbial participles will be discussed in 10.5 (*me/ Ø being the youngster of the lot, I had a bicycle*). The function coding applied in the analyses corresponds to Table B3.

10.1 Structural and cognitive distinctions

It is usually agreed that gerunds like *swimming*, or clauses containing a gerund (*swimming indoors at this time of year*), function as nouns within the larger sentence whereas participles function as verbs (*he keeps swimming fast*). As a logical consequence, the pronominal modifier preceding a gerund in Standard English (*my swimming*) has possessive function and should, therefore, have possessive form (POSS) whereas the pronominal modifier preceding a participle (*me swimming*) acts as a personal pronoun and should, therefore, have personal pronoun form.

116 I am grateful to Liesbet Heyvaert for a very interesting discussion at the ISLE 2008 conference, and the paper given by Heyvaert and De Smet on the same occasion, which sparked my interest in gerund–participle variation.

Historically, gerundive nominals can be traced back to an Old English derivative in *-ing* which functioned as a full noun. According to Heyvaert et al. (2005: 72), action nominals "have preserved most of the latter's nominal properties, still taking nominal determiners other than the genitive/ possessive". This can be seen in the following examples from the COBUILD corpus:

(267) ...when it comes to [the awarding of contracts]... (Heyvaert et al. 2005: 72)

(268) During July we noticed [a browning of our cherry tree leaves]... (ibid.)

The authors also mention several other participial nominals which are no longer considered grammatical today, or only marginally grammatical. Examples for these constructions can be found in FRED, including participials preceded by demonstrative *this* or *that*, as seen in (269), and participials preceded by quantifiers *no* or *any*, as seen in (270).[117]

(269) a. There was all **this cutting of them** with a knife, and you were making your own marmalade. (gerundive nominal preceded by *this*, WesZb, WES_019)

b. But **this taking blood** was a general thing wasn't it in years gone by. (gerundive nominal preceded by *this*, WesZb, WES_019)

(270) a. There was **no buying coal**... (gerundive nominal preceded by *no*, HM, KEN_003)

b. ...you 'd got do what your boss told you to do, there was **no eh favouring any of them**. (gerundive nominal preceded by *no*, NotA85, NTT_010)

c. ...maybe in a village there would be a dress maker, but you had to do a lot on your own because there was **no going and buying anything ready made** then. (gerundive nominal preceded by *no*, NblJB, NBL_006)

d. I've never heard tell of **any feeding on beech**... (gerundive nominal preceded by *any*, WesDX, WES_017)

The fact that distinctively nominal POSS modifiers were never adopted in verbal participials can be taken as a sign that historical boundaries continue to exist. The formal boundary between gerunds and participles has always been permeable in only one direction: POSS modifiers only appear in gerunds whereas the use of non-POSS modifiers in both nominal gerunds and verbal participles has been known since Old English times (cf. Visser 1966). The prescriptive rule that pronominal modifiers of gerunds *must* be POSS is, therefore, not only outdated but never really applied.

The postulation that, in correct English, one should always use POSS in sentences such as *she was opposed to his seeing his daughter* (cf. Stratton 1949) is easily disqualified by comparison with the language in use. In addition, POSS modifiers have undergone a

117 No examples of gerunds preceded by articles, as in *he objected to the sending away the servant*.

considerable decrease in frequency over the last century (cf. Visser 1973: 2364; Heyvaert et al. 2005: 85). While the use of oblique determiners was still considered "distinctly unusual" in literary English at the beginning of the 20th century (Poutsma 1929: 835), categorical claims were later qualified in grammars such as Quirk et al. (1985). In the latter, the authors state that if the *-ing* clause has a subject,

> the item realizing the subject may be in the genitive case or otherwise in the objective case (for pronouns having a distinctive objective case) or common case (for all other noun phrases). In general, the genitive is preferred if the item is a pronoun, the noun phrase has personal reference, and the style is formal. (Quirk et al. 1985: 1063)

More recently, Heyvaert et al. (2005) have shown that oblique determiners are nowadays not only well-established in colloquial, informal style, but that they are starting to predominate in formal registers, too. Based on data from the Bank of English,[118] the authors draw the following conclusion:

> It would, in short, be descriptively inaccurate and (prescriptively) misleading to say that formal language prefers gerundive nominalizations with possessive pronouns. [...] it is the oblique case which is predominant in present-day gerundive nominalizations in general. (Heyvaert et al. 2005: 79)

This view is corroborated by the findings presented in 10.2, which show that pronominal modifiers in gerunds are almost exclusively non-POSS in the spontaneous conversations of the FRED corpus. Accordingly, non-POSS modifiers should be included in general descriptions of the English gerund, at least as a possibility. Ideally, they should also be acknowledged in cross-linguistic comparisons. In the typological overview of action nominals in the *World Atlas of Language Structures* (Haspelmath et al. 2005), for example, English is described as a 'mixed' language with possessive-accusative marking (*Peter's singing the Marseillaise*), ergative-possessive marking (*the destruction of the city by the enemy*) and double-possessive marking (*the enemy's destruction of the city*). The corresponding map by Koptjevskaja-Tamm (2005) is shown in Map D9 in the Appendix.

118 COBUILD project; study based on patterns of the form 'personal/ possessive pronoun + V-*ing*' in formal written texts (newspaper texts from *The Times*, over 5m words) and informal spoken registers (UK spoken corpus recordings, over 9m words). The authors found that oblique case predominates in both registers, although there is a higher percentage of possessives in the formal data (informal 92.3% oblique vs. 7.7% possessive; formal 65.8% oblique vs. 34.2% possessive).

Both Heyvaert et al. (2005) and the results presented in 10.2 strongly suggest the inclusion of a fourth strategy which can be described as double-accusative marking (*Me singing the Marseillaise*). The corresponding tree structure is shown in Figure 10.1.

A variety of relatively simple tests can be used to determine the grammatical status of 'pronoun + V-*ing*' constructions, irrespective of the pronominal modifier's form. They are presented in concise form in Table 10.1.[119]

Table 10.1: Transformation tests for gerund vs. participle classification
(assembled from various sources including Heyvaert et al. 2005)

transformation test	gerund = nominal function	participle = verbal function
passivisation	*He suggested leaving early.* *> Leaving early was suggested.*	*He kept leaving early.* *> *Leaving early was kept.*
replacement with finite clause	*He suggested that we leave early.*	**He kept that we leave early.*
left dislocation	*Leaving early he suggested.*	**Leaving early he kept.*
clefting	*Leaving early was what he suggested.*	**Leaving early was what he kept.*
substitution by pronoun	*He suggested it.*	**He kept it.*
passive voice	*this would lead to [him being replaced]*	---
clause-like objects	*the prospect of [their joining the EU]*	---
auxiliaries of secondary tense	*no sign of [him having been home]*	---

In cognitive terms, the essential difference between verb-like and noun-like participials consists in their profiling behaviour: verbs profile processes whereas nouns profile things or objects. According to Langacker (1991: 551), verbs describe "a relation comprising a series of component states distributed through a continuous span of conceived time and scanned sequentially." Nouns, on the other hand, describe "an abstract region in some domain, i.e. a set of interconnected entities" which are profiled collectively (ibid., p. 19). By adding an *-ing* suffix, the sequential part is removed from both verbal and nominal constructions. In participles, *-ing* turns the process described by V into an atemporal, adjective-like structure which either profiles a snapshot of stative relations from beginning to end, zooms in on a middle part, or shifts the profile to the final stages. In gerunds, *-ing*

119 On the phonological level, Labov (1989: 87) has observed that the *-in'* variant is favoured in progressives and participles but less commonly found in nominal gerunds.

profiles the abstract region (more details in De Smet and Heyvaert 2009). Atemporalisation therefore forms the common semantic ground of both constructions. Imperfectivisation, on the other hand, does not form part of the core semantics of participial constructions. This can, for example, be seen in sentences like *she regretted leaving Belfast*, i.e. she had actually left Belfast, or *I regret eating that chocolate cake*, i.e. I actually ate that chocolate cake (examples from De Smet and Heyvaert 2008).

Differences in the profiling behaviour of participles vs. gerunds entail further semantic distinctions which are occasionally ignored. First of all, both constructions can be used to describe a specific event (*they liked our singing/ they liked us singing > they liked that we sang*), but only the gerund construction can also imply the manner or 'how' of the event (*they liked our singing > they liked how we sang*). On the other hand, the participle construction can be used to express the conditionality or 'if' of the event. This has been illustrated by Poutsma (1929: 837), who has noted "an appreciable semantic difference between the two constructions, as appears from a comparison of [...] *What do you think of my sister's singing?* with *What do you think of my sister singing?*" Only the second sentence "is suggestive of a conditional clause: *What would you think if my sister sang?*"

10.2 Distribution of case forms in the corpus

Moving on to our data, a delimitation of participles against gerunds based on structural and cognitive distinctions is not applied that easily. This is mainly due to the fact that the distinction between verb-like and noun-like behaviour, between process and abstract region, frequently becomes blurred in spontaneous speech. Concerned are not only pronouns like *her* – where a morphological distinction is precluded a priori – but all pronominal modifiers before V-*ing*, given the near-exclusivity of object case forms.

(271) Course her family was very much against **her marrying** him, wasn't, weren't they? ('*her* + V-*ing*', TCA_RA, DEV_001)

(272) Every time I went down Little Dawley, the eldest girl of Lees, Hilda, she used to shout after me, till I got tired of **her shouting** and I thought, I 'll stop **you shouting** after me. ('*her* + V-*ing*' and '*you* + V-*ing*', IBFB, SAL_013)

(273) There 's a possibility they might have done without **me knowing**... ('*me* + V-*ing*', DurNB, DUR_002)

Table 10.2 shows the distribution of modifier case in all 'pronoun + V-*ing*' constructions forming part of a gerund-participle supercategory. The numbers also include those cases where speakers could have chosen a zero realisation (*I remember Ø doing it*).

Table 10.2: Case form distribution in gerund-participles: POSS vs. non-POSS[120]

gerund-participles	1SG	2SG/ PL	3SGm	3SGn	1PL	3PL	TOTAL (%)
S-forms	1	20	2	35	1	0	252 (98.8%)
O-forms	38		45		14	96	
POSS	1	1	0	0	0	1	3 (1.2%)
TOTAL	40	21	47	35	15	97	255 (100%)

The empirical results show a very clear predominance of non-POSS forms in 99% of all occurrences, the default option being object case. As will be elaborated below, O-form modifiers appear to the same extent in verb-like and noun-like occurrences, covering the whole semantic spectrum discussed in 10.3. POSS-form modifiers, on the other hand, are practically never used in the interviews. The FRED results corroborate the conclusion by Heyvaert et al. (2005: 79) that "the oblique case is well established in informal language, to the extent that it has practically replaced the possessive."

10.3 Underlying differences in meaning

Despite the acknowledged predominance of O-form modifiers in present-day English, some linguistic studies still adhere to a surface distinction between participles and gerunds which is supposed to reflect a deeper-level distinction in the meaning of V-*ing*. Abney (1987), for example, distinguishes between ACC-*ing* constructions with object case modifiers, and POSS-*ing* constructions with possessive modifiers, as illustrated in Figures 10.1 and 10.2. In both constructions, -*ing* has the same basic property: it functions as a nominaliser which takes a verbal projection and converts it into a nominal category. The

120 Not included in the analysis are (a) *her* and <mi>; (b) adnominal adjuncts as in *washing tub* or *courting days*; (c) nominal -*ing* forms such as *weaving*; (d) occurrences where V-*ing* is really a progressive verb but where the auxiliary is reduced to such an extent that it becomes inaudible, as in *I talking to 'ee* (SOM_028), *Dick Bean goes over whilst them looking the wrong way* (DEV_005), or *He walking up in Hornsey Road* (LND_001); (e) occurrences where the pronominal modifier is the object of a causative verb such as *have*, *keep* or *make*, or other matrix verbs, as in *they had me working on the main road* (NTT_002), or *he sent me flying backwards*, (SOM_032); (f) sentences where the pronoun and verb are separated by a longer pause, as in *I can see him now -- coming up the Paddock with his pipe* (SAL_031); (g) irregular syntactic constructions, as in *he was quite good to me, letting me carrying on during the summer* (WIL_016); (h) reflexive or emphatic pronouns, as in *I found meself working at the China Works* (SAL_030), or *mi miself being mi father's (gap)* (SOM_028); (i) picture NPs, as in *I've got a picture of miself building that warship* (MDX_001). Note the occasional elision of V-*ing*, for example in *I can't mind her married* (CON_005, 'I can't remember her being married/ getting married/ marrying').

difference between the two phrase structures lies in the scope, or sisterhood, of -*ing*, i.e. in the point on the projection path of V at which the conversion takes place. According to Abney, there is a verbal agreement feature (AGR) in ACC-*ing* assigning common case (object case), and a nominal agreement feature in POSS-*ing* assigning genitive case (POSS). In Figure 10.1, ACC-*ing* affixes to IP. The subject (*me*) behaves like the subject of a sentence whereas the POSS-*ing* subject in Figure 10.2 (*my*) behaves like the subject of a noun phrase.[121]

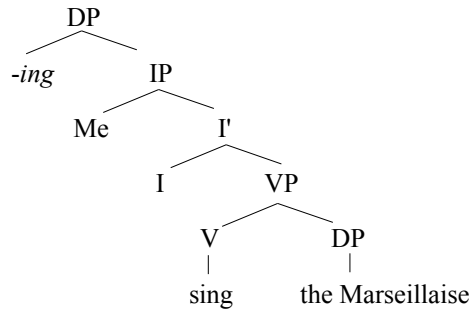


Figure 10.1: Tree structure 'Me singing the Marseillaise' (Abney 1987: 144)

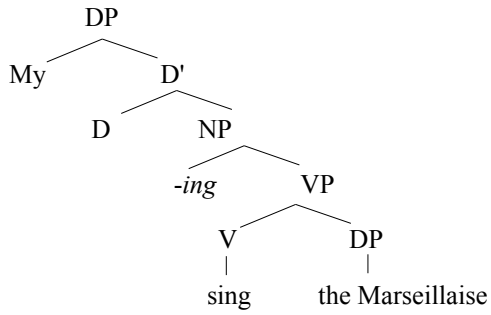


Figure 10.2: Tree structure 'My singing the Marseillaise' (Abney 1987: 144)

Similar analyses are extremely difficult, if not impossible, to apply to the data at hand, due to the simple fact that underlying differences in -*ing* (cognitive/ semantic) are not overtly marked. Considering the distribution of case forms in Table 10.2, it is obvious that, should there be a difference in the underlying structure, case assignment in Abney's terms is not

121 According to Reuland (1983), case is assigned to -*ing* and then transmitted to the subject of -*ing*.

observed in actual conversation. Even in occurrences which could be argued to have an underlying POSS-*ing* structure, speakers resort to non-POSS variants. Consider the following examples:

(274) ...and my sister had left school two years **prior to me leaving school**... (nominal -*ing* with O-form modifier, WesCB, WES_010)

(275) ...a number of full tubs start from Kemberton or from the screens, and they 'd go down one set of rails, pulled along by an endless rope, it used to be an endless chain **prior to them introducing** a rope, as went round a return wheel at the bottom... (nominal -*ing* with O-form modifier, IBGJ, SAL_032)

(276) ...then the one who finished milking first had to corner them in the corner, you see, and with a dog as well, you know, **before you milking** of it. (nominal -*ing* with O-form modifier, SRLM_WB, SOM_011)

Given that the surface distinction between participles and gerunds is close to dissolving, the question arises whether the underlying differences are still recognisable. Only if the answer to this question is yes, and only if we ascertain that the observed absence of modifier distinction is not caused by underlying semantic or cognitive changes, can we safely presume that O-forms have genuinely replaced POSS in V-*ing* constructions.

In the following chapters, it will be demonstrated that semantic differences in 'pronoun + V-*ing*' constructions are recognisable despite the missing modifier distinction. These differences mainly lie in the perspective taken on the action, event or state described by V-*ing*. Some occurrences emphasise the internal, progressive aspect of V-*ing*, accentuating the agent-status of the involved subject, whereas others reflect an external, stative perspective on V-*ing*. In order to facilitate a categorisation of individual examples, two sets of binary oppositions will proposed. First, a distinction will be made between an acting-oriented vs. action-oriented perspective on V-*ing*, similar to the cognitive profiling described above. This distinction is closely linked to an additional distinction between potential or hypothetical actions, states and events, vs. actual actions, states and events (my own labels). If these binary oppositions were to be invariably marked by pronoun case, the first of each pair would be encoded by non-POSS modifiers, the second by POSS modifiers. Looking at the data, however, we find no such distribution.

10.3.1 Acting vs. action

The most salient semantic difference between gerunds and participles in the traditional sense probably lies in the perspective taken on the action or event described by the verb. In the following, examples will be given for an *action-oriented perspective*, in contrast to an *acting-oriented perspective*.[122] Rather than describing a simple opposition, however, these two perspectives should be seen as the extremes of a continuum along which the different corpus examples are situated. The acting-oriented end of this continuum describes an internal, progressive view, in contrast to the action-oriented end, which describes an external, stative view on the result of the acting process (not to be confused with completed action). The corpus data provide strong evidence for asymmetrical case marking: POSS modifiers always reflect an action-oriented perspective whereas O-form modifiers can reflect both perspectives.

The only three occurrences of 'POSS + V-*ing*' in the corpus all express an action-oriented view. Two of them describe an event already completed at the time of the utterance, hence supporting the opinion expressed above that imperfectivisation does not form part of the core semantics of participial constructions:

(277) But **prior to my eh going** onto the Electric Light Company, things happened in my life. (V-*ing* with POSS determiner *my*, MDX_001, MdxCG)

(278) See, Vaughan, I can **mind their all going away**, I can mind thirteen names on that list there. (V-*ing* with POSS determiner *their*, CAVA_DB, CON_001, *mind* 'remember')

(279) ...he says, No I ain't shopped you, he says, But probably some of the others (gap 'indistinct') they don't like the idea **of your going**... (V-*ing* with POSS determiner *your*, A15, NTT_014a.wav)

Most interesting are certainly those examples which contain comparable 'POSS + V-*ing*' and 'O-form + V-*ing*' with nearly identical meanings. In (277), for example, the participial not only fills the same syntactic position as in (275) (*prior to X V-ing*), but both constructions express the same perspective on V-*ing*, describing an actual event completed in the past. Similarly, the 'POSS + V-*ing*' in prepositional complement function in (279) is comparable to the 'O-form + V-*ing*' examples in (280) and (281):

(280) But talking **about me living** there -- I was there some time with mi stepfather's mother... (V-*ing* in prepositional complement position, IBFB, SAL_013)

122 The term action is used in contrast to acting. It is not to be confused with the 'action -*ing*' used by Lees (1960: 65) to distinguish between action nominals (*his drawing of the picture*) and factive gerundives (*his drawing the picture*).

(281) They were jealous **of us being** up there, somehow or other. (V-*ing* in prepositional complement position, LanD1B, LAN_001)

The use of POSS in (278) is directly comparable to other examples with *remember*, as seen in (282) and (283), as well as in the negated examples (284)–(289):

(282) Oh I ’ll I ’ll have to tell you about this, can you **remember me telling** you about the whole hams over the whats-a-names shops, hanging up on hooks. (V-*ing* with *remember*, NotA8, NTT_013)

(283) I do **remember him being trapped** down the mine. (V-*ing* with *remember*, LanF1P, LAN_007)

(284) I can’t **remember us ever doing it**. (V-*ing* with *remember*, WesCE, WES_011)

(285) ʻCause I know my father never had more than two pounds, and I never **remember them being hard up**. (V-*ing* with *remember*, WesCU, WES_012)

(286) A lot of them did, but I can’t never **remember we having** a pig. (V-*ing* with *remember*, WflsFP, WIL_009)

(287) They were four-pound loaves, they were made by Mr Williams from Woodhouse Lane, and I can **remember him coming home**, and **pulling** his bayonet out of his scabbard, **sticking** it in the loaf and **holding** it up in the air. (V-*ing* with *remember*, IBCD, SAL_21)

(288) ...an’ I can **remember vividly ʻe lookin’** at mum an’ ʻe said Come on then Lovey, an’ ʻe kissed ʻer an’ ʻe said, I ’ll send for you an’ the children as soon as ever I can. (V-*ing* with *remember*, TrbrDP, WIL_004a.wav)

(289) I **remember them being made** but I ’ve, I cannot remember who was making them. (V-*ing* with *remember*, IBNB, SAL_006)

In the literature, *remember* is known as a matrix verb which takes both morphological variants (cf. Huddleston and Pullum 2002: 1232 ; Heyvaert et al. 2005: 85).[123] Examples with *remember* are often situated somewhere between the action and acting ends of the continuum. With *remember*, and other verbs, too, O-forms appear in descriptions of completed events or states of affairs, but equally in descriptions of actions in progress. In the latter, they not only profile the progressive aspect of V-*ing* but an internal perspective, by stressing the subject-role of the pronominal modifier. It is therefore not surprising that non-POSS is generally preferred with verbs of perception, like *see*, *watch* or *hear*. Heyvaert et al. (2005: 76), for example, exclude these verbs from their analyses “because rather than offering the choice between GEN and oblique case, they necessarily take oblique determiners.”

(290) He was **watching me hoeing** and **working**, so he must have seen something in me, you know. (V-*ing* with verb of perception, ICS_EW, CON_009)

123 Heyvaert et al. (2005) find that *forget*, to which Huddleston and Pullum (2002) ascribe the same properties as *remember*, is not used with POSS.

(291) ...he ’d rather a Scotch accent but when I **heard him grumbling** about the bill, I knew it was him. (V-*ing* with verb of perception, WesAD, WES_001)

(292) Used to **hear he coming** up through the village. (V-*ing* with verb of perception, WflsGW, WIL_011)

In numerous instances, it is difficult or even impossible to decide the status of V-*ing* as verbal or nominal participial, or to place the participial on the action–acting continuum. In fact, there are so many of these instances that it is plausible to speak of a gerund-participle supercategory (cf. Huddleston and Pullum 2002; 2005). The following examples from FRED support De Smet and Heyvaert (2008), who speak of ‘collapsing categories’ and a supercategory of *ing*-clauses with “further new clusters of *ing*-clauses developing into additional subcategories that are neither strictly gerundial nor participial.”

(293) Well, you ’ll have to get that out yourself. I know she won’t get it, specially **you leaving her**. (gerund-participle, FM, LND_003)

(294) I was so frightened and **it was him bellowin’ down**, it frightened me such a rate I would never use the phone. (gerund-participle, TrbrDC, WIL_018)

(295) Twice I **remember it ripping the railway up**. (gerund-participle, IBEP, SAL_039)

(296) One week we ’d have to have Good Friday off, ‘cause Mr (sname) Cremer (/sname), he didn’t **believe in you working holiday time**...(gerund-participle, HM, KEN_003)

(297) It all **depended on them giving a hand** sometime during the year. (gerund-participle, SRLM_SC, SOM_035)

(298) And then there ’s a good many of ‘em who were unable to buy clothing, so it necessitated **them wearing** their working clothes on a Sunday... (gerund-participle, IBCH, SAL_034)

(299) But these pictures that May had, the big one, framed, showing **them going over Wrynose**... (gerund-participle, WesCY, WES_013)

10.3.2 Hypothetical vs. actual

In addition to the acting vs. action distinction discussed above, a distinction can be made between hypothetical and actual actions or events. Based on the above-mentioned properties – in particular the profiling behaviour of noun-like participials – the logical conclusion seems to be that the use of POSS determiners is appropriate if V-*ing* refers to a completed event or an actual state of affairs. POSS determiners appear less suitable if the action or event is only hypothetical. In 10.1, an early example from Poutsma (1929) was used to illustrate the possibility to express conditionality by ‘non-POSS + V-*ing*’ (*What do you think of my sister singing? > What would you think if my sister sang*), but not by ‘POSS + V-*ing*’ (*What do you think of my sister’s singing?*). Compare the following two

examples:

(300) a. My mother and father wouldn't let me have a bicycle because they were so afraid of **me having an accident**. (V-*ing* with hypothetical meaning, IBPB, SAL_012)

b. My mother and father wouldn't let me have a bicycle because they were so worried about **my having an accident**. (V-*ing* with actual meaning, fictitious example)

Similar to the acting vs. action distinction, the case marking in the hypothetical vs. actual distinction is asymmetrical. Here too, POSS modifiers are the marked option, whereas O-forms appear with both underlying meanings. In other words, POSS modifiers always appear with V-*ing* describing an actual action or event, whereas O-forms are used with both actual and hypothetical V-*ing*. Let us take a closer look at the data.

As was mentioned above, only three POSS examples were found in the corpus. In all three cases, shown in (277)–(279), the participial describes an actual event or action completed before the time of the utterance (in (279), the *going* took place before the reported conversation). Among the numerous 'O-form + V-*ing*' examples, on the other hand, we find examples which describe a hypothetical event, as seen in (301) or (302), and other examples which refer to events or actions which actually took place, as seen in (275) or (280). The different occurrences clearly show that distinctive case marking is not required for the hearer to be able to distinguish between hypothetical and actual meanings. The correct semantic interpretation of 'O-form + V-*ing*' is inferred from the overall context and the extra-linguistic information available to the discourse participants (cf. De Smet and Heyvaert 2009, for the importance of different contextual effects). In example (303), for instance, the action described by V-*ing* could be interpreted as both hypothetical or actual: 'as if I was giving him the sack' or 'like when I gave him the sack'. Similarly, it is clear from the context that the money referred to by *it* in (305) had not actually been taken when the utterance was made.

(301) ...so the copper come 'round, he says, It 's alright, he said, No good **me taking you**, he said, You might be away tomorrow. (V-*ing* hypothetical meaning, CA, LND_001)

(302) And he said, what chance is there of **me starting** on t' commission with a horse of mi own? (V-*ing* with hypothetical meaning, WesDX, Wes_017)

(303) I said, Mr Simons, I said you don't really mean this do you, I mean it 's like **me givin'** the boss the sack? (V-*ing* with hypothetical/ actual meaning, TrbrEGC, WIL_015)

(304) Well, no good **me having** big (v 'laughter') ideas. (V-*ing* with hypothetical/ actual meaning, FM, LND_003)

(305) Besides, he said, In any case, what 's wrong with you **taking it**? (V-*ing* with hypothetical/ actual meaning, TCA_WH, DEV_009)

10.3.3 'O-form + V-*ing*' for a more vivid mode of expression

Occasionally, O-form modifiers are not only found where prescriptive grammar expects the use of POSS, but also where a different grammatical structure could be expected, for example instead of bare infinitives or *to*-infinitives. Among the different verbs involved in such examples, we find *want* in otherwise typical ECM constructions, causative verbs like *make* or *get*, verbs of perception like *see* and *watch*, and for–to constructions as seen in (312)–(314):

(306) I said, what d' you **want me doing**? (V-*ing* instead of infinitive after *want*, WesBR,WES_008)

(307) Particularly if there was anything like a suicide or anything involved, they wouldn't **want it coming** out in the open anyway. (V-*ing* instead of infinitive after *want*, interviewer utterance, Int, DEV_001e.wav)

(308) Yes they all did, we used to **get her singing** and her voice used to wobble... (V-*ing* instead of infinitive after *get*, NotA71, NTT_008a.wav)

(309) Now sometimes she 'd do apples and apples made it better for eh, (unclear) **help** (/unclear) (trunc) an- (/trunc) you know, **it [the juice/ jelly] setting** better. (V-*ing* instead of infinitive after *help*, Lang1p, LAN_003)

(310) Oh yes, I remember it all being closed and it was really a pity to **see it going**. (V-*ing* instead of infinitive after *see*, IBWB, SAL_004)

(311) We used t' **watch 'em puttin' it** [the coal] into the big boats. (V-*ing* instead of infinitive after *watch*, GCS, SFK_010)

(312) ...and then after they went it was all to wash up, towels to wash, tea towels to wash, rub all the tables down again, mop all the floor up again, ready **for them coming** in at dinner time... (V-*ing* instead of for-to construction, DurEL, DUR_003)

(313) I said, he used to go and sit outside the synagogue steps and wait **for them coming** out. (V-*ing* instead of for–to construction, YksWF, YKS_001)

(314) ...he used to have to come from Epperstone in a pony and trap. We used to be waiting ages **on him coming** (v 'laughs'). (V-*ing* instead of for–to construction, NotA80, NTT_009)

Rather than affecting the usual semantic distinction between V-*ing* and infinitive constructions (e.g. recurrent vs. one-time event), it appears that the use of 'O-form + V-*ing*' indicates a subtle shift towards a more personal inside perspective, in analogy with its function as a verbal participial. In sentences like (315a) and (315b), for instance, the grammatical change causes the focus to shift from the verb of the main sentence to the process described by V-*ing*. In addition, the use of V-*ing* conveys the idea of an action in progress, which is more easily turned into a mental picture than the neutral infinitive:

(315) a. he waited for them to come out

b. he waited for them coming out

The shift in perspective is only subtle, but it very much agrees with the nature of oral story-telling in general, and the animated, personal style of the FRED interviews in particular. The effect is comparable to that of the pleonastic pronouns discussed in chapter 12. Through insertion of an overt subject before a non-finite participle, sentences like (316)–(321) are made to sound more personal. This can, but does not have to, be underlined by extra stress.

(316) So I went up with the rest of 'em and, instead o' **me gettin'** my ninety pound, I got seventeen. (V-*ing* with pleonastic subject, EF1, SFK_003)

(317) ...I was on night work -- (unclear) there (/unclear) was a phone message come through, and instead of **me going**, coming home Sunday morning, at nine (trunc) o'c- (/trunc) eh, eh seven o'clock, like with the other lads, I had to eh... (V-*ing* with pleonastic subject, HM, KEN_003)

(318) You 'll hear enough o' that blowin' through the riggin' when there 's a gale, without **you whistlin'** the wind up! (V-*ing* with pleonastic subject, WT, SFK_012)

(319) They said, What about **you having a go**? I said, I 'll have a go at it. (V-*ing* with pleonastic subject, NotA103, NTT_001c.wav, trba.)

(320) And he said, what chance is there of **me starting** on t' commission with a horse of mi own? (V-*ing* with pleonastic subject, WesDX, WES_017)

(321) By **me joining** the eh Oddfellows Club, as a boy, of course, I was entitled to attend Dr Hatch, o' course, free. (V-*ing* with pleonastic subject, MdxCG, MDX_001c.wav)

10.4 Syntactic constraints

The corpus results leave no doubt about the general preference for O-form modifiers in participial constructions with all types of meanings. Nevertheless, it should be noted that the appearance of POSS is occasionally prevented by syntactic constraints. First and foremost, by interposed adverbials and other interjections, as seen in (322)–(324) (cf. Poutsma 1929; Jespersen 1949).

(322) I observed **him** <u>rapidly</u> **drawing** a picture. ('pronoun + V-*ing*' with interposed adverbial, example from Thompson 1973: 374)

(323) Yes, I remember **them** <u>like</u> **dying** and all that sort of thing. ('pronoun + V-*ing*' with interposed filler *like*, LA, LND_005)

(324) Yes, I remember **them**, <u>Ken</u>, **coming** into Wellington... ('*pronoun* + V-*ing*' with interposed hearer address, IBBeBo, SAL_011)

Furthermore, the use of POSS is prevented if the pronominal modifier itself is qualified, for example by a noun as seen in (325). In similar examples, genitive case can only be

marked on the qualified NP as a whole. The use of POSS would change the internal structure of the NP, as illustrated in (325c). Almost a century ago, Poutsma already found that the use of 'POSS + V-*ing*' was "hardly practicable when the preceding modifier of the gerund consists of more members than one, or is followed by some adjunct; thus it could not very well replace the [personal pronoun] in: [...] 'Talk of us girls being vain, what are we to do?' " (Poutsma 1929: 836, quoting Thackeray *The History of Henry Esmond*; also cf. Huddleston and Pullum 2002: 1192–1193). On the other hand, occurrences found with qualifier *all* stand in opposition to Poutsma's claim that "we never find *us* (*you* or *them*) *all* (or *both*) instead of *our* (*your* or *their*) *all* (or *both*) before a gerund" (1929: 836). Two such examples are shown in (326) and (327):

(325) a. Well, when these apprentice boys started, they didn't like **us** <u>trades-boys</u> **using** a file or anything like that, 'cause uh they reckoned we was, well, taking their trade away. (V-*ing* with qualified modifier, KentTAD, KEN_005)

b. ?they didn't like **us trades-boys'** using a file

c. *they didn't like **our trades-boys** using a file

(326) I remember going in there and seeing **them** <u>all</u> **painting**... (V-*ing* with qualified modifier, SRLM_CA, SOM_009)

(327) But I remember **us** <u>all</u> **carrying** these potatoes up and what not on our backs, and **putting** them in the back yard when it was fine and sunny to dry out. (V-*ing* with qualified modifier, WesDT, WES_016)

The use of POSS is clearly disfavoured in coordination. Even if examples are very rare in the data, sentences like (328) and (329) illustrate how 'strange' the use of POSS would appear in a coordinated modifier NP (*?the general's and his staff's appearing*; *?my sister's and my standing these dresses up on the table*).

(328) on **the general and his staff appearing** (Poutsma 1929: 836, OED quotation, my bold type)

(329) And ourselves, we had eh embroidered dresses on for Sundays, and these dresses were done up at the laundry and starched. And I can remember **my sister and I standing** these dresses up on the table on Saturday night, ready to get into on Sunday, and they 'd stand by themselves, they were so stiff. (V-*ing* with coordinated modifier NP, YksMW, YKS_008c.wav)

Last but not least, the use of POSS seems impossible after *with* in sentences like (330). The same tendency has been observed by Huddleston and Pullum (2002). Note that no occurrences were found with POSS after *without*, as mentioned by Heyvaert et al. (2005).

(330) I don't know whether it did any good like, **with him knowing him** or not (V-*ing* after *with*, YksWF, YKS_001)

10.5 Adverbial participles

This last section is concerned with the use of overt pronominal modifiers in cases where V-*ing* is functionally equivalent to an adverbial clause. Such adverbial participles often appear at the beginning of a sentence, replacing adverbial clauses of time, condition or reason, as in *her being the oldest* > *because she was the oldest*. It has been claimed that the subject modifier in adverbial participles requires nominative case, especially in sentence-initial position. A description of the so-called 'nominative absolute' in formal English, for example, is found in Poutsma's grammar:

> The subject-indicating word in the adverbial participle-clauses of the second group is a noun or pronoun in the nominative. As this (pro)noun in the nominative is not connected with a finite verb, it is called a *nominative absolute*, and the same name is often given to the entire clause; e.g.: *He being absent, nothing could be done* [...] In Old English the dative was used in these clauses [...] They were used in imitation of the Latin *Ablativus Absolutus*, and, being foreign to the genius of the language, were not employed by writers who stood outside Latin influence. Also in Middle English they are mostly traceable to foreign, chiefly French originals. Nor can [...] the modern nominative absolute be said to have become natural to the vernacular. [...] Its use with a pronoun as subject is limited apparently to cases in which *being* or *having* is the Participle (e.g. *They having the keys, no entrance was possible*). (Poutsma 1929: 973)

According to Poutsma, object case in adverbial participles came to be supplanted by the nominative, or subject case, in the mid-15th century. In present-day English, subject case is still occasionally postulated. Abney (1987: 144), for example, states that "nominative Case is usually only assigned in absolutive constructions, such as 'Mary was wasting her time on John, [he being a confirmed bachelor]'."

The persistent demand for subject case, however, is refuted by actual linguistic behaviour in native-speaker conversations. A quantitative analysis of the FRED data returns the results presented in Table 10.3, leaving no doubt that in adverbial participles, too, O-forms are the default option. Of the 54 pronominal subjects (not including *you* and *it*), 51 are O-forms (94.4%), and only 3 (5.6%) are S-forms. This comes very close to the results for gerund-participles shown in Table 10.2.

Table 10.3: Case form distribution in adverbial participles (subject function)
(no overt case distinction in *you* and *it*)

adverbial participles	S-forms								O-forms					
	I	*you*	*he*	*she*	*it*	*we*	*they*	TOTAL	*me*	*him*	*her*	*us*	*them*	TOTAL
	0	1	0	0	6	2	1	3 (10)	26	10	1	2	12	51

In addition, a surprisingly large number of empirical examples are preceded by *with*, as seen in (331)–(334). Although this phenomenon is not usually mentioned in English grammars, it accounts for over a third of all adverbial participles in the data of this study (23 cases). In the FRED interviews, the mostly causal or temporal constructions with '*with* + pronoun + V-*ing*' were used by 18 different speakers. Their apparent predominance in MID and N data needs to be confirmed by further studies.

(331) **With me being** at the cafe for so long and Mrs Gascoigne was taken ill, **I** got married and we had a reception... (adverbial participle preceded by *with*, WesCY, WES_013)

(332) She would always, and **with me being** ginger, **she** always made me yellow. (adverbial participle preceded by *with*, LanC5P, LAN_005)

(333) And, o' course, **wi' me bein'** below **they** used t' know exactly what t' do, 'cause as soon as I got the go astern I used t' give her the whack... (adverbial participle preceded by *with*, EJM, SFK_006)

(334) And when he come, he were reeking with drink, and he asked mi dad for some gingerbreads, biscuits. [...] So mi dad had particularly got some in **with him coming**, he knew he liked them. (adverbial participle preceded by *with*, LanD1P, LAN_006)

Regarding possible explanations for the overt use of pronominal modifiers in adverbial participles, the corpus examples can be divided into two groups, each with its own motivation. These groups are equally represented with 28 and 30 cases respectively.

On the one hand are those cases where the subject of V-*ing* is identical with the subject of the main clause, as seen in (335)–(341). In these sentences, the subject of V-*ing* functions as a pleonastic modifier which can be derived logically from the main subject. In a more formal style, the pronoun slot would probably be expected to stay empty, as in 'Ø not knowing any different, I ran home' (compare (335)).[124]

(335) So, **me not knowing** any different, a bit dumb in the brain box, **I** run up home to tell mi mother. (adverbial participle with inferable subject, LA, LND_005)

(336) If he found you, you was supposed to spit over his head, and **me being little I** couldn't spit over his head, I used to spit in his ear. [...] The idea was to get as high up, up to this

124 In the corpus, overt pronouns alternate with zero, e.g. *Being so young and silly, I didn't really know what I was doing* (SFK_003). Unfortunately, a quantification of zero modifiers is beyond the scope of this study.

lad who was upright. To let the other fellows on. **Me being light**, **I** used to jump right at the top. (adverbial participle with inferable subject, YksJV, YKS_002a.wav/ YKS_002b.wav)

(337) **Me being inquisitive I** said, Well what are you going to do with it? (adverbial participle with inferable subject, LAN002, LAN_002)

(338) **You having** casual jobs like, then, sometimes I would eh **you** would pop into a job for a few weeks and -- that was it. (adverbial participle with inferable subject, FP, OXF_001)

(339) **Him being** short-sighted, **he** didn't know it was the same boy... (adverbial participle with inferable subject, IBGJ, SAL_032)

(340) ...**he** used to shout at the men stoppin' at the Rose and Crown and takin' their money there, instead of home, **him bein'** a pious man. (adverbial participle with inferable subject, IBJP, SAL_038)

(341) And **with them being** keen fox-hunters, **they** thought a lot about foxes... (adverbial participle with *with* and inferable subject, YksSL, YKS_003)

The second group consists of so-called 'dangling participles', also known as 'hanging' or 'unattached' participles. In this type of occurrences, the subject of V-*ing* is not identical with the subject of the main clause, and the participial does not modify the main subject: *walking back home yesterday a tree nearly fell on my head*, or *turning the corner, the view was quite different*. Dangling participles are considered unacceptable in formal registers of English if the subject of V-*ing* is not readily identifiable. The *American Heritage Dictionary of the English Language* (Pickett 2000[4]), for example, advices its readers to avoid such participles in writing and to recast the respective sentence so that the subject of the attached phrase becomes clear: *as I was walking back home yesterday, a tree nearly fell on my head*, or *when we turned the corner, the view was quite different*.

The insertion of an overt subject before V-*ing*, of course, allows the speaker to maintain the sentence structure and, at the same time, resolve the latent ambiguity of dangling participles. Examples for this technique can be seen in (342)–(347), as well as (332) above. In the latter sentence, a dangling V-*ing* without an overt subject could refer to either of the two women, i.e. the speaker herself or her mother (*being ginger, she always made me yellow*). The insertion of *me* identifies the speaker as the subject of V-*ing* (*me being ginger, she always made me yellow*).

(342) And they were all to be kept, and of course **me being** the second oldest, the money had to come from somewhere. (dangling participle, YksWF, YKS_001)

(343) And **me being** a lad, mi father had learned me how to catch moles... (dangling participle, WesDX, WES_017)

(344) See, and then, Sir William Earl$_i$ started to build the uh bridge, see, and then when the, when he started, of course, **him$_j$ living there**, they$_k$ [...] got him$_j$ [the speaker's father] to do all the boatin', boating for him$_i$. (dangling participle, WS, KEN_006)

(345) ...but I can't (trunc) rem- (/trunc) recollect much of his life with me, you know, **we not being** entwined very much... (dangling participle, NotA85, NTT_010, trba.)

(346) Oh no, he didn't, no. They were (trunc) la- (/trunc) **over them knowing** all these bosses you see, **he** got him in the yard. But he had to go labourin', but he, he had a good job in the yard. (dangling participle, LanD1B, LAN_001a.wav, trba.)

(347) They [the builders] used to make greenhouses and things like that, with them, **with them [the poles] being rustic** and the (trunc) bar- (/trunc) (trunc) ba- (/trunc) bark was on them you see. (dangling participle preceded by *with*, NotA30, NTT_003a.wav)

The same means also facilitates the use of adverbial participles with split subjects:

(348) We had all the, **her bein'** the eldest, **and me bein'** the youngest, **we** had to bring, we had to help all the family kind of thing... (adverbial participle with split subject, LanEG, LAN_012)

(349) ...and eh as I say, eh, me being, **me or mi mum** or whoever it is **being** that little bit of a boss, **you** 'd got t' do what your boss told you to do... (dangling participle with split subject, NTT_010, NotA85)

Overall, the use of overt subjects before V-*ing* facilitates a reduction of processing effort. In those cases where the subject of the participle is identical with the subject of the main clause, an overt pronoun can be used for emphasis, but it also reduces the effort required to derive zero.[125] While zero may be considered by some to be the more elegant option, an overt pronoun in the same position seems useful considering the swift nature of spoken dialogue. In sentences where the derivation of the participle subject is logically possible but requires some processing effort, the insertion of an overt modifier can significantly reduce that effort. This is, for example, seen in (343) (compare *Being a lad, mi father had learned me how to catch moles*). In dangling participles, in particular, deriving a covert subject can lead to a wrong interpretation of the sentence. The use of an overt modifier allows the speaker to clarify the syntactic and semantic relations within the sentence. It also allows him or her to hold on to the chosen grammatical construction (with relatively little additional effort) without risking a misunderstanding.

10.6 Summary

It is well known that the use of possessive modifiers in English gerunds is in decline, even in the more formal registers. The analyses presented above confirm this trend, showing O-

125 According to Huang (2000: 245), "the tighter the linkage, the more likely that the coreferential arguments will be encoded by zero anaphors or pronouns." In Huang's semantic context hierarchy (lexical NP > pronoun > zero anaphor), the zero option is the semantically least specific realisation (ibid., p. 215).

forms as the supraregional default in all participial constructions of the 'pronoun + V-*ing*' type. A traditional distinction between gerunds and participles based on the case of the pronominal modifier has become extremely hard to maintain. O-forms are the natural choice, even in participials which clearly function as nominals, whereas possessive modifiers have fallen into disuse. The question whether underlying differences in meaning are still recognisable in the absence of a morphological distinction can be answered with yes: semantic differences which were formerly reflected in the case of the pronominal modifier, such as the profiling behaviour of verbal vs. nominal participials, continue to exist. The differences remain recognisable despite the fact that, in the surface realisation, O-forms have taken over the territory of POSS.

Instead of the traditional binary distinction between gerunds and participles, the underlying semantic differences are best described as a continuum reflecting the speaker's perspective on the action or event described by V-*ing*. Two distinctions which span most occurrences are the distinction between an action-oriented vs. acting-oriented perspective, and the distinction between hypothetical vs. actual actions and events. In the surface representation of these distinctions, O-forms are the unmarked variant, covering the entire semantic spectrum. POSS modifiers, on the other hand, represent a marked option which is limited to the action-oriented perspective and descriptions of actual events. In agreement with other recent investigations (e.g. Heyvaert et al. 2005; Heyvaert 2008; De Smet and Heyvaert 2009), the current results speak in favour of a gerund-participle supercategory. This is supported by the general predominance of O-forms and the presence, in the data, of numerous cases for which the exact position on the action-acting continuum is unclear.

Finally, the data show occasional occurrences of 'O-form + V-*ing*' participials used instead of *to*-infinitives, bare infinitives and for–to constructions – occurrences which noticeably contribute to a more oral mode of expression. This equally applies to overt O-forms which appear instead of zero. In adverbial participles in particular, such pleonastic modifiers fulfil an important function. Especially in cases where the derivation of the participle subject is logically possible but requires some processing effort (e.g., dangling participles), the use of an overt modifier can considerably reduce that effort, and can help clarify the syntactic and semantic relations within the corresponding sentence.

Chapter 11
Qualified pronouns

In the previous chapters, we looked at personal pronouns standing on their own in argument and modifier positions. We will now focus on pronouns which combine with so-called qualifiers in more complex NPs. In this study, such pronouns are referred to as qualified pronouns; in the literature they are also known as modified pronouns (e.g., Schütze 2001).

The analysis will include qualifiers from different grammatical categories (see 'specification' and 'notes' columns in the spreadsheets): nouns as in *us girls*, numerals as in *us two*, names as in *we Garbutts*, and combinations thereof, such as *us two girls* or *us Outton boys*. Also included are the quantifiers *all*, *both* and *each*, as in *they all*, *they both* and *they each*. Excluded are appositional cases such as *It's me, John* (cf. Bhat 2004: 44–45).[126] The analysis will be presented in three parts. Chapter 11.1 gives a general overview of post-pronominal qualification as observed in the FRED corpus, including the distribution of case forms and a comparison of synthetically vs. analytically qualified occurrences (*us two* vs. *the two of us*). In chapter 11.2, special attention will be paid to the use of post-pronominal qualification in second person plural NPs. Chapter 11.3 focuses on aspects of pre-pronominal qualification.

Areal distribution patterns will not be analysed in detail since pronoun qualification represents a supraregional phenomenon. No significant differences were detected between the four dialect areas in either synthetically or analytically qualified occurrences.

11.1 Constructions with qualifier nouns and quantifiers

Pronoun qualification occurs with both singular and plural referents. The best-known singular examples are probably second-person vocatives of the *you bastard* type (cf. Dunkling 1990), and pre-pronominal qualifiers as in *poor me* or *soft you*. In the plural, pronoun qualification is used to specify groups with more individuals than just the speaker and hearer, as in *We teachers must resolve this problem* or *You students need to wait until four* (cf. Bhat 2004: 45). Here are some first examples from the corpus:

126 Qualifier nouns have elsewhere been classified as appositions (cf. Kjellmer 1971: 44–45).

(350) And I come home and he said to me, Who are **you boy**? (qualified singular *you*, CAVA_WW, *you boy* forming a prosodic unit, compare CON_006c.wav)

(351) But he said, I ’ll get hold of you, **you bastard**! (qualified singular *you*, CA, LND_001)

(352) Go on, do it **you bastard**, she said,... (qualified singular *you*, 2SG, CA, LND_001)

Among the post-qualified pronouns in the corpus we find singular and plural *you*, 1PL *we* and *us*, and 3PL *they* and *them*. Table 11.1 shows a breakdown of absolute frequencies by qualifier type, including qualifier nouns and personal names, qualifier numerals, and the three quantifiers *all*, *both* and *each*.

Table 11.1: Qualified pronouns, breakdown by qualifier type (synthetic cases)
(additional grey figures indicate non-standard uses)

qualified pronouns	qualifier								
	noun/ name	numeral	*all*–U	*all*–Q	*both*–U	*both*–Q	*each*–U	*each*–Q	TOTAL
you-SG	25	---	---	---	---	---	---	---	25
we	63	4	1	300	0	19	0	18	405
thereof obj./ prep.compl.	9	1	0	0	0	0	0	0	10
us	72	8	24	0	2	0	0	4	110
thereof subj./ subj. compl.	29	4	0	0	0	0	0	0	33
you-PL	24	6	4	36	0	2	0	1	73
they	144	5	0	385	0	36	0	10	580
thereof obj./ prep.compl.	124	3	0	0	0	0	0	0	127
them	1297	19	180	2	10	0	0	7	1515
thereof subj./ subj. compl.	143	10	0	2	0	0	0	0	155
TOTAL	1625	42	209	723	12	57	0	40	2708

In Table 11.1, the second person numbers include generic *you* and *you all*, as in *Everybody started, well you all started* (YKS_007). *They* and *them* were classified as demonstratives when followed by a noun or numeral, as in *them two* or *they days* (POS 90; compare ch. 15). Due to the many demonstrative occurrences, 3PL pronouns have the largest share in the over 2700 examples in the corpus.

Occurrences with *all*, *both*, and *each* were divided into quantificational uses (listed as *all*–Q, *both*–Q, *each*–Q) and universal uses (*all*–U, *both*–U, *each*–U). This differentiation is, for example, made in the *Cambridge Grammar* (Huddleston and Pullum 2002: 427),

where inseparable 'universal pronouns', as in *She likes you all* (*She had liked [you all]*), are distinguished from separable 'quantificational adjuncts', as in *We all enjoyed it* (*We had all enjoyed it*). The question whether composite NPs like *we all* should be treated as lexical units will not be discussed. On the one hand, the relatively high frequencies of such composite NPs, and the contiguous pronunciation of the two component parts (y'all /j כ:l/) speak in favour of such an interpretation; the fact that they can be separated by verbs and other parts of speech speaks against it.

Based on the empirical examples, the most salient properties of pronoun qualification can be described as follows:

- The primary function of pronoun qualification consists in referent specification. Pronoun qualification provides additional information used to specify individual referents or groups of two or more referents.[127]
- Pronoun qualification can be used to distinguish number in the second person, where the distinction is no longer morphologically marked in English (compare 3.1).
- Pronoun qualification is a means for quantifying plural referents, and for indicating distinctions which are not marked in the Modern English paradigm. Among the different qualifiers, the quantifier *both* and the numeral *two* encode duality – a concept which in earlier stages of English was expressed by dual forms (see Table A9, Appendix; Map D7 shows the areal distribution of variants of 'just we two'). While there is, theoretically, no limit to the number of referents which can be encoded by numeric quantifiers, it is interesting to see that speakers seem to prefer analytic constructions for larger, multi-morphemic numbers, e.g. *us two*, *us thre*e, but *twelve of us*, *thirty-two of us*, and so forth. Synthetic equivalents with large numbers such as *us thirty-two* or *us two hundred* rank intuitively low in acceptability, indicating that the productivity of numeric qualifiers in synthetically qualified NPs is not unlimited.
- Pronoun qualification is not used to mark inclusive–exclusive distinctions. Similar to other European languages, English has no such distinctions in its pronominal paradigm, i.e. no morphological or lexical distinction between pronouns which do or do not include the addressee(s) (*we*: 'I + other(s)' vs. 'I + you (+ other(s))'). Covert clusivity can only be inferred from the situational context: if a speaker A welcomes a

127 No examples were found for comic uses as suggested by Ihalainen (1994: 111), who describes *us girls* as a historically comic rustic expression.

hearer B with the words *I'm glad we two could meet*, for example, it can be assumed that *we two* means A and B.

- A certain degree of persistence can be observed in the linguistic behaviour of the interviewees, who show personal preferences for specific pronoun–qualifier combinations. For instance, speaker IBME is a frequent user of '*we* + qualifier noun', whereas speaker WesDT is a frequent user of *we all*.

In the following, the phenomenon will be approached from two perspectives:

1. Does pronoun qualification have a positive impact on non-standard case, and does it affect subject as well as object functions?
   (The distribution of case forms will be investigated in synthetically qualified NPs as compared to non-qualified NPs, both non-coordinated.)

2. Are certain qualifiers preferred in specific syntactic contexts?
   (Synthetically qualified pronouns will be compared against equivalent analytic occurrences.)

In order to answer these questions, we will need to distinguish between synthetic and analytic occurrences. Pronouns followed by an adjacent qualifier noun, numeral or quantifier will be referred to as synthetic (*we two*); pronouns in complement position as analytic (*two of us*).

11.1.1 Qualified vs. non-qualified NPs

In order to answer the first question, i.e. whether pronoun qualification is effectively conducive to non-standard case, we first take a look at the distribution of case forms in synthetically qualified NPs such as *we boys* or *us girls*.

As in many other areas of the grammar, variation in qualified pronouns is perceived as critical by some language-related institutions. The Queen's English Society, for instance, refers to qualified pronouns as an "area of confusion" on its website, similar to pronouns in coordination. Both phenomena have been registered as imprecision and error prone by these guardians of 'good English' (http://www.queens-english-society.com).

Variationist studies, of course, take a different view. Although non-standard case in qualified NPs, for instance subject NPs like *us kids*, has not been analysed in great detail, it has not gone unnoticed in the literature (e.g., Anderwald 2004: 178; Emonds 1986: 100).[128] Qualified pronouns have, for example, been described by Schütze as one of various default case environments where pronoun case can vary due to missing or ambivalent case assignment.[129] From a structural perspective, the pronoun in utterances such as (353) can be argued to occupy the position of head in absence of a determiner, in which case it would require subject case. If we assume, however, that case in similar occurrences is not assigned by syntactic rules, the appearance of object case can be explained as default case (cf. Schütze 2001: 215–216). Along the same lines, the pre-qualified pronoun in (354) (more examples in 11.3), can not head the subject, either because the preceding material fills the corresponding position (e.g. the determiner in *the real me*), or it can only occur to the right of the determiner position, as is the case with *soft me*. It follows that the pronoun in the subject NP is not assigned subject case and therefore receives its case by default (O-form). Schütze's examples include qualifier nouns and numerals in subject position, as in *We/ Us linguists are a crazy bunch* or *We/ Us three have to be leaving now*. With other adnominal adjuncts variation seems less acceptable, for example *How much would us/ ?we with insurance have to pay?*

(353) ...but if it was a job it were needed it were, was carried out, **we**, **us bairns** used to carry it out to them... (post-qualified *us* with subject function, NblJB, NBL_006)

(354) So ʻcourse, **soft me**, I took them. (pre-qualified *me* with subject function, WesBL, WES_006a.wav)

If Schütze's claims are true, ambivalence in case assignment should not only affect qualified NPs in subject function, but also in object function. The view that non-standard O-forms are more acceptable than non-standard S-forms, however, is quite common. In his article "Us Anglos are a cut above the field", for example, Kjellmer (1986) focuses on ʻ*us* + noun' NPs in subject function in both American and British English, including written examples from newspapers such as *The Observer* (BrE) and the LOB and Brown corpora.

128 Emonds (1986: 98) proposes a generative explanation based on the so-called Adjacent Head Condition: "two heads of phrases can be related by a transformational rule only if one governs the other." In qualified NPs like *us girls*, this condition is not satisfied. The categorical conclusion that the pronoun in such an NP must therefore have object case is not supported by the distribution found in FRED.

129 While default case may present a universal concept, default O-forms do not. In German, for example, qualified NPs with subject function have nominative forms, as in *Ihr beiden kennt euch?* (ʻYou two know each other?'), and qualified NPs with object function have accusative forms, as in *Gehört das Auto euch beiden?* (ʻDoes the car belong to both of you/ you both?'). My intuition as a German native speaker tells me that other case forms would be judged unacceptable by most speakers in these sentences.

In Kjellmer's opinion, 'O-form + noun' NPs in subject function are "reasonably acceptable in colloquial English" (ibid., p. 445) whereas 'S-form + noun' NPs in object function are described as hypercorrections. As an explanation, Kjellmer mentions the general drift towards O-forms in English:

> When by the conjunction of a pronoun and a noun phrase in standard English the subject can behave in the normal noun-phrase fashion, viz. retain its objective form and stay invariable, and the reason for preserving a special distinctive subjective form of the pronoun is thus removed, the drift [towards default object forms] need not be checked any more but is free to assert itself. As a result, the type *Us Anglos are*... arises. (Kjellmer 1986: 448)

Let us take a look at the case distribution in FRED:

Table 11.2: Case form distribution in qualified vs. non-qualified pronouns (synthetic)

qualified pronouns		subject/ subject compl. function		object/ prep.compl. function	
		qualified	non-qualified	qualified	non-qualified
1PL	*we*	395	16757	10	19
	us	33	116	77	1491
	% ns*	**7.7%**	0.7%	**11.5%**	1.3%
3PL	*they*	453	21675	127	49
	them	155	63	1360	6779
	% ns*	**25.5%**	0.3%	**8.5%**	0.7%

*non-standard

The results in Table 11.2 show that non-standard case in qualified NPs is a frequent phenomenon in both subject and object function. Perhaps unexpectedly, qualified *us* in subject function is relatively less frequent than qualified *we* in object function (7.7% vs. 11.5%; the frequencies for *they* and *them* can not be taken at face value due to the large proportion of prepositional *in they/ them days*). The following examples show different occurrences with post-qualified *we*, *us*, *they* and *them* in different syntactic functions (more demonstrative *they*/ *them* in ch. 15):

(355) And mi father used to have home-brewed beer and bread and cheese and onion, and **we lads** had dripping cake and a mug of milk. (qualified *we* subject, IBJD, SAL_019)

(356) A Frenchman had this farm before **we Garbutts** come... (qualified *we* subject, YksWG, YKS_009)

(357) ...it might be raining hard tomorrow, oh, yes, that ʻs what she told **we boys**. (qualified *we* object, SRLM_RF, SOM_005)

(358) ...he didn't touch this pottery stuff, you know left it all to **we lads**. (qualified *we* in prepositional complement function, SRLM_CA, SOM_009)

(359) ...and she 'd come down in between **we two**, throwing out ropes each side... (qualified *we* in prepositional complement function, SRLM_WT, SOM_028)

(360) **Us Outton boys** Sunday dinnertimes plenty o' times [...] we used t' walk t' the Maid... (qualified *us* subject, JRS, SFK_011)

(361) That was all right amongst **us small people**... (qualified *us* in prepositional complement function, RR1, SFK_007)

(362) ...because they couldn't call **us both** Charlie, because we wouldn't know which one they 'd be calling! (*us both* object, SRLM_CA, SOM_009)

(363) **They both** came from Bath. (*they both* subject, TrbrKM, WIL_019)

(364) Then after a bit we bought the New Hotel and we ran **them both** for one year. (*them both* object, WesAD, WES_001)

(365) All my frocks, or, or owt, that was torn or too small, **them all** used to go in. (*them all* subject, LanC5P, LAN_005)

(366) That was a herrin' voyage, when ol' Oscar Pipes and **them all** had motor boats, little motor boats out there. (*them all* subject, HED, SFK_005)

As was mentioned above, non-standard case in qualified NPs occurs in subject as well as object function. However, in order to test whether qualification really has an impact on case selection, the obtained frequencies need to be compared against those in non-qualified pronouns (previously discussed under pronoun exchange in ch. 8). These additional frequencies are also included in Table 11.2, for both syntactic functions.

The numbers strongly indicate that qualification has a positive effect on non-standard case, irrespective of the NPs syntactic function. Among the different pronouns, the results for *we* and *us* are the most reliable, since the figures for *they* and *them* are heavily affected by their use as demonstrative modifiers. In subject function, we get 7.7% *us* in qualified NPs, as compared to only 0.7% non-qualified *us*. A similar result is obtained in object and prepositional complement function, where 11.5% qualified *we* stand in opposition to only 1.3% non-qualified *we*. Among the different qualifiers, qualifier nouns and numerals have the strongest impact on non-standard case, as can be seen in Table 11.1. The quantifiers *all*, *both* and *each*, on the other hand, do not appear to have any particular influence on pronoun choice: all combinations consisting of 'S-form + *all*/ *both*/ *each*' in the corpus have subject or subject complement function, and almost all combinations consisting of 'O-form + *all/ both/ each*' have object or prepositional complement function. The only two

exceptions were shown in (365) and (366). Furthermore, the tendency for non-standard case to appear with qualification is also illustrated by sentence-internal switches. Compare example (353), where simple *we* changes to qualified *us* in the same syntactic function.

11.1.2 Synthetically vs. analytically qualified NPs

The second question asked at the beginning of this chapter was whether certain qualifiers are preferred in specific syntactic contexts. For this purpose, synthetically qualified pronouns need to be compared against equivalent analytic occurrences. While synthetically qualified pronouns are widely accepted in spoken English, their analytic equivalents are often perceived as more formal: compare *us all* vs. *all of us*. How are the two structural variants distributed in the spontaneous conversations in FRED?

The numeric results are presented in Table 11.3. They include all instances of proper interchangeability, i.e. all cases where the speaker could have used either of the two structural variants (e.g., *us all* or *all of us*; coded POS 62 and POS 62a respectively). Not included are, (a) qualifier nouns (*we boys*, but **the boys of us*); (b) other qualifiers and quantifiers restricted to synthetic occurrences (e.g., *us others*, but **(the) others of us*); (c) quantifiers restricted to analytic occurrences, such as *one*/ *none*/ *half*/ *dozens*/ *several*/ *any*/ *a load*/ *lot*/ *few*/ *crowd of* (compare *one/ none of us*, but **us one/ none*; *half/ dozens of them*, but **them half/ dozens*); (d) cases with potential *of*-elision (e.g., *all they*/ *all them*); and (e) obviously demonstrative or adverbial uses, as in *two of them that was up in the draughtsmen department* (WIL_019), or *he used to live out here Broomborough, they days* (DEV_007). When combined with a numeric qualifier, *they* and *them* are mostly demonstrative, but they were included for the sake of completeness. It should also be noted that synthetic NPs with numeric qualifiers can be paraphrased by both definite and indefinite analytic constructions. The phrase *us four*, for example, can refer to a group of four people and four people only, in which case it is probably paraphrased as *the four of us*. The indefinite variant, *four of us*, more likely refers to four people who form part of a larger group. Based on corpus concordances, it is virtually impossible to infer which paraphrasis is correct; this particular distinction was therefore not considered.

In Table 11.3, the numbers in brackets show that syntheticity (s) and analyticity (a) are variably correlated with pronoun case. Both structural variants appear with S- and O-form pronominals, but S-forms are generally preferred in synthetically qualified NPs, and O-forms are generally preferred in analytically qualified NPs. In addition, the data show

frequent alternation between the two structural variants, sometimes even within the same text passage or sentence. Two examples are shown in (367) and (389).

Table 11.3: Qualified pronouns: synthetic vs. analytic variants
(the statistical values indicate how significant the difference in frequency is between the synthetic and analytic occurrences of each qualifier; the additional grey figures indicate non-standard uses)

qualified pronouns	1PL (*we* + *us*)			2PL (*you*)***			3PL (*they* + *them*)		
	s	a	σ	s	a	σ	s	a	σ
numeral	12 (4+8)	**207** (4+203)	**40.9	6	**24**	**5.8	24 (5+19)	**122** (2+120)	**15.5
all	**325** (301+24)	27 (1+26)	**42.2	**40**	1	**27.9	**563** (385+178)	49 (1+48)	**54.1
both	21 (19+2)	12 (0+12)	2.3	2	0	---	46 (36+10)	17 (0+17)	**5.8
each	22 (18+4)	4 (0+4)	**6.9	1	0	---	17 (10+7)	2 (1+1)	**7.9
TOTAL	380 (342+38)	250 (5+245)	**7.5	49	25	*4.2	650 (436+214)	190 (4+186)	**26.8

* significant at $p < .05$, ** significant at $p < .005$, ***sampled; s = synthetic, a = analytic

Overall, the figures reflect a strong preference for quantifiers *all*, *both* and *each* to be used in synthetically qualified NPs. Especially *all*, as in *we all*, *you all*, *they all*, etc., is much more common as a post-qualifying element. The differences in frequencies between the two structural variants are highly significant for all three persons (1PL, 2PL, 3PL). Here are some examples:

(367) They took **all of them** into t' yard and sorted them, catched **them all**. (*all of them* and *them all* objects in same sentence, WesBR, WES_008a.wav)

(368) Oh, how **we all** nearly died, **we all** screamed,... (*we all* subjects, ALI, NTT_012)

(369) ...well (unclear) father (/unclear) was a navy man, they made **us all** do a job at home. (*us all* object, LanD1B, LAN_001)

(370) **They all** ran it as a a rest place... (*they all* subject, MdxCG, MDX_001)

(371) She died, she was the last of **them all**... (*them all* in prepositional complement function, LanEG, LAN_012)

(372) ... 'cause she had to do for **all o' we**, see, when we come in the world. (*all of we*, prepositional complement, SRLM_RF, SOM_005)

(373) I, ah, oh, **all of they** were down there, see (reg sic=thy) they (/reg) (unclear) used to be good chapel people. (*all of they* subject, WflsWGP, WIL_010)

(374) ...and **each of they** used to go in the pitcher. (*each of they* subject, SRLM_WT, SOM_028)

(375) I couldn't keep **the both of them**, mind,... (*the both of them* object, TCA_RA, DEV_001)

Unlike the quantifiers *all*, *both* and *each*, numeric qualifiers are more commonly seen in analytic constructions. The corpus speakers generally prefer analytic variants like *the two of us* or *two of us* over synthetic variants like *us two* (also note the above-mentioned restrictions regarding larger multi-morphemic numbers). Once again, the differences in frequencies are highly significant for all three persons. In the 3PL, they are even underestimated since the pronouns' frequent use as demonstratives favours their appearance in synthetic constructions.

(376) Big families. [...] There was **ten of we**! (numeric qualifier in analytically qualified NP, W, SOM_030)

(377) Your mother 'd send you t' chip shop for t' family of us and there were **thirteen of us** altogether in my family, with sixpence... (numeric qualifier in analytically qualified NP, WigWHJ, LAN_020)

(378) There was **us two** went through. (numeric qualifier in synthetically qualified NP, HJT1, SFK_013)

(379) And **we three** used to saw, go and do the timber sawing... (numeric qualifier in synthetically qualified NP, SRLM_RM, SOM_014)

The corpus findings have major implications for the classification of qualified pronouns in variational linguistics. Especially the common use of *all* in synthetically qualified NPs shows that this quantifier is by no means restricted to the second person, and that the increasing use of *you all* in different varieties of English can not be discussed out of context. It is very likely that *you all*, including contracted *y'all*, forms part of a broader development which consists in the increasing use of 'pronoun + *all*' NPs. This of course implies that the common argument for the spreading of *you all* – as a compensatory strategy for the missing 2SG–2PL distinction in English – only explains part of the phenomenon (more on *you all* in 11.2).

Regarding the alleged preference for analytic constructions in non-standard varieties, qualified pronouns are a case in point that such a tendency, if present at all, can not be generalised regardlessly (compare ch. 13).

11.2 Second person plurals

In the second person, quantifying qualifiers like *all*, *two* or *both* serve the additional purpose of marking plural number on otherwise unspecified pronominal heads. While other Germanic languages have distinct 2PL forms which can also be qualified (e.g., German *ihr alle* 'you–2PL all–PL'), the loss of an overt 2SG–2PL distinction in English has resulted in one unspecified second person form *you* for both singular and plural addressees (diachronic details in 3.1). In addition, the loss of this distinction was accompanied by a loss of social deixis, making English "the most weakly socially encoded European language" of modern times (Mühlhäusler and Harré 1990: 134; compare Map D12 for a typology of politeness distinctions in pronouns). In recent decades, the underspecification of *you* has led to the diffusion of alternative 2PL forms in different varieties of the language, the official reintroduction of a distinct 2PL pronoun being among the most likely candidates for a future grammatical change in English (cf. Kortmann 2001; Wright 1997). Should this change ever be accomplished, English would change from the current 5 person system, which is typologically rare, to a 6 person system, which is the typologically most common system (cf. Forchheimer 1953).

Dialectal 2PL forms represent one of the most prominent pronoun features in English worldwide (cf. Kortmann and Szmrecsanyi 2004; Ingram 1978). Among these forms feature the dialectal homophones *yous* and *youse*, which are well-known in British English varieties influenced by Irish English,[130] as well as in American English. The usefulness of a *you* (SG) vs. *youse* (PL) distinction is illustrated in the following utterance from a female speaker from Belfast:

(380) So I said to our Jill and our Mary: 'Youse wash the dishes.' I might as well have said : 'You wash the dishes', for our Jill just got up and put her coat on and went out. (Harris 1993: 146)

Two other variants, *yunz* and *yinz* (also *yinz guys*), are commonly regarded as possible contractions of ScE/ IrE *you ones* (cf. Miller 2004: 49). These 2PL forms are so characteristic for Pittsburgh English that Pittsburgh natives with a heavy accent are often referred to as *yinzers*. Besides the different variants in *-s* – 'motivated forms' according to Brown and Levinson (1987: 23) – *you* can be combined with quantifiers like *all*, *both* and *each*, as described above, and it is highly productive with qualifier nouns. The large variety

130 Especially areas of Irish immigration such as the metropolitan areas of Liverpool, Newcastle and Manchester, also Glasgow, New York, and urban Australia; compare Trudgill (1999[2]: 92); Beal (1993: 205); Filppula (1999: 55); Cheshire, Edwards and Whittle (1993: 81). For Scots, see Miller (1993: 108); for Irish English, see Harris (1993: 139–140).

of strategies used to accomplish a 2SG–2PL distinction in English points towards a concrete need to mark an opposition which is no longer marked in the standard variety. Cheshire and Stein have pointed out:

> It is the marking of the distinction that is important to speakers, rather than the form that the distinction takes: this can be seen from the fact that the particular pronoun that develops in a new variety of English is not necessarily dependent on there being distinct singular and plural second person forms in the older varieties from which they have developed. (Cheshire and Stein 1997b: 8)

In the data investigated in this study, the two main strategies for marking plurality on *you* are the use of quantifiers (*you all*, *you two*, etc.; 49 cases), and the overt marking of number on the qualifier noun (e.g., *you bastard**s***; 24 cases). The examples include generic occurrences of the type shown in (388). It is interesting to note that none of the distinct 2PL forms known from other studies were found in the England data.[131] Three instances of *yous* which were found elsewhere in the corpus are shown in (381)–(383), but the corresponding recordings are unclear. One instance of *you guys* was found in the Scotland component of the corpus; it is shown in (384).[132]

(381) Then we used to have a game of kick back waters. {<u Int> How do you play that, Arthur?} Well, **yous** eh get a (gap 'indistinct') (2PL *yous*, LEI_001a.wav, not in dataset)

(382) ...and some of them had handles on, (trunc) wh- (/trunc) you see you could run and run with these at side on **yous** through this here... (2PL *yous*, NTT_016a.wav, not in dataset)

(383) I didnae recognise her at all, mind you it was hard to recognise any o' **yous**. (2PL *yous*, ELN_011a.wav, East Lothian, Scotland, not in dataset)

(384) ...I thought I was gonna stop it and start rugby and carry on playing rugby, and then I didn't want to get trampled by you. [...] I didn't want to get trampled by **you guys**, right. (2PL *you guys*, ELN_015, East Lothian, Scotland, not in dataset)

Unlike plural qualified *you*, singular qualified *you* is usually either pejorative or has an obviously informal quality. Out of 25 cases in the corpus, 21 have '*you* + *bloody*/ *stupid*/ *old*/ *daft* + *bastard*/ *bugger*/ *devil*/ *rascal*/ *villain*/ *bitch*/ *fool*'; the rest have *you boy*. This clearly distinguishes these occurrences from plural cases where qualifiers are used for number distinction and referent specification. Compare the following examples (also (351) and (352) above):

131 No instances were found of *you...together*, as in *Where are you together?* or *Come you on together!*; see Trudgill (2004a) for examples from East Anglia.

132 *You guys* is especially common in American English. The increasing gender-neutrality of *guy* has been considered as a motivating factor for the spreading of *you guys* (cf. McLennan 2004).

(385) I just happened to be stood there you know. So I says, he says, Oh dear, what have I to do, fine miself two shillings? I says, No, get out, get out! (v 'laughs') {<u IntER> (gap 'indistinct')} I (trunc) sa- (/trunc) I says, **You bloody local preacher**, **you bloody local shithouse**! I says, You 're not fit for your job, I says, And you go around preaching! I says, **You filthy, flatfooted, codfish-eyed swine**, you 're nothing else. (qualified pejorative *you*, singular, LanH1P, LAN_009e.wav)

(386) I thought, Oh **you naughty boy**! But of course you couldn't say anything. Oh I used to bless him when he done that to me I thought, Oh **you little rascal**. (qualified *you*, singular, TrbrAS, WIL_023)

(387) Here you are Mrs. Bestwick -- that 's for mi mother -- Here you are Mrs. Bestwick, here 's a lovely cabbage (trunc) to- (/trunc) it 'll (trunc) ma- (/trunc) it 'll serve **you all** today. They 'd they 'd give you the surplus you see, so that 's how you got a lot of that stuff... (*you all*, universal use, NotA65, NTT_007, youall6.wav)

(388) Uh, no, not a trap, a brake, you know, where **you all** sat around. (*you all*, generic, TCA_FP, DEV_005, youall1.wav)

(389) Well, and then the other three, you chucked it in and **you all three** took, in the summertime, **you all three** shared out what was left; you used to share out the, **the three of you** used to have the same. (*you all three*, HM, KEN_003, youall7.wav)

(390) I 'm going t' buy **you crew** somethin' special, when you go away t' Padstow. (qualified *you*, plural, GCS, SFK_010)

(391) All the old stuff what -- well, **you people** wun't eat it now. (qualified *you*, plural, HED, SFK_005)

(392) No, mi mother used to say, I 'd sooner bury **you lot** than put you on a farm, it was that bad. (*you lot*, plural object, FM, LND_003)

(393) ...you couldn't throw a ball until **you** had **all** gone round lamp, and then you threw it... (*you all*, quantificational use, Lang1p, LAN_003)

(394) {<u Int> Where did they have the swimming races?} Oh just off the island, you used to have to, for the swimming races, **you all** went out in the boat and dived off the side, to do your swimming. (*you all*, TCA_FP, DEV_005, youall2.wav)

<u>You all</u>

A brief note is in order regarding the well-known plural *you all*. In the corpus, *you all* occurs in its universal use (compare (387)) as well as in its quantificational use (compare (393)). Even if the latter is more common (see Table 11.1), a development towards the lexicalisation of *you all* is indicated by three empirically substantiated observations. First of all, a comparison of *you all* against *all of you* shows how popular the synthetic variant is in the corpus (29 *you all*, plus 11 '*you* V *all*', vs. only 1 *all of you*). Secondly, an attempt to paraphrase occurrences with '*you all* + numeral', like those in (389), shows just how much of a unit *you* and *all* represent. Following the acceptability judgement in (395), a qualified NP like *you all three* is more likely to be analysed as '[*you all*][*three*]', i.e. 'complex pronoun + numeric qualifier', than '[*you*][*all three*]':

(395) **all** of **you** > **you all**

all three of **you** > **you all** three/ ***you** three **all**

Thirdly, and perhaps most importantly, the contiguous pronunciation of *you all* shows the fusion of two components which have grown very close on the phonetic level.[133] The effect of lexicalisation on internal word boundaries is most noticeable in contracted *y'all*, which has been reported to be spreading in Southern American English (cf. Bernstein 2003).[134] In FRED, *y'all* is still outnumbered by non-contracted *you all* as heard in youall1.wav through youall15.wav on the CD-ROM. The situation may of course present itself differently among younger speakers.

11.3 'Oh deary me': delexicalised exclamations

This last section considers the use of pre-qualification by a qualifier adjective. Even if there are not enough examples in the corpus for a profound analysis, some interesting observations can be made regarding exclamations such as *Dear me* or *Good gracious me* (more frequent are occurrences with '*poor*/ *poor old*/ *lucky*/ *bloody*/ *silly*/ etc. + noun'). The only non-exclamatory case found in the data is repeated in (354, rep.). All other examples consist of interjections or exclamations with *me* as the pronominal head (15 cases, coded POS 19).

(354, rep.) I said, Well would you be happier if you were all together? And they said oh yes, they would. So 'course, **soft me**, I took them. (pre-qualified *me*, WesBL, WES_006a.wav)

The grammatical status of exclamations such as *Dear me* is unclear. They have been described as informal abbreviations of longer phrases such as *Dear God, save me* (cf. Mayhew 1908), but at the same time they are also being perceived as rather exalted.

In FRED, similar exclamations are rare despite the relatively high emotionality in some interviews. Nevertheless, exclamations of the *Dear me* or *Good gracious me* type were used by a variety of speakers, both male and female. A regional preference appears to be indicated for *dear me* type exclamations in the North: all *dear*/ *dearie me* were found in data from N, except two cases with *good gracious*/ *goodness me* in the SW.

133 Compare the Spanish 2PL pronoun *vosotros* ('you others'), and the polite forms of address *usted/ ustedes* (Catalan *vusted*), originally *vuestra merced* 'your Grace/ Highness'.

134 For the use of *y'all* with singular reference, see Hyman (2006). For an early reference to *you all* as a polite expression in southern American English, see Morrison (1926).

(396) **Oh dearie me**, I just, you know, Sally knows him better than me. (pre-qualified *me* in interjection/ exclamation, WesBR, WES_008b.wav)

(397) Nowadays you 'd think, **Good gracious me**, that 's a waste of time, you know. (pre-qualified *me* in interjection/ exclamation, CorJLN, CON_011)

(398) **What the deary me** did they call the blacksmith? Yes, bless us what did they call him. (pre-qualified *me* in interjection/ exclamation, WesBX, WES_009)

Even if the use of *me* in all of these examples seems so natural that it appears insignificant, the exclusive use of object case corroborates Schütze's claim that pronoun form variation does not occur with pre-nominal modifiers. Schütze's own examples include *The real me/ *I is finally emerging*; *Lucky me/ *I gets to clean the toilets*; *Dear me/ *I!*; *Lucky us/ *we!*; and *Poor them/ *they!* (cf. 2001: 215).

One of the above examples deserves special attention: in (398), the exclamatory NP not only has the colloquial form *dearie me* (one of five *dearie me* in the corpus), but it is also preceded by a definite article. A possible explanation, of course, is analogy to expressions such as *what the hell* or *what the heck*. Nevertheless, this example presents a particularly clear case of delexicalisation – a process which has been defined by Tognini-Bonelli as:

> the process through which a lexical item loses at least some of its original lexical value and often acquires other meanings and other functions within a larger unit. This kind of semantic impoverishment is again triggered when the dividing line between an item and its environment becomes blurred; then, strong collocational and/ or colligational patterns combine to create multi-word units where the function of the whole is of course different from the function of the individual parts. (Tognini-Bonelli 2001: 116)

It is well known that lexical items in collocation and components of phraseological chunks have a tendency to become delexicalised (cf. Sinclair 1991). This is certainly the case in exclamations like the ones presented above, where both components, *dear/ gracious/ goodness* and *me*, have undergone semantic bleaching. From a syntactic point of view, we can also speak of grammatical impoverishment, regarding both the grammatical function of each component and the grammatical relation between the two components. In the particular case of (398), *deary me* even loses some of its exclamatory power by being imbedded into the larger *what the...* phrase, which is exclamatory in itself. Similar examples are extremely hard to find. By chance, one was found in a novel by the American

writer Mary J. Holmes, and two others were spotted online. These additional examples show that the construction in (398) is extremely rare, but not unique:

(399) "**The deary me!**" said John Jr., mimicking his sister's manner, "how much lower is her origin than yours?" (*Lena Rivers*, 1856 novel by Mary J. Holmes, ch. XXXV; my bold print)

(400) In our little village there has been a tragedy
Oh the deary me, what a terrible tragedy,
Mary Ellen Bottomley, today should have been wed
It's a good job that she didn't now 'cause everybody said...
(*Mary Ellen at the church turned up*, old music hall song; my bold print)

(401) "**oh the deary me** the tension is brewing"
(online forum, http://www.planetcricket.net/forums/cricket-games-stories-strategies/heros-xi-return-cricket-07-a-28766-10.html; my bold print)

11.4 Summary

Simple personal pronouns like *you*, *we*, *us*, *they* and *them* are frequently accompanied by a noun or numeral, or by quantifiers such as *all*, *both* and *each*. Pronoun qualification has a variety of different functions, most importantly the specification of the pronominal referent and the quantification of plural referents. In English, it can be used to mark grammatical distinctions which are not, or no longer, marked in the standard pronominal paradigm, such as duality or a 2SG–2PL opposition.

In particular conversational situations, the drawbacks of under-specification are obviously felt by the speech community, leading to the introduction, or re-introduction, of overt morphological distinctions. Among the best-known examples in English are distinct 2PL forms such as *yous*/ *youse*, which are currently spreading in different varieties. The phenomenon ranges among the most prominent dialectal pronoun features of English worldwide and represents one of the most likely candidates for a future grammatical change. Among the different markers of plurality, qualifier *all* in 'pronoun + *all*' NPs appears to be on its way to grammaticalisation, especially in combination with *you* in *you all* (*y'all*).

Regarding variation in case, qualified pronouns are frequently mentioned in connection with non-standard O-forms in subject function, for example in sentences like *us boys used to go there*. It was shown in the analysis that qualification has a strong positive impact on the use of non-standard case in subject and, less frequently, object functions.

Comparing the synthetically qualified pronominals in the corpus to their analytic equivalents in the same dataset (*us all* vs. *all of us*, etc.), a strong preference was observed for speakers to use quantifiers *all*, *both* and *each* in synthetic constructions, and numeric qualifiers such as *two*, *three*, etc. in analytic constructions. This identifies 'qualifier type' as an important determinant of structural variation, adding another layer to the discussion about structural preferences in the different varieties of English.

Chapter 12
Pleonastic pronouns

Several pronoun phenomena in FRED involve the use of so-called *pleonastic pronouns*, i.e. pronouns which are deemed unnecessary in prescriptive grammars of English.[135] In the present study, however, the term pleonastic is not understood as 'superfluous', 'uneconomic' or even 'wrong'. Even if pleonastics are redundant insofar as they do not change the fundamental proposition of a sentence,[136] they have other functions which are often linked to emotive language use. Both the doubling of pronouns, as seen in disjunctives and resumptives, and the use of overt pronouns instead of zero pronouns are iconic of intensity and provide useful means for expressing emphasis and contrast. Overall, there is no doubt that pleonastic pronouns are more than just violations of an 'avoid pronoun' principle.[137]

Different types of pleonastics will be investigated in this chapter, and it will be shown that all of them represent supraregional features used by speakers throughout the country. Section 12.1 describes the use of overt *you* in second person imperatives such as *believe you me*. Sections 12.2 and 12.3 explore the speakers' use of additional object pronouns which turn standard intransitive verbs into transitive verbs (*I washed myself*), and standard monotransitive verbs into ditransitive verbs (*he was afraid to take him one*). Section 12.4 discusses the use of so-called ethic pleonastics in sentences like *Don't go to sleep on us!*; 12.5 focuses on the discourse functions of disjunctive pronouns, as in *I think so, me*; and 12.6 takes a look at resumptives in utterances like *A thing that you could wash en*. The first pleonastic feature described in this study was the use of dummy *it* in section 6.2. The use of pleonastic modifiers in adverbial participles was discussed in 10.5, remember *Me being light, I used to jump right at the top* (336). A brief outline of the respective historical developments was given in 3.1.

Note that pronoun dropping, which is also typical for informal English, is beyond the scope of this study (*Ever been there?*).

135 I use the term pleonastic according to Mitchell (1985: 114).

136 In a recent study, Szmrecsanyi and Kortmann (2009: 72–73) found that traditional L1 varieties exhibit the highest degree of grammatical redundancy in the sense of repetition of information, in comparison with higher-contact varieties.

137 According to Chomsky (1981: 65), the 'avoid pronoun' principle "might be regarded as a subcase of a conversational principle of not saying more than is required, or might be related to a principle of deletion-up-to-recoverability, but there is some reason to believe that it functions as a principle of grammar."

12.1 Overt *you* imperatives

One of the best-known pleonastic pronoun features in colloquial English is probably the use of overt *you* in second person imperatives such as *believe you me* or *look you*.[138] In the FRED interviews, most of these imperatives form part of conversations from the past which are retold by the speakers. The concerned utterances refer to requests or commands which somebody directed at the speaker at some point in the past, or which the speaker directed at somebody else. From a pragmatic point of view, the use of overt *you* in these utterances can be felt to reflect the speaker's opinion of the request or of the way it was made.

Overall, 127 overt *you* imperatives were found in the corpus (coded POS 60). This does not include occurrences with the common discourse marker *mind you* (POS 61) or prepositional phrases with *with*, as shown in the following examples:

(402) And it were a pocket watch, for me, **mind ye**. (*mind you*, LanEG, LAN_012)

(403) ...you see, he were one of the best dads as anybody could wish to have. [...] **Mind you**, eh I didn't get mi own way with him... (*mind you*, LanD1P, LAN_006e.wav)

(404) He looked across an' 'e said, Is tha' you Myrtle? An' she said, Yes Dad. Get on home **wi' ye** then, 'e said. (TrbrDP, WIL_004)

All pleonastic imperatives in the data contain second person *you* or corresponding dialectal variants. The only occurrence with overt *thee* is shown in (405). It forms part of a religious text which the speaker remembered from his childhood home. There are no examples with 1PL or 3SGm pronouns as known from earlier stages of English, for example in sentences like *retyre we to our chamber* (Shakespeare *Macbeth* II.ii) or *die he like a theefe* (Shakespeare *Pericles* IV.vi; cf. Rissanen 1999: 279).

(405) ...one big feature I 'd better mention was always two or three texts hanging on the wall. Why our people in those days were more religious than they are today. My mother used to say, I believe every word what that says up there: **Look thee**, put your trust in the Lord, you 'll never want. (overt *thee* imperative, IBME, SAL_023)

12.1.1 Structural subtypes

A first major result in the classification of the corpus examples was the confirmation of all imperative patterns known from earlier stages of English (cf. Rissanen 1999: 246). This continued use of pleonastic imperatives describes a relatively stable picture, in contrast to

138 Compare the following typological tendency described by Greenberg (1966: 44): "in imperative and hortatory forms, the second person is evidently the unmarked form and frequently has zero expression."

accounts which describe the phenomenon as an innovation.[139]

In the corpus, overt *you* in imperatives is used in combination with main verbs, with *do*-support, and in prohibitives with *don't* and *never*. In order to assess the phenomenon in its full extent, additional searches were conducted for *you*, adding cases which were lost in the random sampling process described in 4.2. The different syntactic patterns present themselves as follows:

'*you* + V': 105 cases
Speaker Number: 44; Speaker Proportion: 21%;
Mean Speaker–Case Ratio: 2.4
verbs *go/ don't go/ go back/ on/ off/ out* (36), *come* (10), *get/ get down/ off/ on* (5), *take* (5), *ask/ carry on/ give/ put/ tell* (3 each), *be/ bring/ count/ listen/ mention/ show/ stop* (2 each), *bet/ come out/ do/ eat/ have/ hit/ keep/ keep shut/ leave/ lie down/ look after/ move/ pay/ say/ send/ sit down/ start/ suit yourself/ think/ try* (1 each) (not including *you better...*)

'V + *you*': 10 cases
SN = 7; SP = 4.8%; MCR = 1.4
verbs *go* (3), *believe* (2), *come on* (2), *come/have a go/look* (1 each)

'*don't you* + V': 10 cases
SN = 10; SP = 4.8%; MCR = 1
verbs *be* (3), *go* (2), *talk/ give/ ask/ play/ make* (1)

'*do you* + V': 1 case; verb *come*

'*never you* + V': 1 case; verb *go*

There are several aspects which need to be considered in the classification and analysis of overt *you* imperatives. Regarding the most frequent pattern, '*you* + V', it can be difficult to distinguish between imperative and descriptive occurrences in sentences such as *you just make it up with some water* (SFK_007), or *you don't drink, you sip* (LAN_005). It can also be difficult to distinguish between imperatives and directions, for instance in sentences like *you go down Tollington Park that way* (LND_004). The above list only includes

139 Trudgill (2009: 109), for instance, attributes the use of pleonastic imperatives to an innovative increase in redundancy in the traditional low-contact Norfolk variety (2009: 109).

unambiguous examples. Regarding the second most frequent pattern, 'V + *you*', the numbers listed above do not include non-prototypical reflexives, as in g*et you back into bed* (LAN_002); these will, however, be discussed in 12.2.

Prosody and pausing can be decisive for a correct classification: sentences (411) and (412), for example, were included because of the contiguous pronunciation of verb and pronoun (*come on you*, not *come on, you*; see the soundfiles). Also included were two cases with exceptional *'s* variants, which were accidentally found when searching for 1PL *'s* 'us'; they can be seen in (413). It is clear from the surrounding co-text that the addressee in these occurrences is singular. It is unclear, however, whether the *'s* is pronominal, or whether it represents some sort of second-person inflectional ending on the verb ('goes on'; similar to 2SG endings in other varieties of English, e.g. AAVE). The *'s* could even be an abbreviation of *us*, in which case *Go's up and get* could be paraphrased as *Go up and get ... for us/ me* – an interpretation tying in with the benefactive objects discussed in 12.3.

Here are some empirical examples for the five patterns identified in FRED:

(406) And uh I said, Hello, here comes old shiney buttons, coming up over the -- I could see his head coming up, see. She said, **You come in**, she said, Have you have you been drinking? I said, I, well, I 've had a little, I haven't had a lot, I said, 'tis tiredness really. (overt *you* imperative, '*you* + V' pattern, TCA_RA, DEV_001a.wav)

(407) Well then, **you bring** her out then will you. (overt *you* imperative, '*you* + V' pattern, SRLM_RM, SOM_014)

(408) She says, Well go to work and come back on Friday at four o'clock and **you tell me** what they 've decided. (overt *you* imperative, '*you* + V' pattern, LanH1P, LAN_009)

(409) 'Cause when I worked up there for Palmers, he was the Mayor. She used to go 'round doing a little bit of local preaching, don't ever work for a religious woman 'cause there 's nobody worse to work for, **believe you me**. (overt *you* imperative, 'V + *you*' pattern, TCA_K, DEV_008c.wav, trba.)

(410) Go and fetch me that foc'sle funnel, I said and so and so. **Go you** and do that. That 'll help. (overt *you* imperative, 'V + *you*' pattern, JRS, SFK_011)

(411) ...eh my mother used to know when she was there eh she 'd, she 'd holler out, **Come on you**, Elsie, 'bout time you come in. (overt *you* imperative, 'V + *you*' pattern, FM, LND_003g.wav, trba.)

(412) And, if you went to the toilet, there was a woman in there, **Come on you**, out of it. (overt *you* imperative, 'V + *you*' pattern, FM, LND_003f.wav, trba.)

(413) ...and then he said to William, **Go 's up** and get the gilmat [gimlet] out my, my boat house, he said, and we put a hole in him, see. [...] Anyhow old William, William went down with this now and put a hole in 'im, and now they was looking on one another, so he said to William, **Go 's on**, try him, No, he said, William said, **You try him**. Now anyhow in the end the old man says, Well, I 'll have a go, see what it is -- salt water, I suppose. (second person *'s* in imperative, and '*you* + V' pattern, CAVA_WW, CON_006e.wav/ CON_006f.wav)

(414) They were no good (v 'laughs'). Now **don't you be afraid** to eat one of those [apples] -- they 're all right. (overt *you* imperative, '*don't you* + V' pattern, SRLM_SC, SOM_035a.wav, trba.)

(415) But you had to be careful if anybody came to help. You 'd say, **Don't you go in** with that one. (overt *you* imperative, '*don't you* + V' pattern, CorJLN, CON_011)

(416) Well, one thing led to another and he was (reg sic=gorn) going (/reg) t' take a rope and draw me across the behind with it. And I say, **Do you come** and try it! (overt *you* imperative, '*do you* + V' pattern, EJM, SFK_006)

(417) ...there used to be some baths on the Embankment, open air baths and I got (e 'clock chimes') going nicely, and I got the cramp and my friend as was teaching me he says, **Never you go in** no more, he says, You nearly lost your life. (overt *you* imperative, '*never you* + V' pattern, NotA30, NTT_003b.wav, trba.)

12.1.2 Distribution in the corpus

Overt *you* imperatives are a typical generalised feature with a high speaker number (52) and a Mean Case Ratio of 2.4 occurrences per speaker. They are used by a fourth of all speakers in the corpus (24.8%). Some speakers use different syntactic patterns (e.g., speaker FM uses '*you* + V', 'V + *you*' and '*don't you* + V'); others use the same pattern throughout the conversation (e.g., speaker NotA8 with 8 '*you* + V' cases).

A closer look reveals an interesting tendency: while most of the identified patterns are used only occasionally (e.g., 'V + *you*' with 10 occurrences by 10 different speakers), speakers who use the pre-verbal '*you* + V' pattern use it more regularly. This tendency, together with the pattern's general high frequency, indicates that '*you* + V' is more deeply entrenched, or somehow comes more naturally to speakers, than the other patterns. This is likely to have implications for the future coexistence of the different variants.

The corpus results put into perspective some of the claims which have been made in the literature. Most importantly perhaps, the results are only partly consistent with descriptions of the historical development of imperatives after Early Modern English. For instance, the two main historical changes described for second person imperatives in Modern English are the rise of overt pronouns in pre-verbal position and their disuse in post-verbal position (cf. Ihalainen 1994). Judging by the FRED results, the post-verbal pattern is effectively less common, but it is still in use. Furthermore, 'V + *you*' is also attested in the SED incidental material (21,6,iJ VII.4.11. *look you after him*), and its use has been noted in other varieties of English, such as Belfast English (Henry 1995, for instance, mentions examples like *Believe you me* or *Go you away*).

Two aspects that cannot be ignored in this context are, one the one hand, collocational preferences of certain verbs, and, on the other hand, differences in stress between the pre-verbal and post-verbal patterns. Regarding the first aspect, it can be observed that whilst overt *you* imperatives are not collocationally conditioned themselves (see the great variety of verbs in the pattern list), certain verbs like *believe* are restricted to certain imperative patterns (compare *believe you me* vs. **you believe me*). Regarding the second aspect, it appears that the stress in post-verbal examples automatically falls on V (compare (418b)), whereas the pre-verbal pattern allows for stress to shift to either V or *you*, depending on which element the speaker wants to emphasise (compare (418a); also (413) *No, William said, Yóu try him*). By allowing for stress and emphasis to shift, the pre-verbal pattern possesses an inherent flexibility which is lacking in the less common post-verbal variant. The two patterns are, therefore, not totally synonymous.

(418) a. **You go** and do that.

b. **Go you** and do that.

12.2 Pleonastic reflexives after standard intransitive verbs

The next three sections, 12.2 through 12.4, are dedicated to pleonastic objects in positions where prescriptive grammars stipulate no pronoun, or the use of a zero pronoun. An overt object in such cases increases the verb's standard valency by adding an extra argument to the sentence. The pleonastic pronoun, which in some cases expresses a non-prototypical object or reflexive function, is grammatically unnecessary in so far as it is not required for the main proposition. It is, however, not meaningless. In many cases, pleonastic objects are associated with emphasis and speaker involvement. They can also be used to underline the agent-status of the sentence subject.

In the current section, we take a look at pleonastics in direct object position. These are pronouns which turn standard intransitive and inherently reflexive verbs into overtly reflexive verbs, as in *I washed myself* or *get you back into bed* (irrespective of the pronoun's case). In 12.3 we will look at pleonastics in indirect object positions, for instance in sentences like *I'll have me a pint*. The historic development of both types was described in 3.1. In the spreadsheets, the two uses are coded POS 52 (reflexive function) and POS 53 (non-standard benefactive).

Here are some first examples from the corpus:

(419) ...mi mother had to pull mi shirt off. For I couldn't **wash me**. (pleonastic *me* with inherently reflexive *wash*, NTT_011, NotA99)

(420) ....we used to know how fat mam was getting but we didn't know anything. And then eh I heard a a noise one night and I got up and I thought, well mi mamma must be bad and I, I went in the room and I said, Is mi mamma bad? And then mi daddy he said, Now go and **get you back** into bed mi lass, he said, She 'll be all right come morning. (pleonastic *you* with intransitive *get back*, LAN_002, LAN_002b.wav)

(421) And they [two horses] come all the way home to Warren's stables, and anyhow **they$_i$ turned them$_i$ round**, and somebody was up the stables and they knew where they were and they went back, met the old man coming a-puffing and blowing... (pleonastic *them* with intransitive *turn round*, CAVA_WW, CON_006g.wav)

12.2.1 Historical and typological perspective

Examples from earlier stages of English show that overt self-forms have a long tradition of co-occurring with inherently reflexive and self-directed verbs, despite the general tendency for intensifiers to attach primarily to object pronouns "in those cases where a verb expressed an other-directed activity" (König and Siemund 1997: 103). Reflexively used O-forms also occur with inherently reflexive verbs in Old, Middle and Early Modern English texts (both dative and accusative case forms; see 3.1).

From a typological perspective, the phenomenon is not uncommon. In Standard Dutch, for example, weak reflexives occur as objects of inherently reflexive verbs such as *zich gedragen* 'to behave' or *zich herinneren* 'to remember': *Jan herinnert zich dat verhaal well* 'Jan remembers REFL that story well' (example from SAND, cf. Barbiers et al. 2005). In Standard German, we find equivalent constructions with *sich benehmen* and *sich erinnern*.[140] Dutch and German behave differently from present-day Standard English, where inherently reflexive verbs require a null implied object as the most heavily reduced available option with a reflexive interpretation (cf. Safir 2004: 65). This is not always the case in other varieties of English. In the dialectological literature, pleonastic reflexives are frequently mentioned in connection with verbs of grooming such as *wash*, *shave* and *dress*, verbs of posture such as *sit* and *stand up*, and non-translational motion verbs such as *turn*, *bow* and *stretch*.[141] Various occurrences with these and other verbs are documented in the SED; for instance:

140 Zwart (2005) lists examples with Middle English reflexive verbs (motion, cognition, psych, behaviour) where the pronoun has turned to zero in Modern English whereas Dutch has kept an overt pronoun.

141 Gestural reflexives of the type *Steve stretched (himself)* have been described by Safir (2004: 122).

(422) *have yourself a sit-down/ sit yourself down/ sit you down* (VIII.3.3)
lay/ lie you/ yourself down/ he laid himself down/ he lied himself (VIII.3.6)
get yourself away (VIII.7.9)
he drowned himself (IX.9.6; also passives of the type *whose uncle was drowned*)
I wash myself/ wash me (IX.11.1; incid.mat. 33,1 IX.11.1.)

12.2.2 Pleonastic reflexives in FRED

In different previous studies, the use of pleonastic reflexives has been described as a predominantly Northern or Scottish English feature (e.g., Wright 1905: 276, *get thee dressed while I wash me*). In the empirical data in this study, however, pleonastic reflexives represent a supraregional phenomenon used by a large number of speakers in the corpus (SN = 38, SP = 18.1%, MCR = 1.4.). Their areal distribution is visualised in Figure 12.1.

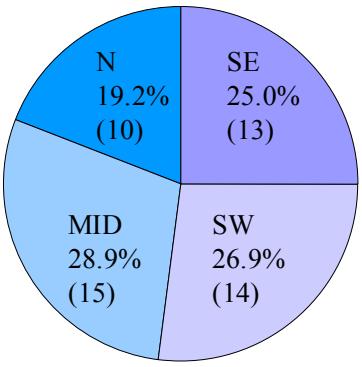


Figure 12.1: Areal distribution: pleonastic reflexives (weighted proportions, weighted absolute numbers in brackets)

Table 12.1 shows the distribution of different case forms (O-forms and self-forms) among the pleonastic reflexives in the corpus. The numbers include all cases with a coreferential subject and direct object (excluding other pleonastic objects, as in *it's wondering me for years* (LAN_009)). As expected, the results show a clear preference for self-forms in 88.5% of all occurrences. At the same time, the continued use of O-forms shows that they have not died out as markers of derived intransitivity, a concept which has been defined by

Siemund (2003: 490) as "the non-referential use of reflexive pronouns with the function of intransitivizing otherwise transitive predicates." Siemund himself states that "what we do know with certainty is that the use of personal pronouns as markers of derived intransitivity (pleonastic pronouns) suddenly dies out when complex intensifiers and reflexives become grammaticalized (around 1500)." However, he also mentions that "the original situation is preserved in regional varieties" (ibid., p. 491). This is certainly the case in the FRED data. After all, O-forms account for over 11.5% of all occurrences in this context.

Table 12.1: Case form distribution in pleonastic reflexives

intransitive→ transitive	O-form	self-form	TOTAL
1SG	2	2	4
2SG/ PL	2	14	16
3SGm	0	2	2
3SGf	0	5	5
3SGn	0	8	8
1PL	1	9	10
3PL	1	6	7
TOTAL	6	46	52

12.2.3 Possible explanations

According to Faltz (1985), the acceptability of O-form reflexives with inherently reflexive verbs has a logical explanation:

> If we believe that the reason that it is natural for an emphatic to mark reflexive coreference is that such coreference is normally a marked situation, we would predict that the emphatic would *not* be used to mark the objects of verbs like *wash* or *dress* when they are coreferent with the subject, since for these verbs coreference is normal. (Faltz 1985: 242)

However, O-form reflexives also appear where coreferentiality or self-directedness is not automatically given, as seen in examples (420) and (421). Regarding the verbs in our data, we find plenty of examples which go beyond the categories usually mentioned in the literature, including behavioural verbs such as *behave*, *misbehave*, *show off* and *adapt*,

action verbs like *choke* and *smash*, and different constructions with *get* (*get ready*/ *pissed*/ *out*/ *round*/ *wet*).[142]

(423) They say, **Get yourself ready**. (pleonastic *yourself* with *get ready*, EF1, SFK_003)

(424) ...and I seed my sister die. Her **choked herself** with this thing in her throat. (pleonastic *herself* with *choke*, IBJP, SAL_038)

(425) You mustn't tip it away, or else you 'll **drown yourself**, **get yourself all wet**! (pleonastic *yourself* with *drown* and *get wet*, SRLM_CA, SOM_009)

(426) And he was in a boat one day, **showing himself off** a bit among the rowers, you know, with his private boat,... (pleonastic *himself* with inherently reflexive *show off*, IBAD, SAL_018)

(427) Doesn't drink or anything like that, goes not to **booze herself** but to (pause) starts talking... (pleonastic *herself* with intransitive *booze*, KentAW1, KEN_009)

(428) ...and out they 'd come, so we used to have to **move ourselves** and get up on the corner there. (pleonastic *ourselves* with intransitive motion verb *move*, LA, LND_005)

From a wider perspective, the corpus examples fit in nicely with well-known typological tendencies in the coding of middle situations. According to Kemmer (1993), grooming verbs are the one type of body action verbs which most strongly requires the signalling of coreferentiality between the agent/ experiencer and patient/ recipient in a sentence (more than self-induced motion, body posture or non-translational motion verbs). In FRED, over one third of all examples belong to this category (19 out of 52), including the verbs *wash*, *shave*, *bath*, *dress*, *dress up*, *change* and *keep clean*. A comparison of grooming verbs with and without pleonastics supports the first impression that the former option is very common. The verb *wash*, for example, has 12 standard occurrences and 7 occurrences with pleonastic reflexives. The common saying quoted in (432) also shows how widespread the use of overtly reflexive *wash* is.

Kemmer (1993) situates the use of overt reflexives with grooming verbs somewhere between direct-reflexive and one-participant events in her radiational model. A logical justification is provided by Langacker (1990), who explains that the subject in such utterances acts on itself in a slightly different way from how it would act on another participant or object, since it only acts on an 'active zone' of the subject. The overt reflexive – no matter which case is used – refers to this active zone and is, therefore, not

142 More examples are available in the pronoun spreadsheets. Verbs include: *adapt*, *bath*, *behave* (in Safir's view, pronominals used with *behave* are really manner adverbs), *booze* (occasionally used with overt self-forms online), *change*, *choke*, *dress (up)*, *drown*, *dry out*, *get back*/ *ready*/ *wet*/ *round*/ *out*/ *pissed*, *keep clean*/ *warm*/ *to sth.*, *merge*, *misbehave*, *rest*, *run*, *shave*, *show off*, *slag*, *smash*, *turn*, *wash*. Excluded are cases with uncertain coreferentiality, e.g. *she caught her at the top of the tunnel* (LAN_012c.wav). Also excluded is *tat*, as in *She used to tat 'erself up* (SAL_035). *Tat* originally refers to the making of a special type of lace; *to tat oneself up* does not usually appear in dictionaries but is occasionally found with overt self-forms online.

out of place. In other words, the self-directed *washing* in sentences like *I washed me* or *I washed myself* differs from the other-directed *washing* in sentences like *I washed him*. In pleonastic occurrences, this distinction is overtly marked.

(429) And, uhm, we used to have to stand in this copper and **bath ourself**, wash our hair and all. (pleonastic *ourselves* with inherently reflexive *bath*, LA, LND_005c.wav)

(430) No -- well, you was supposed to wash and shave before you went on parade, and I mean to say you 'd got to shave and **wash you** in cold water, no hot water, no, oh, ah. (pleonastic *you* with inherently reflexive *wash*, IBLE, SAL_025)

(431) ...I got m' best rig out, what I 'd got with me, and **dressed m'self** so I made m'self look right smart. (pleonastic *myself* with inherently reflexive *dress*, EJM, SFK_006)

(432) Well, that 's an old common saying, **it 's washed itself** in t' beck, now, whether it has or it hasn't I don't know, I 've never deliberately seen one, stop in the beck and **wash itself**. That 's an old common saying, it 'll have **washed itself** in t' beck! Aye. (pleonastic *itself* with inherently reflexive *wash*, WesAE, WES_002, beck = small stream)

With other types of verbs, pleonastics appear only occasionally and less systematically. We find, for instance, one example with the action verb *run*, shown in (438) (as compared to over 80 standard *run*); one example with *rest*, as seen in (434) (1 pleonastic, 7 standard occurrences); and some examples with *get* and *get back*, as seen in (420) and (433) (8 pleonastics, more than 500 standard occurrences). The only non-translational body verb found in the data is *turn*, which was shown in example (421).

(433) {<u IntLW> So when did you tend to get more women weavers in the mill?} Well they were gradually coming to that when, at about -- when I -- second, third year I was there, that was around early -- then it eventually more or less **got itself** like that, that it was pretty well all women weavers. (*itself* with intransitive *get* 'become', TrbrEGC, WIL_015)

(434) ...and on returning he felt rather tired so he thought he would close his eyes to see if he could **rest himself**,... (pleonastic *himself* with intransitive *rest*, IBCH, SAL_034)

Contrary to what one might expect, disambiguation does not usually explain the appearance of pleonastic reflexives. Even verbs which are self-directed by default show a preference for the overtly marked option. One such verb is *behave*, which is only listed with an overt reflexive in some dictionaries, probably due to the fact that *behave oneself* is considered slightly informal. In the corpus, *behave* is represented by 9 pleonastic vs. 4 intransitive examples, as seen in (435a) and (435b).

(435) a. Yes. You had to **behave yourself**, you know, or you 'd get a bang on your ear-hole. (*behave* with pleonastic reflexive, IBTJ, SAL_033)

b. Aye. Very much so, there was no messing about then, was it, mi father he used a stick then so you had to **behave**. (*behave* without reflexive, NotA52, NTT_005)

Some examples from the corpus suggest that pleonastic pronouns fulfil specific discourse functions which are not usually mentioned in the literature. For instance, speakers use them to underline the volitional agent status of the sentence subject, as shown in (436) and (437). In negation, the same feature can be used to stress that the agent-status is impeded, as seen in (419). A similar effect can even be obtained with inanimate referents, as illustrated in (438). In a certain way, the use of *itself* in this example animates the subject *coach*, giving it a personal quality as if acting of its own accord. The impression is reinforced by the use of gendered *him* in the same sentence.

Similar uses can be observed in (439) and (440). Note that *merge oneself into* and similar constructions are frequently found online but are not usually listed in dictionaries. In sentence (439), the meaning corresponds to 'established itself' or 'was introduced/ implemented'. When used in narratives, pleonastics with a comparable animation effect can serve as a stylistic means to make a story or event appear more vivid, even with clearly non-agentive and non-volitional subjects as in (440).

(436) ...why we **kept ourselves** more or less to one quality, as a rule,... (pleonastic *ourselves* with intransitive *to keep to sth.*, TrbrEGC, WIL_015)

(437) They **adapted themselves** to conducting. (pleonastic *themselves* with intransitive *adapt*, WigWO, LAN_023)

(438) Then **the coach would run itself down** and they used to turn him into the platform at Wellington. (pleonstic *itself* with intransitive motion verb *run*, referent: a slip coach, IBBeBo, SAL_011)

(439) Now that 's all single yarn and there 's a particular weave for that and it was made for the cavalry actually but of course it **merged itself** into the breeches trade like with the hunters, farmers, all that type of people had these cloths... (pleonastic *itself* with inherently reflexive *merge*, TrbrEGC, WIL_016)

(440) He jumped out of her, let her run away, and she crashed through the wall, and dropped down about sixty feet in this ravine, and **smashed herself** all to pieces. (pleonastic *herself* with inherently reflexive *smash*, referent: an engine, SRLM_RM, SOM_014)

Two last occurrences which are slightly different from the above examples but nevertheless deserve mentioning are shown in (441). Both sentences represent middle voice cases with the usually intransitive verb *hurt*. The subject in both cases is *belly*, i.e. a body part, or active zone, of the speaker, and the direct object is the speaker himself. The subject here is clearly non-volitional, although not quite non-agentive, which is why the situation is not a prototypical two-participant event (cf. Kemmer 1994). At the same time, however, the described event is also not a one-participant or reflexive event. The subject referent is not coreferential with the object referent, the former being a body part of the latter. What makes these examples so interesting is that the factual one-participant event is

subjectively perceived by the speaker as a two-participant event, adding a different dimension to the usual distinction. The situation in (441) is best described as a non-prototypical or perceived two-participant event. The reiteration of pleonastic *me* can be explained by the speaker's dissociating himself (the one suffering the pain) from his own body (causing the pain), as two separate entities. This feeling is iconically translated into a transitive construction. Similar examples are known from other languages, for instance German *mein Arm tut (mir) weh* 'my arm hurts (me)'.

(441) If you were constipated my ma would say, What 's up with you? I says, Oh **my belly's hurting me**. She, Have you been to the back, I says, No, I says, **Mi belly hurts me**. (pleonastic object *me* after *hurt*, NotA8, NTT_013d.wav, trba.)

12.3 Benefactive objects

The use of overt indirect objects after standard monotransitive verbs is frequently mentioned as a characteristic of English vernaculars. Coreferential occurrences of *me*, *you* or *us* in sentences like *I'll have me a beer* or *we gonna have us a great time*, for example, are well known from American English varieties (cf. Siemund 2003). Besides their use of pleonastic objects, similar sentences also attract attention for the fact that these pronouns frequently have object case. This specific aspect is probably the reason why the phenomenon is often discussed in connection with the use of O-form pronouns after standard ditransitive verbs, as in *why don't you go and buy you some chocolate* or *I'll fetch me a drink* (instead of *buy yourself* and *fetch myself*).[143] In these sentences, too, the indirect object can be considered obsolete if the self-directedness of the action is inferable from the context. Compare the pleonastic pronoun after monotransitive *have* in (442a) against the O-form reflexive in (442b). In the following, we will try to keep these two uses separate.

(442) a. Why don't you **have you** some chocolate? (non-standard benefactive)

b. Why don't you **buy you** some chocolate? (non-standard reflexive, replacing *yourself*)

The fact that, in the history of English, reflexives with *self* established themselves earlier in direct than in indirect object function is a possible explanation for the continued use of O-forms in sentences like (442a) and (442b). König and Siemund (1997: 100–101) have noted that in texts from the 16th and 17th centuries "plain pronouns can still be found as markers of co-reference in the position of indirect objects." Occasional examples from the

143 Compare the SED incidental material, Sa, 7 IV.6., fieldworker note: "I'll fetch me (myself) a drink."

literature show that such pronouns continued to exist later on, for instance in Defoe's *Robinson Crusoe* "and we made us a little tent" (first published 1719; 1831 edition, p. 180). In 1949, Stratton called these "needless objectives" a "quaint old-fashioned habit":

> A quaint old-fashioned habit is to put unnecessary pronouns after verbs. This is tolerated in old persons, but in anyone else it is frowned upon. Don't imitate others and fall into this habit. 'Down at Perk's general store I bought *me* some thick underwear.' 'First, find *you* some nice plank about 3/4 inch thick.' 'I would get *me* an interesting story and read it, and then I would make *me* some good candy.' Put the ideas above into expressions like 'bought *for myself*,' 'find *yourself*,' 'get *myself*,' 'make *myself*.' (Stratton 1949: 161)

Technically speaking, pleonastic objects in sentences like *I'll have me/ myself a beer* are not reflexive, even if they happen to be self-forms. A more adequate term proposed by Siemund (2003: 492) is that of 'benefactive datives', where S is the beneficiary of the action performed by S. The benefactive role, of course, applies to pleonastic and standard occurrences in equal measure. In the following, pleonastic occurrences of the type *I'll have me a beer* will be referred to as *non-standard benefactives*.

12.3.1 Non-pleonastic occurrences

In the FRED corpus, benefactive objects are striking but not very common. Non-pleonastic objects of verbs like *get*, *borrow*, *fill* or *take* 'get' do not, strictly speaking, fall into the category of non-standard benefactives. In many dictionaries, these verbs are followed by self-forms in indirect object position, and the benefactive in such cases can be paraphrased by a prepositional phrase with *for*: *I borrowed me/ myself some money* or *I borrowed some money for me/ myself*. The use of non-pleonastic benefactives is of course noteworthy for the observable variation in case.

The following occurrences were not included in the analysis: cases where coreferentiality is uncertain, as in *She liked a drop of wine, used to go and get* ***her*** *a bottle of wine* (LND_004c.wav); displaced intensifiers, as in *I forget* ***myself*** *what I 'm talking about* (NBL_006); and vernacular uses of certain verbs, for example ditransitive *learn* 'teach', as in *My dad* ***learned me*** *the loom tunin'* (WIL_018), or *take* 'bring', as in *a lady once* ***took me*** *one* (LAN_012).

(443) I took ʻem to Stokesley, and I lost five shillings a head, and five shillings a head in those days was a lot o' money. [...] it hit me very hard ʻcause I **borrowed me money** to buy them with off mi dad. (benefactive *me* with verb *borrow*, YksSL, YKS_003a.wav, trba.)

(444) Cause I wanted some clothes, and I thought that 's the only way I could **get me clothes**. (benefactive *me* with verb *get*, FM, LND_003h.wav)

(445) Yes, but I loved the sweet peas that were growing in the garden. I loved to go across and **get myself** a bunch. (benefactive *myself* with verb *get*, WesZs, WES_019)

(446) ...and then of course when you 'd **got you** so much water out again you used to go and fill it up again, fill your fill your boiler up again, oh it used to last all day. (benefactive *you* with verb *get*, NotA58, NTT_006, NTT_006b.wav, trba.)

(447) So I say, Look here [...] Go across t' ol' Mabel's and **get y'self** a bloomin' cheese roll or (reg sic=suffin) something (/reg) before she shut. (benefactive *yourself* with verb *get*, EAA1, SFK_002)

(448) It was a sitting down job, but anyone$_i$ could **get him$_i$** a job at Coalport. (benefactive *him* with verb *get*, IBMrsD, SAL_022)

(449) But I used to go into the coal-face and fill the tubs. I used to maybe go in at -- maybes twenty minutes before the miners, the fillers, and I 'd **fill myself** a couple of tubs, well we used to get one and ten pence a tub. (benefactive *myself* with verb *fill*, DurML, DUR_001c.wav, trba.)

(450) And I got two or three out [sweets], put them in mi pocket, and he was afraid to **take him** one... (benefactive *him* with verb *take*, LanEG, LAN_012)

As Siemund (2003) has illustrated with similar sentences from Southern American English, examples like the ones shown here are not specific to British English. This agrees with the areal distribution in FRED, which identifies benefactive O-forms as a supraregional feature, hence potentially a general feature of spoken English (4 SE, 5 MID, 11 N; SN = 12, SP = 5.7%, MCR = 1.7). In addition, the distribution of case forms shows that benefactive objects are preferably encoded by O-forms (20 O-forms vs. 5 self-forms).

12.3.2 Pleonastic occurrences

The most defining characteristic of non-standard benefactives is that they are not directly licensed by the argument structure of the verb. Unlike the benefactive objects described above, they can not be paraphrased by a prepositional phrase with *for*. In the corpus, only 6 examples were found, all of them after *have* or *get down* meaning ʻeat' or ʻdrink':[144]

(451) There 's always the bloke there playing the piano and you used to **get a few pints down you**, maybes a game of dominoes,... (pleonastic *you* with verb *get down* ʻdrink', DurML, DUR_001)

144 Other constructions listed by Siemund (2003: 493) do not appear in FRED but may of course be found elsewhere. These include verbs denoting psychological states such as *want*, *love* or *like*, as in *I love me some baked beans*, or double objects, as in *I'm gonna write me a letter to my cousin*.

(452) Now, we drove to t' house, and while they were putting t' coffin in, they brought us o'er a drink o' neat whisky. Well, I 'd never been used to that and I didn't want to take it, but t' chap who was driving t' next cab to me, said, **get it down you**. It 'll warm you up. (pleonastic *you* with verb *get down* 'drink', YksSL, YKS_003)

(453) I was standing there **having me** a drink and I didn't know who he was, and I said to the landlord afterwards, I said, Who is that gentleman? (benefactive *me* with *have*, IBFB, SAL_013)

(454) But as children eh, there was little eh, see, you gamble with alleys eh, because -- {<u IntER> Hm.} Eh, **have you** what you won, and if you were a good player at alley well you always have plenty in the house... (benefactive *you* with *have*, Lang1p, LAN_003)

(455) ...he was half-way through mi hair and he popped in and **had him** half a pint, and he didn't finish mi hair. (benefactive *him* with *have*, IBHP, SAL_037)

(456) ...we went in the pantry and popped a finger in the water, and up jumped the pike, **have him** a meal if he could get one. (benefactive *him* with *have*, IBCD, SAL_021)

In addition to the benefactive aspect, the pleonastics shown in examples (453)–(456) express an additional meaning which is inaccessible to a formalist approach. By using an overt indirect object in (455) and (456), for instance, the speakers indirectly express their personal attitude towards the described action, or towards the described agent. In (455), the non-standard benefactive stresses the hairdresser's laziness and indifference towards his customer in an almost humorous way, and the non-standard benefactive in (456) stresses the fish's go-get-it attitude. In (453), on the other hand, the additional pronoun serves to underline that the speaker was *just* having a drink and thinking of nothing bad when somebody he didn't know suddenly started talking to him. Overall, the pleonastics in these examples do not change the fundamental proposition of the sentence but express the speaker's attitude towards the proposition of the sentence – a mechanism which will be revisited in the next section on ethic datives.

Last but not least, we also find some benefactives where the indirect object is not coreferential with the sentence subject. Similar cases are not usually mentioned in the literature. They are, however, not unusual in requests like (457), (458) or (459), where they appear to aim for a mollifying effect on the addressee. By expressly mentioning the beneficiary or recipient in the request, these benefactives also emphasise the speaker's personal interest in seeing the request followed through.

(457) ...so o' course I got the old-fashioned wooden bucket and I lowered it down the hold and I say t' the chaps what was down there, **Put me** some herrin' in there, Will you please? (benefactive non-coreferential *me* in request, JRS, SFK_011)

(458) ...and they always used to come from all over, like, to mi mother, and, Will you **patch me** this, like, or Will you **patch me** that? (benefactive non-coreferential *me* in request, LanC5P, LAN_005)

(459) And we used to say to her, **Save us** your jam jars, and we used to collect these jam jars and take them 'round the rag shop, where we can get a ha'penny each for them, you see. (benefactive *us* with verb *save*, non-coreferential, LA, LND_005)

(460) I says, Well they **put me** one beam in, I says, But they 've **put me** a new Jacquard up and it were all strings, you know,... (benefactive *me* with verb put, non-coreferential, LanH1P, LAN_009)

All of the benefactive objects we have seen in this section express some sort of speaker involvement, which brings them close to the ethic constructions discussed next.

12.4 Ethic datives and ethic PPs

The so-called ethic dative – also ethical dative or dativus ethicus – is a feature known from earlier stages of English. In Early Modern English, ethic datives were a popular stylistic means for higher intensity of expression which could give the dramatic dialogue a more vivid, maybe also a more colloquial flavour (compare 3.1.3). Compare the following example from Marlowe' *Doctor Faustus*:

(33a, rep.) I seeing that, tooke him by the leg, and never rested pulling, till I had pul'd **me** his leg quite off. (Marlowe *The Tragical History of Doctor Faustus* 1248)

Similar to the non-standard benefactives discussed above, ethic datives are not directly licensed by the argument structure of the verb. Unlike non-standard benefactives, however, the pleonastic object in this usage is not the beneficiary of the situation (compare *I had me a drink* vs. *I pulled me his leg off*). In a somewhat moral tone (hence 'ethic'), the additional pronoun expresses "an emotional attitude of a speaker in a non-explicit form toward a remarkable propositional state of affairs as a whole or restricted to the predication part of it." (Hübler 1998: 61). Gutzmann has defined the conventionalised meaning of ethic datives as follows:

> The Ethic Dative expresses the speaker's attitude towards the proposition P_s of the sentence in which it stands by an independent proposition P_{DE}.
> (Gutzmann 2007: 282; my translation from German)

Gutzmann's addition that "the independent proposition P_{DE} can be paraphrased as 'the speaker expresses his or her personal interest that the demanded action be executed or performed'" (ibid., p. 282) only applies if an action is demanded (which is not the case in

sentences like (33a) or (33b)). Sentence (463), for example, does not fit into the two subgroups he identifies for German ethic datives, which are the 'Aufforderungs-Ethicus' (ethicus of request) and 'Ausrufs-Ethicus' (ethicus of exclamation).

In a theory of conventional implicatures, which Gutzmann (ibid., p. 283) proposes as a basis for a formal logic of ethic datives, the non-standard benefactives discussed above are similar to ethic datives if they share the following properties:

- if they express their own proposition and contribute to the meaning of the utterance independently from the 'what is said' proposition of the sentence;
- if removing the pleonastic pronoun from the sentence does not alter the at-issue proposition of the sentence ('multidimensionality');
- if the conversational implicature of the pronoun is detachable, i.e. if the pronoun can not be paraphrased by a different construction without losing its specific meaning ('detachability' of the implicature; compare *I was having me a drink* vs. *I was having a drink for/ to me/ myself*);
- if they express the point of view of the speaker ('speaker referentiality');
- if the involvement of the speaker can not be revoked without making the sentence unacceptable ('linearity'; e.g., (456) can not be paraphrased as **the pike would have him a meal if he could get one -- and I don't care if he did*).

The fact that no regular ethic datives of the type seen in (33a) were found in the corpus, supports their classification as an archaic feature.[145] Gutzmann's definition is, however, directly applicable to some prepositional constructions found in FRED. Despite the different structure, these PPs express meanings which are very similar to those of ethic datives. I therefore refer to them as *ethic PPs*:

(461) He says, You talk to me like that, he says, I, I 'll knock your head **off to me**. And so I picked a tin of fruit up, I says, Now you come near me, I say, And yours 'll be off (v 'laughs') It cured him. (ethic PP, NotA52, NTT_005a.wav)

(462) <u WesZb> I think we all tend to think that we had some lovely summers, but I don't think we altogether remember the rain we had as well. [...] But you see, we don't -- don't go to sleep **on us**! <u WesZs> I 'm not. (ethic PP, WES_019)

145 The *OED Online* lists the following examples: *They also killed us Captain Feluck* (Travestin Siege Newheusel 48, year 1685) and *They wounded us only one Man* (London Gazette No.4864/1, year 1711).

(463) For the quantity we were making you see we used to -- when we were very busy -- put on six presses a day. We really had to get rid of ʻem once the apples were rotting **on us**. (ethic PP, SRLM_HM, SOM_024)

Just like ethic datives, these ethic PPs or ʻpronouns of caring' (cf. Shetter 2000) express the personal and emotional involvement of the speaker as well as his or her attitude towards the proposition of the sentence. One small reservation with respect to the conventional implicatures defined by Gutzmann concerns the detachability of the implicature, which is not always given. If we paraphrased an example like *I'll knock your head off to me* (ethic PP in (461)) as *I'll knock me your head off* (ethic dative as found in earlier stages of English), the fundamental implicature would be preserved in both constructions: ʻthe speaker is angry and threatens the hearer by adverting to his physical strength'.

Interestingly, the ethic PPs found in FRED do not only correspond in meaning to archaic ethic datives, they are also translatable into present-day ethic datives in other languages: sentence (462), for instance, translates into German as *Schlafe uns bloß nicht ein!* ʻgo-to-sleep us–DAT not', and (463) can be translated as *als die Äpfel uns verfaulten* ʻwhen the apples us–DAT rotted'.[146]

Similar examples involving the use of an *on* PP have been described by Rot (1994: 478) for Yorkshire, Scottish, Irish, American and New Zealand English. According to Rot, such additional PPs can be paraphrased by an *in spite of* construction, so that a sentence like *The cow died on me* would be interpreted as *The cow died in spite of all my efforts to save it*. The two FRED examples in (462) and (463), however, suggest a different interpretation which is closer to the ʻethic of disadvantage' described by Filppula (2004) for Irish English. In this meaning, the ethic PP implies "a disadvantage of some kind or another from the point of view of the speaker or some other person" (ibid., p. 97). Filppula's own Irish English recordings include, for example, the sentence *this day the fire went out on him*, where the additional proposition expressed by the *on* PP is similar to that in (462) and (463).

Overall, a semantic definition of ethic datives (as found in earlier stages of English) and ethic PPs (as found in FRED and other present-day non-standard varieties) needs to consider subtle differences in meaning between individual examples. The conventionalised meaning of ethic datives suggested by Gutzmann (above) is broad enough to encompass these differences. I suggest rephrasing it as follows:

146 For a description of ethic datives in German compare Werner (1973).

(464) Working principle of ethic datives and corresponding constructions (based on Gutzmann 2007)

The ethic dative, and corresponding constructions such as ethic PPs, express the speaker's attitude towards the proposition P_s of the sentence in which they stand by an independent proposition P_{DE}. Ethic pleonastics can be used to express the emotional involvement of the speaker, including his or her opinion, approval or disapproval of the action or event described by P_s. The more specific meaning of each utterance has to be inferred from the context.

Ethic pleonastics can, for example, express the speaker's personal interest that a request be performed (cf. Gutzmann 2007: 282, 'ethic of request'; not included in the conventionalised meaning); they can imply a disadvantage of some kind from the point of view of the speaker or some other person (Filppula's 'ethic of disadvantage'); or they can imply that something happened despite the speaker's effort to prevent it (Rot's 'in spite of' paraphrasis). Other occurrences may yet have slightly different meanings.

"I Can Do Science Me!"
(Brainiac: Science Abuse,
UK television programme)

12.5 Disjunctive pronouns: emphasis, topicalisation, specification

Disjunctive pronouns are another type of pleonastics which add meaning to a given sentence without changing its fundamental proposition. In the literature, they are also known as 'unconnected' or 'absolute' pronouns. Similar to these other terms, *disjunctive*, from Latin *disjunctio* 'separation' or 'dissociation', refers to the pronoun's dislocated position outside the main sentence structure in utterances like *I like flowers, me*. As marked syntactic structures, disjunctive utterances are "more associated with the vernacular, oral end of the varieties continuum" (Stein 1997: 41). As doubling devices, they are iconic of intensity (ibid., p. 40).

It is well known that in some languages personal pronouns in disjunctive positions take default case: consider French *C'est moi/ Moi, je l'aime/ Je n'aime pas, moi* ('It's me'/ 'Me, I love him'/ 'I don't like it, me'). In English, disjunctives have long been recognised

as a syntactic environment where variation in case is observed. In 1886, for instance, Elworthy proposed a subcategorisation of English personal pronouns into 'full form', 'unemphatic', 'interrogative enclitic' and 'unconnected'. In his *Grammar of the Dialect of West Somerset*, Kruisinga (1905: 35) named disjunctives as one of four nominative case environments where pronouns may change their form ('emphatic', 'unemphatic', 'interrogative', and 'unconnected/ absolute'). And in the same year, Wright listed no-verb disjunctives as one of various O-form environments (*Who did that? Her.*), next to cleft sentences (*it was her that did it*), coordinated NPs (*him and me did it*), and antecedents of relative pronouns (*him that did that ought to be hanged*). More recently, Beal (1993: 210–11) described the use of post-positioned O-forms in the Tyneside variety in sentences like *They're useless, them* or *I could just go a toasted sandwich, me*,[147] and Miller (2004: 68) quoted the 3SG example *Oh it was a loss it* from broad Scots.

12.5.1 Disjunctive uses in FRED

In this analysis, the focal point will be the distribution of pronoun case in different disjunctive constructions, complemented by a qualitative description of the different functional applications observed in the data. Two main structural parameters will be distinguished: the disjunctive's position relative to the co-indexed NP (which is usually the subject of the main sentence) and coordination. Regarding the first parameter, a binary distinction will be made between disjunctives following the coindexed NP, as shown in (465), and disjunctives preceding the coindexed NP, as shown in (466). Based on their position in the sentence, disjunctive pronouns will be referred to as either *pre-positioned* or *post-positioned* (for no-verb utterances, see 8.3.4 and 8.5.4).[148]

(465) ...and then me for flowers, I don't think (trunc) th- (/trunc), buttercups, daisies, dandelions I 'd pick -- eh I love flowers, **me**. (post-positioned disjunctive *me*, LanD1P, LAN_006b.wav)

(466) **Me**, I go twice every Sunday. (pre-positioned disjunctive *me*, WesZS, WES_019)

Regarding the second parameter, the internal composition of the NP containing the disjunctive will be considered with the aim of determining functional and distributional differences between simplex and coordinated cases, as shown in (467) and (468).

147 Also disjunctive *this*, as in *My skirt's too short, this*.
148 The concept of subject and object territory is not useful for this analysis (cf. Quirk et al. 1985: 337).

(467) He was always dressed, **him**. (post-positioned disjunctive *him*, NblMN, NBL_003a.wav)

(468) I can always remember my old dad on my mother's bike, when he used to go out Sunday morning, **him and** one or two more friends. (post-positioned disjunctive *him* in coordination, MdxCG, MDX_001f.wav)

The four structural subtypes distinguished in the analysis are shown in Table 12.2. In the spreadsheets, they are coded as POS 70/ 71/ 72/ 73 respectively.

Table 12.2: Disjunctives, 4 structural subtypes

disjunctives	post-positioned	pre-positioned
simplex	Type 1: *I always go, me.*	Type 2: *Me, I always go.*
coordinated	Type 3: *We always go, me and him.*	Type 4: *Me and him, we always go.*

Table 12.3. shows a breakdown of simplex vs. coordinated disjunctives by pronoun case. Most occurrences have pronouns of the 1SG (*I*, *me*, *myself*), 3SGm (*he*, *him*) and 3PL (*they*, *them*).

Table 12.3: Simplex vs. coordinated disjunctives, breakdown by case form

disjunctives	simplex				coordinated			
	S-form	O-form	self-form	TOTAL	S-form	O-form	self-form	TOTAL
1SG	2	11	8	21	35	52	5	92
2SG + PL		0	0	0		1	0	1
3SGm	4	5	1	10	6	12	0	18
3SGf	0	0	0	0	1	4	0	5
3SGn		0	0	0		0	0	0
1PL	0	2	1	3	0	1	0	1
3PL	5	15	1	21	0	2	0	2
TOTAL	11	33	11	55	42	72	5	119

The numbers in Table 12.3 are based on all pronominal disjunctives found in the data, including cases where the corresponding subject NP has to be inferred, as in *[They are] Limmers them* (DUR_001). Also included are cases where the disjunctive is followed by an impersonal structure, as in *I and this boy I mentioned [...], it was our job to go out and feed all the cattle* (SOM_011). In cases which could classify as either Type 3 or Type 4

(compare Table 12.2), the prosody and pausing had to be considered. One example is the sentence *{she 'd been to a dance [her and her sister who was the Landlady up at the London Inn] they 'd had a dance up there}* (DEV_001b.wav; classified as Type 4). Self-form occurrences were generally included if not in adnominal or adverbial intensifier function (compare ch. 3.3). The numbers in Table 12.3 do not include:

- disjunctive nominals or personal names, as in *my father, he had the dairy over there then* (SOM_006), or *Bob Edwards, he used to come too* (CON_006);
- repetitions of the type 'pronoun + V', although these are also characteristic of spoken language, for example *he 's a marvellous lad, he is* (SAL_035), or *they were all our bosses, them were* (SAL_022);
- preposings, as in *her I didn't know* (DEV_007);
- qualified pronouns of any sort, whether qualified by a nominal (*us lads, when we did start work in the brick field* (KEN_003b.wav), *we were at school, us kids* (YKS_009a.wav)), relative clause (*Of course us what could see a bit when it was time to go in, we used to...* (NotA58)), or locative adjunct (*Same if anybody were confined, next door neighbour or her from down here up to Woodcock Road* (LAN_017));
- self-forms forming part of a 'pronoun + self-form' NP (*I myself, I...*);
- anacolutha, as in *to us, we classed it in three categories* (CON_010);
- and potential transcription errors, as in *Aye I get down line I* (Yks_010), where the second *I* is not audible in the recording.

The corpus results show several interesting tendencies regarding the distribution of case forms and disjunctive subtypes. These tendencies will in the following be summarised under the headings of 'Preference for object case', 'Disjunction and coordination', 'Discourse functions' and 'Areal distribution'.

12.5.2 Preference for object case

First of all, the distribution of the different pronouns supports a classification of disjunctives as a default case environment (cf. Schütze 2001). In the data, there is a clear preference for O-forms, in alternation with S- and self-forms. According to Ihalainen's description of the spreading of O-forms in English disjunctives – a change from below

which "has not yet gone to completion" (1994: 108) – a further reduction of S-form proportions can be expected in the future. Speakers in FRED, however, still use a substantial amount of S-form disjunctives, especially in the 1SG.

The distribution is typical for an environment where morphological case is not determined structurally. As a result, there is a noticeable preference for default O-forms (60.3%, see examples (470)–(473)) but also a substantial number of other case forms (30.5% S-forms, 9.2% self-forms, see examples (474)–(477) and (478)–(480)). This quantitative distribution is similar to the distribution observed in no-verb utterances (see ch. 8), but with a more substantial amount of non-default forms. In some utterances, for example (469), it appears that the lack of case assignment, or of easily accessible case assignment rules, caused the speaker to vacillate between different pronouns forms.

(469) That was the first bike I ever had -- **myself**, **me**, **I** eh, as I 'm now talking, Charlie Greenfield. (post-positioned disjunctive *myself/ me/ I*, MdxCG, MDX_001b.wav)

(470) Oh well I would be about eight year old **me** but t' others were a bit older... (disjunctive O-form, subject function, WigMB, LAN_017)

(471) Well, he 'd give me permission, **him**, to get the manager's stand there with the mint or whatever I had, see. (disjunctive O-form, subject function, JG, LND_004a.wav, trba.)

(472) Eh, then Christmastime we used to go and fetch a, they used to have a great big tree brought to 'em in, in Bridgerow school. And **us**, eleven, twelve and that, we used to have to decorate the lamps (unclear) with them (/unclear) and fill all this up and little presents for all the youngsters and... (disjunctive O-form, subject function, Lei2, LEI_002b.wav)

(473) Might have worked at the Choppington brickyard, I couldn't tell you who worked at the Choppington one. But them 's all Battleton Station ones, **them**. (disjunctive O-form, subject function, NblMN, NBL_003c.wav)

(474) {<u IntER> What about when you were married, did you go for a holiday?} Yes, I did, **I.** (disjunctive S-form, LanW1P, LAN_004a.wav)

(475) ...he died uh, about three years ago. He started with nothing **he**. I don't know whether you noticed on Mapperley just when you get this side... (disjunctive S-form, NotA99, NTT_011b.wav)

(476) {<u IntMG> So did you sell from the works?} From the works, yeah. A lot of people used to come in, visitors and that sort, yeh. This is one I made, **he**. Don't know what I made him for as I say (v 'laughs') (disjunctive S-form, referent: a pot, SRLM_CA, SOM_009a.wav, trba.)

(477) They go around to each places every year. They don't uh very often go to two places in one year **they**. (disjunctive S-form, NotA99, NTT_011c.wav)

(478) {<u Int> Was there any difficulty in getting um ingredients during the war?} Oh yes, very, very hard job to get eggs. Now we used to use an a (trunc) s- (/trunc) material called flu-egg. It was a sort of a liquid egg but uh, I don't think it ever saw egg, **miself**. (disjunctive self-form, TCA_FP, DEV_005d.wav)

(479) That, ehr, the captain used to make a little bit on buying the food by twisting, one thing and another, so, **hisself**, he only got a pound a week. (disjunctive self-form, SRLM_WT, SOM_028)

(480) And **ourselves**, we had eh embroidered dresses on for Sundays. (disjunctive self-form, YksMW, YKS_008a.wav)

12.5.3 Disjunction and coordination – two distinct mechanisms

In the corpus examples, there exist perceivable differences in the syntactic distribution and functionality of simplex vs. coordinated disjunctives. Simplex, non-coordinated disjunctives are more frequently post-positioned, i.e. they more frequently follow than precede the coindexed NP. In post-position (Type 1) all three case forms occur, whereas the pre-position (Type 2) seems confined to non-S forms in simplex cases as seen in (472) and (479). This is all the more surprising considering that all pre-positioned cases have subject function. The use of S-forms in simplex Type 1 disjunctives, therefore, represents a typical case of context-restricted variation (S-forms: 11 x Type 1; O-forms: 25 x Type 1, 8 x Type 2; self-forms: 3 x Type 1, 8 x Type 2).

Coordinated disjunctives show a different distribution. S-forms – according to Emonds (1986) an unnatural option in disjunctives – are noticeably more frequent in coordinated cases. The two syntactic positions (Type 3 and Type 4) are equally frequent, with 63 post-positioned vs. 56 pre-positioned occurrences. In both positions, all three case forms appear.

Regarding the use of non-standard case forms, disjunction has a clear positive influence on the use of simplex O-form subjects. The comparison of non-coordinated disjunctive vs. non-disjunctive cases in Table 12.4 shows that O-form subjects are significantly more frequent in disjunctive utterances. Disjunction also has a positive influence on the use of simplex self-form subjects, as shown in Table 12.5.

Table 12.4: O-form subjects in disjunctive vs. non-disjunctive occurrences (based on all pronouns with distinct S- and O-forms)

O-form subjects	disjunctive (f1)	non-disjunctive (f2)	Δ(f1,f2)
simplex (f3)	60.4% (32/ 53)	0.3% (293/ 101606)	**σ = 8.94*** **$p = 7*10^{-6}$**
coordinated (f4)	58.6% (65/ 111)	52.4% (87/ 166)	σ = 1.01 $p = .61$
Δ(f3,f4)	σ = 0.22 $p = .91$	**σ = 13.45*** **$p = 2*10^{-11}$**	

* significant at $p < .005$

Table 12.5: Self-form subjects in disjunctive vs. non-disjunctive occurrences

self-form subjects	disjunctive (f1)	non-disjunctive (f2)	Δ(f1,f2)
simplex (f3)	20.8% (11/ 53)	0% (0/ 101606)	**σ = 3.73** ***p* = .06**
coordinated (f4)	4.5% (5/ 111)	5.4% (9/ 166)	σ = 0.35 *p* = .86
Δ(f3,f4)	σ = 2.75 *p* = .17	σ = 3.09 *p* = .12	

An interesting observation if we compare the influence of disjunction against that of coordination is that, in coordinated cases, there is no statistically significant difference between O-form frequencies in disjunctive (58.5%) vs. non-disjunctive subjects (52.4%), and the same holds for self-form subjects (4.5% vs. 5.4%). Furthermore, a comparison of simplex vs. coordinated disjunctives shows that both syntactic environments have similar O-form proportions (60.4% vs. 58.6%).

In combination, these results show that disjunction has a noticeable impact on the use of non-standard case in simplex subjects, but not in coordinated subjects. In other words, pronoun variation in simplex disjunctives is conditioned by disjunction, whereas pronoun variation in coordinated disjunctives is conditioned by coordination. This distinction is easily overlooked if simplex and coordinated cases are not analysed separately.

12.5.4 Discourse functions

In the investigated conversations, the four disjunctive subtypes distinguished in Table 12.2 have slightly different functions. Based on the empirical examples, three main discourse functions were identified: emphasis as the main function of post-positioned disjunctives; topicalisation as the main function of pre-positioned disjunctives; and specification as the main function of coordinated disjunctives. Let us take a closer look at these functions.

Post-positioned disjunctives were shown in several examples above. Some more occurrences are shown in (481)–(486). The usual function of these disjunctives is to emphasise the coindexed NP by repetition. This is of course most noticeable where the pronoun carries heavy stress, as heard in (481):

(481) ...they can't do that (k 'car passing') {<u IntER> No.} Anyway, she was before me, and I couldn't hear what they were saying and she went out and she was still sobbing

bitterly, and eh I can feel mi blood boiling **me** (v 'laughs') (post-positioned disjunctive *me*, LanH1P, LAN_009d.wav)

(482) The big mansion kind of house, poor black sheep, **me**. (post-positioned disjunctive *me*, H3L, LAN_010a.wav)

The relatively common technique of pronoun doubling (compare the resumptive pronouns in 12.6) was, for example, described by Joseph Wright (1905: 270), who stated that "[i]n all the dialects of Sc. and Eng. there is a tendency to introduce a redundant personal pronoun after a noun when emphasis is required." These *emphatic disjunctives* are functionally similar to adnominal intensifiers, with the difference that disjunctives can be emphatic without implicating comparison. Consider the following two sentences:

(483) a. **He** was always dressed, **him**. (emphatic disjunctive)

b. **He himself** was always dressed. (adnominal intensifier)

In self-form occurrences, an additional distinction is required between emphatic disjunctives and 'misplaced' intensifiers, i.e. intensifiers which do not occupy a standard intensifier position (see (484)). In (485), for example, the speaker's actual answer in (485a) is probably closest to (485b). In addition, it should also be noted that disjunctives are not always separated prosodically from the rest of the sentence, as seen in (486).

(484) I don't believe being in a Trade Union and acting like these miners, because I think **miself**, really, that 's a lot of agitators ('misplaced' intensifier, TCA_WH, DEV_009a.wav)

(485) a. {Which was the rough end?} I think the bottom end was the roughest, *miself*. I think so. (CA, LND_001b.wav)

b. I think the bottom end was the roughest, *me*. I think so. (disjunctive pronoun)

c. I *myself* think the bottom end was the roughest. I think so. (adnominal intensifier)

d. I think the bottom end was the roughest myself. I think so. ('misplaced' intensifier)

(486) {<u IntAH> You didn't do too much when you were working?} Oh no, no, we, we, we couldn't afford to go off **us**, you know. (disjunctive *us*, SRLM_RF, SOM_005e.wav)

Less frequent than post-positioned disjunctives are pre-positioned disjunctives as shown in (487)–(489) and some of the above examples. Here, too, emphasis is most noticeable where the pleonastic pronoun receives extra stress, for instance in (487). From a discourse analytical perspective, the left-dislocated disjunctive functions as a topicaliser by anticipating the coindexed NP. Topic fronting by use of disjunctives is one of various

topicalisation techniques used in the interviews (also, for example, changes in word order, as in *Fifty-odd now I is*, WES_017).

In their function as emphatic pronouns and topicalisers, pre-positioned disjunctives have additional discourse-structuring and emotive components. They can also, for example, be used as floor-holding devices or to buy encoding time (cf. Pawley and Syder 1983: 563). According to Stein (1997: 43), pre-positioned disjunctives "model the actual production of speech: what comes first to the mind comes first in the utterance, a natural production principle that applies to all topicalizations." Furthermore, these disjunctives are typical oral features that "mirror the pragmatics of the production process and directly make reference to non-verbalized context" (ibid., p. 43). Their low acceptance in written Standard English is hence predictable.

(487) I 'm sure **me**, I could go and control them. (disjunctive topicaliser, NotA85, NTT_010a.wav)

(488) ...and of course the recession started in nineteen twenty-six and there was quite a few laid off and it carried on, and **myself**, I was suspended,... (disjunctive topicaliser, WesBX, WES_009)

(489) Now **them**, it must have been one of someone of these people from the cottages must have complained... (disjunctive topicaliser, Lang1p, LAN_003)

The main function of coordinated disjunctives is to specify the referents of a coindexed plural NP (usually the sentence subject). In addition, coordinated disjunctives also possess certain discourse-structuring, emphatic and even stylistic functions which become clear if we rephrase examples from the corpus as non-disjunctive sentences (see (494b) and (495b)).

(490) **I and** the wife then we were living over the Lamb Inn, we had rooms over the Lamb Inn... (coordinated disjunctive, TrbrGR, WIL_001d.wav)

(491) Mum used to play the piano and we used to sing hymns around the piano, **me and** father, I can't sing in tune, I 'm either sharp or flat... (coordinated disjunctive, TCA_K, DEV_008d.wav)

(492) We had a postman when I went to school three days a week. And then there was, uh, some of them was wanting him to come oftener, so they see who got most letters, **us or** people inland here and... (coordinated disjunctive, KentSPR, KEN_008b.wav)

(493) ...and it was which could swear the hardest they said, **her or** the doctor. (coordinated disjunctive, WesDB, WES_015)

(494) a. **Him and I, we** 've fished together for many years,... (coordinated disjunctive, MdxCG, MDX_001g.wav)

b. **Him and I** have fished together for many years... (non-disjunctive variant)

(495) a. And then in our dinner hour **we** had to go round with the flipping butter, **I and my sister**. (coordinated disjunctive, SRLM_PT, SOM_007)

b. And then in our dinner hour **I and my sister** had to go round with the flipping butter. (non-disjunctive variant)

12.5.5. Areal distribution

Disjunctive pronouns are a supraregional phenomenon used by a considerable number of speakers in the corpus (SN = 87, SP = 41.4%, MCR = 2). This includes both simplex and coordinated occurrences, and both pre-positioned and post-positioned occurrences. Figure 12.2 visualises the areal distribution of disjunctive pronouns in FRED.

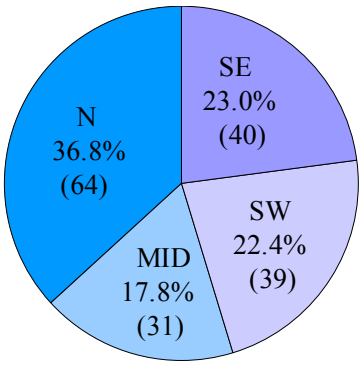


Figure 12.2: Areal distribution: disjunctive pronouns
(all case forms, weighted proportions,
absolute occurrences in brackets)

12.6 Resumptive pronouns

The last type of pleonastics in the corpus are so-called resumptive pronouns (coded POS 74). These are typically pronouns which repeat the antecedent of a relative clause, appearing "in the position of the (alleged) extraction site" (Salzmann 2006: 172). According to Salzmann (2006: 329), resumptives in English "only occur where extraction is impossible," for instance in sentences like *This is the man that I don't know [why nobody likes *Ø/ him]*. The empirical findings discussed in this section, however, show resumptive pronouns in places where extraction is actually possible. In most cases,

resumptives appear in the alleged zero slot of a simple relative clause, but examples like (504) show that they can also be used elsewhere (note that (504) has no noticeable pause between the left-dislocated NP and the main clause).

Pleonastic resumptives are a minor phenomenon in terms of frequency and are easily overlooked in data that are not tagged accordingly. An additional search routine was therefore applied in this study, looking for all combinations of relative pronouns (*that*/ *what*/ *which*/ *who*/ *as*) + S-/ O-/ self-forms, as well as all sentence-final pronouns. Not included were (a) grammatically ambiguous occurrences such as *my dad were one of them as [since?]* ***he*** *had to read the work to the nook* (LAN_016); (b) resumptives after left dislocations containing a relative clause, as in *the frocks that we used to have* ***they*** *used to be leg o' mutton sleeves* (LAN_002); and (c) other pronoun types, as in *they could only come at certain times you see, according to the tide, which probably you 'd know* ***that*** (LAN_010).

All of the 27 occurrences in FRED have third-person pronouns (*he*/ *him*/ *she*/ *it*/ *they*/ *them*). They appear in different syntactic roles (usually within a relative clause), including subjects (15 cases),

(496) ˈCourse it was, in them days there were a lot of treacle about, cow treacle, you know. {What 's cow treacle?} Well it 's what they fed, you used to put it on hay **that it** was mouldy, you know, hay, bad hay,... (resumptive subject *it*, WesBR, WES_008e.wav)

(497) I think, I think Mrs Johnson was the sister to this woman, **which she** (trunc) ju- (/trunc) just recently died. (resumptive subject *she*, CA, LND_001d.wav, trba.)

(498) ...and I used to get two buckets of boiling hot water **which ˈey** was a penny each which went to the infirmary. (resumptive subject *they*, H3L, LAN_010f.wav, trba.)

(499) We had one, one of our old customers, Old Dickie, **that he** had to go up the tree with a toast. (resumptive subject *he*, SRLM_HM, SOM_024)

objects (8 cases),

(500) And the eel comes up and bites at the eels, at the worms, and then they 've got their big tank besides them that they just flicks **it** out and the eel drops off into the bath. (resumptive object *it*, SRLM_CK, SOM_004a.wav; repeated from (264))

(501) Or, he said, a pound of butter that they give 'm **it** in a bowl... (resumptive object *it*, WesAY, WES_004a.wav)

(502) Before we were eh using those pressure lights that you had to fill **ˈem** up first with eh methylated spirits... (resumptive object *them*, TCA_RA, DEV_001c.wav)

(503) When we were youngsters going to school, we ˈad a, like a rubber collar. Like a rubber collar thing you know that you could wash **en**. (resumptive object *it*/ *him*, WflsGW, WIL_011)

and prepositional complements (4 cases):[149]

(504) Some pots here we used to put scollops in ʻ**em**, the other ones we used to leave plain on the top. (resumptive prep.compl. *them*, SRLM_CA, SOM_009d.wav, trba.)

(505) There ’s a great big walled round ash pit as you used to put your ashes in **it**, and when that was full they used to come and empty that. (resumptive prep.compl. *it*, IBLE, SAL_024)

(506) I think it was Bristol engineering that they did a lot of work for **them** during the war, yes. (resumptive prep.compl. *them*, WesBX, WES_009)

Similar to the other pleonastics described above, the resumptives in these examples are not functionally void. According to Bakker (1974), fo instance, the use of a pleonastic resumptive, or ʻpronomen abundans’, is caused by the weakening of the relative pronoun, typically in non-restrictive relative clauses. The resumptive is needed for additional transparency and emphasis, especially in unplanned or fast speech (cf. Filppula 1999: 195; Miller 1993: 112–113, ʻfast speech hypothesis’). Stein, too, notes that resumptives are “typical of oral language, which does not have the benefit of editability, or of preserving the sensory input, and which therefore is in need of more redundancy” (1997: 39).

For Mühlhäusler and Harré (1990: 251), pleonastic resumptives are “a candidate for ʻnatural syntax’,” based on their use in child language and natural second-language acquisition. The authors even state that the absence of pleonastic resumptives “from standard languages such as Standard English may then illustrate an instance of cultural suppression of a natural construction” (ibid.). As has been noted by Stein (1997: 39), one idea about language which is particularly noticeable in 18th century writings is that “languages ought to be ʻlogical’,” and that “there ought not to be double segmentalizations.” According to Stein (ibid.), this “logical factor would account for the injunction against structures that seem to have the same referent, such as resumptive pronouns.”

Strong evidence against the alleged logic of the English prescriptive constraint is provided by other standard languages where the use of pleonastic resumptives is fully accepted. Such languages are, for instance, Greek (cf. Bakker 1974) or Catalan, as shown in the following example:[150]

149 Corresponding German examples can be found in Salzmann (2006: 158).

150 According to Fischer (2002: 26), “the origin of clitic pronouns [in Catalan] is to be found in contexts involving pronominal resumption of a topicalized argument. From these contexts it generalizes to others in which the co-referential or antecedent item is elsewhere in the discourse or extra-linguistic context.” For a transformational approach to resumptives compare Ross (1967).

(507) El llibre **el** hi vaig donar al Jordi.
the book$_1$ it$_1$ him$_2$ I-gave to Jordi$_2$
'I gave the book to Jordi'

In Standard German, the coreferential element is obligatory in sentences like (508), where the use of zero would be ungrammatical:

(508) Der Ring, von dem ich hoffe, dass du **ihn** morgen kaufst.
the ring$_1$ of which I hope that you it$_1$ tomorrow buy
'the ring which I hope you will buy tomorrow'

(adapted from Salzmann 2006: 299)

Regarding their distribution in the FRED corpus, resumptive pronouns are not particularly frequent, but they can be shown to appear in different parts of the country (1 SE, 10 SW, 6 MID, 10 N), similar to the other pleonastics discussed above. Although occurrences are rare in the individual interviews, the feature was used by a substantial number of speakers (SN = 23, SP = 11%, MCR = 1.2).

12.7 Summary

The different types of pleonastic pronouns discussed in this chapter are so diverse that they seem unconnected at first view. In sections 12.1 through 12.6, we had a look at overt *you* imperatives (*believe you me*), pleonastic reflexives following inherently reflexive verbs (*I washed myself*), non-standard benefactives following standard monotransitive verbs (*I'll have me a beer*), ethic PPs (*I'll knock your head off to me*; *don't go to sleep on us*), disjunctive pronouns preceding the sentence subject (*Me, I go twice every Sunday*) or following the sentence subject (*I love flowers, me*), and resumptive pronouns (*hay that it was mouldy*; *a collar thing that you could wash en*).

From a structural point of view, all of these pleonastics appear where Standard English requires a null implied subject or object, which explains their low acceptance in formal writing and prescriptive grammars. Pleonastic pronouns have not only been criticised as unnecessary and uneconomic but also as illogical, especially concerning instances of pronoun doubling. What their castigators usually ignore is the fact that the very same uses were considered acceptable in earlier stages of English and are typologically attested in other languages today.

The main focus in the discussion was on functional aspects. While it is true that pleonastic pronouns do not influence or change the primary proposition of the sentence, they carry a certain functional load. It is generally agreed that pleonastics are typical of oral language, which is naturally in need of more redundancy and transparency. The functional core distilled from the different phenomena described in this chapter consists in an extra dimension, or independent proposition, which is added to the proposition of the sentence (compare the Working principle of ethic datives in (464)). The exact nature of this independent proposition varies, depending on the overall context and intention of the speaker. Pleonastics are, for example, naturally iconic of intensity, both as doubling devices (disjunctives, resumptives) and in contrast with standard zero. In the discourse, they can be used to add emphasis to a given utterance, or as topicalisers. It was shown how pleonastics are used to reflect the speaker's perspective on a particular situation or event (perceived two-participant events, active zone), the speaker's emotional involvement with, or personal attitude towards, that situation or event (non-standard benefactives, ethic PPs), and, last but not least, the perception of other subject referents (volitional agent status, animation effect).

Chapter 13
Genitive variation

13.1 The two genitives of English

Dialects and contact varieties have been claimed to show a general preference for analytic structures, as opposed to standard language varieties and the Latin synthetic model (cf. Stein 1997). The reasoning behind this claim is that analyticity, being the "less dense and more transparent means of expression" (Biber et al. 1999: 302), facilitates a reduction of processing effort. In this chapter, the distribution of synthetic and analytic variants in dialect data will be tested on genitive constructions with pronominal modifiers. The main interest here lies in the distribution and possible determinants of variation of semantically equivalent or near-equivalent *s*-genitives (synthetic genitives as in *its door*) and *of*-genitives (analytic genitives as in *the door of it*). The analysis is kept separate from the distribution of pronoun case in English genitives, which will be discussed in chapter 14.

The following terminological conventions apply:

The two genitive structures will be referred to as *s*- and *of-genitives* (coded POS 80 and 81 in the spreadsheets). In *s*-genitives, the analysed syntactic position is that of the pre-nominal determiner, as in ***her*** *house*. This is the position of the genitive clitic *'s*, or Saxon genitive, with nominal NPs in English, as in ***Mary's*** *house*. To refer to genitives like *her house* as *s*-genitives is justified by the view that pronominal determiners represent fusions of an underlying structure [pronoun + POSS], for example [*she* + POSS] > *her*, which corresponds to the nominal equivalent of [noun + POSS], for example [*Mary* + POSS] > *Mary's* (the same view has been expressed by Taylor 1996: 1).

In the *of*-genitive, also known as the Norman genitive, the analysed position is that of the prepositional complement, as in *this friend of* ***mine***, or *the name of* ***it***.

In the following, the term *possessive* is used interchangeably with *genitive*; it includes, but is not restricted to, relations of ownership and possession.

The different head–modifier relations are divided into the following subcategories:

possession	alienable: *his house, the dog of him*; inalienable: *his fingers*
relationship	*his wife*
partitive	substance: *head of lettuce*; element: *group of kids*; source: *portion of rice*
referent property	*their size*, *the size of them*
state	*the beginning/ end of it*
participation	subjective: *the boy's arrival*; objective: *his murderer*, *the release of the boy*
origin	*the people of London*, *Beethoven's music*
reference	*of the Gaulish War*
description	*maid of honour*
classification/ measure	*two week's notice*, *a dollar's worth*
purpose	*women's shoes*, *children's games*
appositive	*the great city of Berlin*
double genitive	*a picture of the Queen's*
adverbial genitive	*of a summer's day*

Synthetic and analytic genitives have been in use since Old English. Historically, the *of*-genitive is the younger and initially less common structural variant. It became increasingly frequent over time, especially with inanimate possessors (Rosenbach 2002: 177 and 233). From around 1400 onwards, however, preferences began to change. The *s*-genitive started to spread on the expense of its analytic counterpart, a development probably linked to the development of possessive *'s* from an inflectional ending to a clitic-like element.[151] The contextual distribution of the two structural variants appears to have developed roughly to what it is today by the 16th or 17th century (see the distribution of *its* vs. *of it* analysed below). Different studies have tried to capture the exact nature of this distribution, including factors like economy (cf. Rosenbach 2002: 235) and different registers (cf. Rissanen 1999: 201). It has been argued that the revival of the *s*-genitive was largely determined by three conceptual factors, along the hierarchy of Animacy > Topicality > Possessive Relation. In Modern English, the relative importance of the same three factors has spread to the inanimate domain (cf. Rosenbach 2002: 235).

151 Compare Rosenbach (2002: 231) for parallel developments in other Germanic languages.

Focusing on the distribution of pronominal modifiers, the present study is not directly comparable to other studies in the field. However, previous studies on nominal modifiers are a point of reference and comparison. They include manifold approaches to genitive variation in English, examining potentially influential factors as diverse as information structure, phonology, syntactic complexity, semantic properties of the modifier and head, and the modifier–head relation. Szmrecsanyi (2006: 87–107) has presented the first multivariate analysis of the most important determinants, based on FRED and the Corpus of Spoken American English. Among Szmrecsanyi's predictors for genitive choice, the ones most relevant for a study on pronominal modifiers are animacy ("the more animate the possessor, the more likely the *s*-genitive", p. 205) and persistence, an additional important concept which not only affects genitive choice but linguistic behaviour in general. Szmrecsanyi found that "switch rates between the two genitives are exceedingly low, thus genitives are 'sticky' variables." (p. 107). Regarding genitives with nominal modifiers in the FRED corpus, he found that "usage of an *s*-genitive reduces the odds that an *of*-genitive will be used next time by 98% [...] This effect, however, weakens as textual distance between two genitive sites increases. The forgetting function that describes this relationship is logarithmic." (ibid).

Some factors which appear to have a negative influence on the choice of *s*-genitives in nominal occurrences need not be considered in this study. These include phonological exclusion criteria as that of nominal modifiers ending in /s/, /z/, /Θ/, and the issue of syntactic complexity, or end weight (cf. Rosenbach 2003: 384; Quirk et al. 1985: 323), since the modifiers and heads investigated in this study are generally mono-lexical.

Out of the many papers on genitive variation, three studies will be outlined which represent three different approaches: Rissanen (1999), Anttila and Fong (2004), and Stefanowitsch (2003). Their respective claims will be tested on the FRED data.

Previous studies have given special attention to semantic properties of the head noun as a major determinant of genitive choice, as well as the relationship between head and modifier. Rissanen (1999), for example, has observed that *s*-genitives are often preferred with human names (*John's number*), in cases where the modifier stands in a subjective relation to the head (*the boy's arrival*), and in certain quantifying expressions (*an hour's talk*). *Of*-genitives, on the other hand, seem to be favoured with inanimate nouns (*the colour of your shirt*) or where the modifier stands in an objective relation to the head (*the release of the boy*). Rissanen's claims regarding the modifier–head relation will be tested in 13.3.

Anttila and Fong (2004) have studied genitive variation from an Optimality Theory perspective. Similar to Rissanen, they have observed that *s*-genitives lend themselves to a subjective reading whereas *of*-genitives more likely receive an objective reading. An expression like *his thought* is most likely interpreted as 'he had a thought', whereas *the thought of him* is probably read as 'somebody thought of him'. At the same time, the authors found that *s*-genitives often have potentially ambiguous readings. A phrase like *his picture*, for example, can receive a lexical reading (a picture depicting the referent of *his*), or an extrinsic reading (a picture which belongs to, or was painted by, the referent of *his*). The analytic variant, on the other hand, is bound to receive a lexical reading: *a picture of him* will probably be understood as a picture depicting the referent of *him*. According to Anttila and Fong, variation between the two genitives is not entirely free but serves to express more or less subtle differences such as subjective vs. objective reading, new vs. old information, and so forth. In their study, the authors found that "corpus examples demonstrating free variation between the '*s*-genitive and the *of*-genitive are not easy to find. This is because the two variants will almost never occur in exactly the same linguistic context" (2004: 1254). I intend to show that Anttila's and Fong's observations tally, in general, with the distribution found in FRED, but they are less categorical than previously assumed. *Of*-genitives occasionally receive a subjective reading, and although free variation between the two genitive structures is rare, it is present in the data. Consider this first example from FRED:

(509) See, and then we used to do **their backs** when they were in the, in the cowshed, before they (gap 'indistinct'). As the ewe did go out, we did do **the backs of 'em** and then let them out, see. (free variation between *s*- and *of*-genitive with same referent in same paragraph, SRLM_PT, SOM_007)

Stefanowitsch (2003) has presented a constructional semantics approach to genitive variation which demonstrates that the distribution of the two structural variants is not (primarily) linked to information structure, but rather to specific semantic roles (p. 418–419, 438; supported by Hinrichs and Szmrecsanyi 2007: 460). Stefanowitsch describes the prototypical meaning of *s*-genitives as that of possession or ownership, where the roles of POSSESSOR and POSSESSUM are assigned to the modifier and head of the construction: {*John's*–POSSESSOR *dog*–POSSESSUM}. *Of*-genitives, on the other hand, prototypically encode part–whole relations, where the modifier and head are best described as ENTITY and INTRINSIC ENTITY: {*the roof*–ENTITY *of that house*–INTRINSIC ENTITY}. Despite these preferences, there is some overlap between the two constructions. Stefanowitsch gives

several examples for variation in identical contexts: *the desert's beauty* and *the beauty of the desert* both encode the same ATTRIBUTE–HOLDER relation; *the baby's eyes* and *the eyes of the baby* both encode the same COMPONENT–WHOLE relation; *the University's budget* and *the budget of the University* both encode the same POSSESSOR–POSSESSUM relation (cf. 2003: 421).

Of the three studies outlined so far, Stefanowitsch's analysis comes closest to the results obtained in the current investigation. Nevertheless, I aim to show that, while the identified semantic relations are useful for assessing the determinants of genitive variation in general, there are exceptions to some of his proposed constraints. Especially interpersonal relations, which have been argued to be encoded by *s*-genitives only, are also encoded analytically in the FRED interviews.

13.2 Quantitative distribution in the corpus, with focus on 3SGn cases

Tables 13.1 and 13.2 show a breakdown of *s*- and *of*-genitives by pronominal modifier category as found in the FRED corpus. For the sake of completeness, Table 13.2 also includes possessive pronouns, as in ***mine*** *was a good house in a bad area* (LAN_012) or *she's paid for* ***hers*** (OXF_001).

The numbers include all instances of potential interchangeability, i.e. all cases where the speaker could have theoretically chosen between the two genitive structures. For this purpose, the following occurrences were excluded by manual post-editing:

1. *of*-genitives of the 'partitive' group where the pronominal modifier is categorically found in complement position; for example, *one of mine* but **my one*;
2. *of*-genitives, mainly of the 'relations' group, with an indefinite possessum and no semantically equivalent *s*-genitive, especially where the possessum is pre-modified by a demonstrative or adjective; for example, *I 'm still a very great friend of him* (MDX_001), *they were quite big customers of ours* (WIL_015), or *they gave him six weeks, this pal of mine* (YKS_007); also excluded are occurrences with an indefinite possessum where the possessum is post-modified, for example by a local adjunct, as in *a friend of mine down in Bridgeford* (NTT_002), or a relative clause, as in *these men of ours that get around the military* (DEV_006). Indefinite possessa are included if they have a direct *s*-equivalent, as seen in (510). They are also included if they are one of a group and can therefore be paraphrased by *one of*, as in

she was with a friend of hers 'she was with one of her friends' (CON_009), *I was talking to a mate of mine* 'one of my mates' (LAN_012), or *I helped an uncle of mine* 'one of my uncles' (CON_011);

(510) There was three. There was miself, and mi sister who 's four years younger than me. And then there 's eh **a brother of mine**, he 's another four years younger than mi sister. (KEN_005; speaker has only one brother, consequently 'my brother')

3. *of*-genitives with 'the rest/ gist/ glory/ shock/ idea of + pronoun';
4. *s*-genitives with *own* where the modifier is categorically expected in determiner position: *her own gardening* but **the own gardening of her*, or *it's her own* but **it's the own of her*;
5. *s*-genitives with participle-gerunds of the 'participation' group, such as *my singing*;
6. s-genitives where the head is a personal name, as in *your Joyce is having a baby* (LAN_009), where the head is followed by a personal name, as in *his uncle Jim* (LAN_016), or where the modifier is part of a name, as in *they call it Our Lady's* (LAN_013);
7. double *s*-genitives of the type *my uncle's pub* (WIL_005) or *mi father's [team]* (WES_004), where an analytic paraphrasis would require a sequence of two *of*-genitives: *the pub of an uncle of mine*, *[the team] of the father of mine*;
8. local *s*-genitives of the type *these birds were left in the sacks* ***at my grandfathers*** *till dinner time* (WES_017);
9. *s*-genitives where the head describes the modifier referent's age, as in *she was well in her forties* (LAN_001).

Vernacular or colloquial features which were also excluded are:

10. qualified pronouns, as in *the delights of us little lads* (SAL_023);
11. colloquial *of*-genitives with no direct synthetic equivalent, as in *I went skipper of her* (SFK_010), or *there were a big family o' them* (SFK_010, not 'there was their big family' but rather 'they were a big family'/ 'their family was big');
12. occurrences with prepositions *to* and *with* instead of *of*, as in *this was a brother to him* (CON_006), or *the favourite one with us* (LAN_007), as well as any

ambiguous cases where *on* is possibly a pronunciation variant of *of*, as in *Harry Riches his name was what was skipper on 'er* (SFK_013);

13. impersonal *of it*, as in *that were the way of it in they times* (SOM_030);
14. and cases with *the top of* meaning 'on top of', as in *we had people living over the top of us* (LND_005).

Other restrictions which only apply to nominal modifiers play no role in this study. One of them is the constraint 'no non-specific specifiers' in *s*-genitives, as in *a ring of gold*/ **a gold's ring*, or *a state of shock*/ **a shock's state* (cf. Anttila and Fong 2004: 1265). Also, the corpus has no pronominal modifiers in independent genitives of the type *her memory is like that* ***of an elephant*** (cf. Szmrecsany 2006: 91–92).

Table 13.1: Genitives with pronominal modifiers: personal pronoun forms
(S: subject form, O: object form; no self-form cases in corpus)

structural variant	1SG		2SG/ PL	3SGm		3SGf		3SGn	1PL		3PL	
	S *I*	O *me*	S/ O *you*	S *he*	O *him*	S *she*	O *her*	S/ O *it*	S *we*	O *us*	S *they*	O *them*
s-genitive	---	///	6 (*thee*)	---	---	---	///	---	20	24	---	1 (*them's*)
of-genitive	1	1	4 (*you*)	---	8	---	12	86	---	4	1	32

Table 13.2: Genitives with pronominal modifiers: possessive determiner/ pronoun forms
(det.: pronominal determiner form; pr.: possessive pronoun form; poss. pron.: possessive pronouns as in *mine was a good house* or *she's paid for hers*)

structural variant	1SG		2SG/ PL		3SGm	3SGf		3SGn	1PL		3PL	
	det. *my/ mi*	pr. *mine/ mine's*	det. *your*	pr. *yours*	det./ pr. *his*	det. *her*	pr. *hers*	det./ pr. *its*	det. *our*	pr. *ours*	det. *their*	pr. *theirs*
s-genitive	5706	1	2262	---	2459	782	---	117	1149	2	1369	---
of-genitive	---	49	---	1	6	///	4	---	--	6	---	---
poss. pron.	---	83	---	22	26	---	4	---	2	---	---	19

The data show a clear preference for *s*-genitives in cases where it is theoretically possible to use both structures. Within a relatively small range of frequencies for the different pronouns, *of*-genitives are, on average, used in only 1.5% of all cases (also compare Table

13.3). This result corroborates findings of high *s*-genitive proportions in previous studies of spoken or informal text types (e.g., Altenberg 1982; Rosenbach 2002). It also supports a generalisation stated by Anttila and Fong (2004: 1260) that "[p]ronouns are preferred in the Specifier position and dispreferred in the Complement position."[152] According to this generalisation, speakers prefer *the removal of the tree* over *the tree's removal*, but *its removal* over *the removal of it*. This could explain why the *s*-genitive proportion among the pronominal modifiers in this study is even higher than among the nominal modifiers in the same corpus (cf. Szmrecsanyi 2006: 92; Szmrecsanyi and Hinrichs 2008: 296)

Table 13.3: Distribution of *s*- vs. *of*-genitives in FRED

genitives	*s*-genitive	*of*-genitive	% *of*-genitive	TOTAL
1SG	5,707	51	0.89	5,758
2SG/ PL	2,268	5	0.22	2,273
3SGm	2,459	14	0.57	2,473
3SGf	782	16	2.01	798
3SGn	**117**	**86**	**42.36**	**203**
1PL	1,195	10	0.83	1,205
3PL	1,370	33	2.35	1,403
TOTAL	13,898	215	1.52	14,113

One striking exception for which the above-mentioned generalisation is less applicable is the 3SGn. Although the general preference for *s*-genitives still holds as a tendency, the 3SGn has an unusually high proportion of over 42% *of*-genitives. The ratio between analytic *of it* vs. synthetic *its* is roughly 1 : 1.3.

In search for explanations, the first determinant which springs to mind is animacy. Rosenbach (2002: 233) has shown that animacy was already a decisive factor in the distribution of the two genitives in Old English, the *s*-genitive first being preferred for [+human], later [+animate] possessors. Szmrecsanyi's multivariate analysis (see above) has shown the continued importance of animacy as a predictor for *s*-genitives in 20th century English. Elsewhere, Stefanowitsch (2003: 438) has claimed that animacy has a strong influence on the choice between genitive constructions, following the general preference of English for placing items with animate reference before items with inanimate reference (especially in ATTRIBUTE–HOLDER relations). And Altenberg (1982: 146) has

152 Anttila and Fong's generalisation is based on ranked and violable constraints. According to the authors, the generalisation is supported cross-linguistically (cf. 2004: 1257). Their second generalisation "Non-pronouns show the reverse pattern" (ibid., p. 1260) could not be tested in this study.

mentioned the tendency to place items with inanimate reference after the head noun, and the tendency that possessors ranking higher on the Animacy Hierarchy are more likely to occur in *s*-genitives (inanimate > nonhuman animate > human > proper name > third person pronoun > first/ second person pronoun; cf. Croft 1990: 130).

Following this train of thought, the lack of animacy in 3SGn *it* would explain why *it* is more frequently found in complement position than pronouns typically associated with [+human] or [+animate] referents. The effect is visualised in Figure 13.1, which is an adaptation of the 'Preference structure for English *s*-genitive' by Rosenbach (2002: 168).

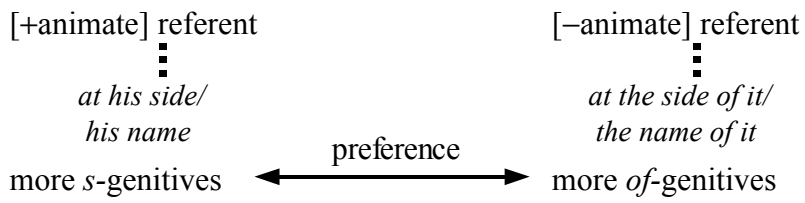


Figure 13.1: Animacy as a determinant of genitive choice

On the other hand, the corpus results rebut any more restrictive or categorical claims (cf. Biber et al. 1999: 303). In the corpus, 3SGn modifiers are placed both after *and* before animate heads: we get examples like *the foreman/ manager/ organisers of it*, but also *its chauffeur/ cubs/ lamb/* etc., as seen in (514)–(516). While it is true that most *of it* in the data indeed have inanimate reference (only 2 out of 86 have animate reference, e.g. in (513)), half of all *its* occurrences also have inanimate reference, devaluating categorical statements for either construction. In the whole dataset, inanimate reference in genitive modifiers is encoded by 57 *its* vs. 84 *of it*.

Note that inanimate referents of *its* are as varied as *shop*, *land*, *cottage*, *moustache*, etc.; animate referents are mostly animals, such as *horse*, *bird*, *crab*, *chicken*, *badger*, etc. ('REFERENT' column in spreadsheet). Collective referents like *the Church of England*, *school class*, etc. fall between categories; in the analysis, they were classed as inanimate. Five cases from the semantic categories of 'dead animal' and 'spirit' were excluded.

(511) And they waited until they got over, and the engine then, the weight of the engine put the bridge$_i$ down into **its**$_i$ proper position. (*its* with inanimate concrete referent, WS, KEN_006)

(512) The men used to indulge in what today is a filthy habit -- spit into this spittoon$_i$ -- that 's how it comes to get **its**$_i$ name... (*its* with inanimate concrete referent, IBME, SAL_023)

(513) And I got to Dale End and they had this sheep in a hull, and it had lambed. [...] Anyway, they let it out. Ding dong drive. It went off. But I had a lamb in mi arms you see. And Bragg [the dog], I set him off, And old Jonathan said, It 'll get down in there in t', among them houses, and you 'll never see it no more. Well, Bragg went right side **on it** [the sheep] and he turned it up... (*of it* with animal referent, WesDB, WES_015)

(514) And then as the time went on we got a pool, each member of the town had his own car and a chauffeur, the Mayor's car, the engineer's car, the health department car had **its chauffeur**... (*its* before animate head, WigWO, LAN_023)

(515) ...spring of the year, when old bitch fox is very bad to hunt, yes it is, when it has **its cubs**, it 's very bad to hunt. (*its* before animate head, WesAE, WES_002)

(516) ...and he carried it home and it had a lamb, and it -- we were surprised that it held **its lamb**, but it had done. (*its* before animate head, WesDB, WES_015)

A different important determinant in pronominal cases is topicality. As typically anaphoric expressions, pronouns are referentially given modifiers. They usually refer to entities which are already prominent in the discourse and, therefore, prominent in the interlocutors' minds. Pronominal modifiers hence represent highly accessible [+topical] expressions. In a right-branching language like English, pronominal modifiers are thus likely to undergo early serialisation (given > new information). On the phrase level, this mechanism leads to pronominal modifiers being realised as part of an *s*- rather than *of*-genitive (also cf. Osselton 1988; Szmrecsanyi 2006).

Interaction between the influence of topicality and animacy is easily imaginable, leading to an even distribution of the two genitives and, consequently, balanced 3SGn proportions as shown in Table 13.3 (58% *s*-, 42% *of*-genitives). The interaction between the two determinants is visualised in Figure 13.2.

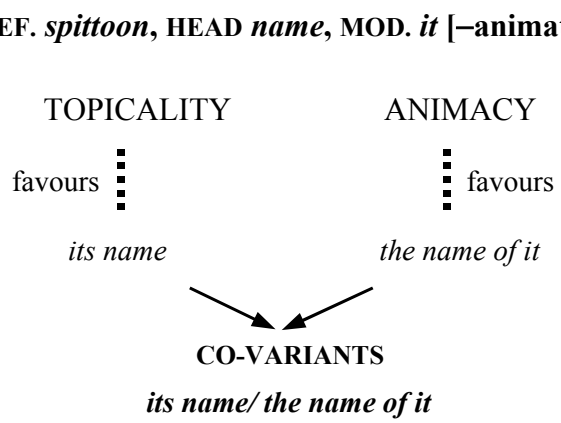


Figure 13.2: Interacting forces: animacy and topicality in 3SGn genitives

Besides determinants like animacy and topicality, which provide logical explanations for synchronic results, the distribution of English genitives also has an important historical dimension. The corpus findings for 3SGn *its* and *of it* are easily integrated into an ongoing linguistic change which has been described as the increasing use of *s*-genitives with inanimate modifiers. This development of the 20th century (cf. Denison 1998) has, for instance, been described by Rosenbach (2003) for genitives with nominal modifiers.[153]

Based on a questionnaire elicitation test with British and American English speakers, Rosenbach (2003: 380) found that "the s-possessive has systematically extended its range of application from Late MidE to present-day English both in terms of compatible contexts and frequency." Rosenbach's data confirm predictions derived from iconic and naturalness principles, the *s*-genitive being more frequently used with [+animate] and [+topical] modifiers in prototypical possessive relations.[154] Notwithstanding this dominant tendency, she observes the increasing use of *s*-genitives with [–animate] modifiers. In addition, Rosenbach provides evidence that the change is not lexically restricted to certain noun classes but productive "in an inanimate noun class hitherto assumed not to be participating in this change, i.e. concrete nouns." (p. 399). This observation, of course, contravenes the common assumption that the lexical class of the nominal modifier is one of, if not the most important, determinant of genitive choice, as maintained in the grammars of Quirk et al. (1985: 1277) and Biber et al. (1999: 302–303).

Rosenbach's findings for nominal modifiers are fully confirmed for pronominal modifiers in the present study. In the FRED corpus, the referential realm of inanimate *its* not only includes concrete referents, i.e. referents that can be physically experienced with the human senses, but it is practically confined to this category. About 90% of all genitives with inanimate *its* are concrete (51 out of 57, compare 'NOTES' column in spreadsheet). The referents are mainly buildings and vehicles, as shown in (511) and (514) above, but also machinery, crockery, a mole heap, and the spittoon in (512). The results can be interpreted in two ways: (a) the increasing use of *s*-genitives with inanimate modifiers, as visualised in Figure 13.3, is already well advanced in FRED, or (b) the animacy and noun

153 According to Rosenbach (2002: 189), "the increasing use of the s-genitive does not necessarily contradict the general drift towards more analyticity: the s-genitive becomes more productive as a clitic/ determiner, not as an inflection." Note that, in non-standard varieties of German, a drift has been observed towards analytic possessives of the type *dem Vater sein Haus* 'father his house' or *der Ruth ihr Kleid* 'Ruth her dress' (cf. Auer 2004: 74–85).

154 Rosenbach uses a speaker-based approach to grammatical variation, arguing that formal variation is determined by the need of speakers to place easily available information first in linear order (serialisation principle) and to encode more prototypical, inherent and therefore more predictable relations in the more bounded construction of the *s*-genitive (principle of conceptual distance). The iconicity of constituent order in English possessives has also been discussed by Taylor (1996: 18).

class restrictions proposed for nominal modifiers do not apply to pronominal occurrences. At this stage, both interpretations are possible; further research is needed to clarify this point.

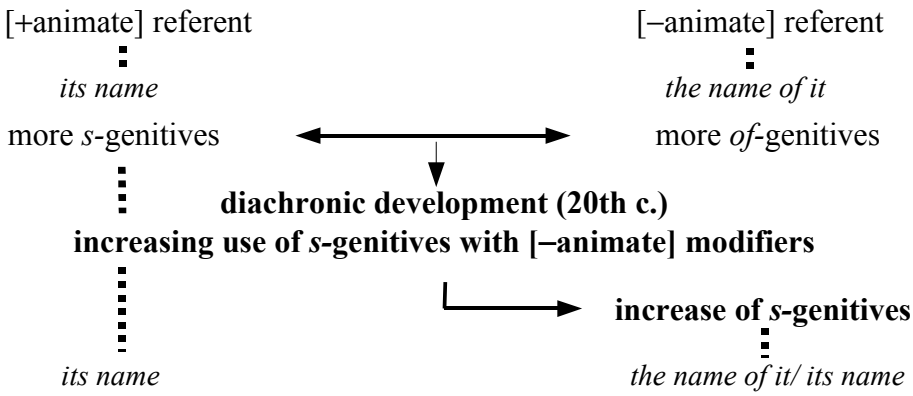


Figure 13.3: The increasing use of *s*-genitives with inanimate modifiers

If we go back even further in the history of English, we find yet another possible explanation for the genitive distribution in the corpus data. I propose the following hypothesis:

(517) The higher frequency of *of*-genitives with 3SGn *it*, as compared to other pronominal modifiers, can be traced back to the absence of a synthetic neuter variant *its* in 16th century English.

A brief outline of the history of *its* illustrates the difference in the development of English genitives in the 3SGn, as compared to other pronouns. According to Nevalainen and Raumolin-Brunberg (1994), the shift from grammatical to notional gender in English, together with the reinterpretation of possessive determiner *his* as a masculine-only form (previously masculine and neuter), entailed a severe reduction of *his* with inanimate or non-human referents. In view of the absence of a neuter pronoun which could take the place of *his* in determiner position, the possessive paradigm in 16th century English was left with a systemic gap. As long as this gap existed, analytic *of it* was the only variant with exclusively neuter reference. It was often preferred over pre-nominal *his* and post-nominal variants like *thereof* and *of the same* in oral genres and colloquial language (ibid., p. 197).

In addition, *of it* was "an analytic form at a time when a typological change from synthetic to analytic [had] been continuing for several centuries"; as a post-nominal alternative, *of it* was, therefore, "well suited for inanimate reference" (ibid., p. 198). Nevalainen and Raumolin-Brunberg find that "during the first half of the sixteenth century, *his* was still chosen in 26% of the cases, and in the rest of the cases, posthead alternatives were employed." (ibid., p. 190). The gap in the possessive paradigm was filled when *its* was introduced, probably spreading from the South of England with remarkable celerity within the span of only one generation (cf. ibid, p. 200, for socio-historical connections). From 1660 onwards, *its* became the main neuter variant, and in 1672 John Dryden, one of the members of the Royal Society's committee for the improvement of the English language, already referred to neuter *his* as "ill syntax". According to Nevalainen and Raumolin-Brunberg, "the Present-day Standard English usage with *its* as the main variant and *of it* as its postnominal alternative had become standard by the end of the seventeenth century" (ibid, p. 200).

Notwithstanding the fast propagation of *its*, the time during which *of it* was the only exclusively neuter variant may have left traces. From a usage-based perspective, a historic link to present-day data like FRED could be seen in the continued above-average exposure of English speakers to *of it*, starting at a time when synthetic variants already existed for all other pronouns. Although we are talking about a prolonged period of time spanning many generations of speakers, it is easy to imagine how *of it* could have transmitted itself mimetically from one generation to the next, resulting in the distribution shown in Table 13.3.

In this regard, it is interesting to note that the distribution of 3SGn modifiers in FRED (20th century) is surprisingly similar to their distribution in 17th century English, shortly after the introduction of *its* into the English paradigm. Nevalainen and Raumolin-Brunberg (1994: 190) found the following genitive distribution in the diachronic part of the Helsinki Corpus: 62% pre-head variants (*its*/ *it*) vs. 38% post-head variants (28% *of it*, 7% *thereof*, 3% *of the same*).[155] This compares to roughly 58% *s*-genitives and 42% *of*-genitives in FRED, suggesting that neither time, nor register, nor the linguistic medium (spoken/ written) have so far exerted a strong influence on genitive choice in the 3SGn. Even the two major Modern English developments mentioned above – the general revival of the *s*-genitive and, more recently, the increasing use of *s*-genitives with inanimate modifiers – seem to have left the distribution of *its* vs. *of it* unaffected.

155 *Thereof* and *of the same* were longest retained in the language of law. They do not occur in FRED.

13.3 Modifier–head relations

We now take a look at a different determinant of genitive variation, the modifier–head relation. Several constraints on this relation have been formulated in the literature for nominal modifiers. The present study offers the opportunity to check these constraints for pronominal occurrences. The variation between the two genitive structures in the corpus will also be used to reassess whether certain constraints are absolute or variable. Two main aspects to be tested concern the subjectiveness of the modifier–head relation and the relationality of the head noun.

13.3.1 Subjective vs. objective readings

Regarding the first aspect, it has been claimed that there is a general preference among speakers of English to encode subjective relations by the *s*-genitive, and objective relations by the *of*-genitive (compare the approaches by Rissanen, and Anttila and Fong, described in 13.1). This preference is exemplified by the two phrases *the boy's arrival* (subjective: the boy arrived), versus *the release of the boy* (objective: somebody released the boy; cf. Rissanen 1999). In order to test the claim, all cases were compared for which a subjective or objective relation could be ascertained. The results are shown in Table 13.4 (spreadsheet subjectiveness.xls).

Table 13.4: Genitive structures and modifier–head relations (subjectiveness) (ambiguous cases excluded, e.g. with head noun *training*; modifier-head relations assigned by manual post-editing; spreadsheet: 'MODIFIER–HEAD RELATION' column)

modifier–head relation	*s*-genitives	%	*of*-genitives	%
subjective	100	95.2	5	4.8
objective	23	79.3	6	20.7

There is clear evidence against an absolute constraint. Furthermore, the results show that the *s*-genitive is the preferred structural variant for both modifier–head relations. Subjective relations are usually encoded by the *s*-genitive as expected, with a few exceptions,

(518) ...these little impurities gets on during the course of the plate, or whatever it is, cup or saucer, whatever it is -- during the course of **its progress** round the works... (subjective relation encoded by *s*-genitive, IBAL, SAL_030)

(519) ...well he has got a bigger farm than me, it was **his suggestion**, he has got a bigger farm than me, than the two of us combined. (subjective relation encoded by *s*-genitive, SRLM_CH, SOM_008)

(520) She was sitting by the window, and, uh, by all accounts (v 'interviewer blows nose') this landmine, the, the, uh, **blast of it** blew her out the window. (subjective relation encoded by *of*-genitive, LA, LND_005)

and objective relations are also encoded by *s*-genitives, with slightly more exceptions:

(521) ...but he helped other people with all sorts of -- he helped other people to **our detriment**. (objective relation encoded by *s*-genitive, YksML, YKS_007)

(522) And I applied for **my release** two or three times, and I was put in front of a tribunal at Barrow, but they wouldn't let me go. (objective relation encoded by *s*-genitive, WesDX, WES_017)

(523) I mean there 'd been very little common sense gone into the actual **planning of it** and the **cropping of it**, it was in a terrible state to say the least... (objective relations encoded by *of*-genitives, WesCB, WES_010)

While the numbers show that objective relations receive an analytic encoding more frequently than subjective relations, it would be wrong to translate the results into a preference statement for the analytic encoding of objective relations. The encoding of objective relations is still predominantly synthetic (79.3%).

At the same time, the comparable number of subjective and objective relations among the different *of*-genitives in the corpus (5 and 6 cases respectively) requires a qualification of Anttila's and Fong's statement that the *of*-variant more likely receives an objective reading. While their own example *his thought* vs. *the thought of him* leaves little room for discussion (see 13.1), there are other head nouns for which semantic differences need to be considered first. In FRED, for example, two occurrences with the head noun *memory* show readings which are diametrically opposed to the *thought* example: in (524), the *s*-genitive encodes an objective reading (people remembered by the speaker), whereas the *of*-genitive in (525) encodes a subjective reading (the speaker remembers the house).

(524) The various old old friends that -- and old people of Pinner, they 've still got one or two of their sons and their daughters left, that I still remember, and still know, and cherish **their memory**, of course. (head noun *memory*: objective modifier–head relation, MdxCG, MDX_001i.wav)

(525) Another **memory of mine** were Wilton House, Lord Pembroke's. (head noun *memory*: subjective modifier–head relation, TrbrCS, WIL_005b.wav)

Based on the corpus results, the preference structure for the encoding of modifier–head relations can be summarised as follows:

(526) Encoding of modifier–head relations in genitives with pronominal modifiers
Subjective and objective readings are both predominantly encoded by *s*-genitives, and, less frequently, by *of*-genitives. At the same time, objective relations are more frequently encoded by *of*-genitives than subjective relations.

13.3.2 Relationality of the head noun

The second aspect to be tested is the relationality of the head noun. It has been claimed in the literature that *of*-genitives do not usually appear with non-relational nouns, i.e. referents that may or may not belong to somebody (cf. Stefanowitsch 2003: 428). The claim has been statistically verified for pronominal modifiers. Anttila and Fong (2004: 1270), for example, found no occurrences of *of*-genitives with 'non-relational head + pronominal modifier' in their data (BROWN corpus), as opposed to 611 *s*-genitives of this type. They consequently concluded that non-relational nouns with pronominal complements, such as *the hospital of us* or *the music of him*, are "generally no good" (ibid, p. 1266).

The structural variability of genitives with relational head nouns, on the contrary, is not usually put into question. Stefanowitsch (2003: 24), for example, finds that non-relational heads as in *Kate's shoes* can not precede the modifier (**the shoes of Kate*), but relational heads as in *the University's budget* or *John's friend* can (*the budget of the University*, *the friend of John*). Similarly, Anttila and Fong (2004: 1267) state that a "crucial fact about relational nouns is that they allow variation: the of-genitive can be generally paraphrased by the 's-genitive." Their findings support the so-called Barker-Dowty Generalization: "If a noun can take a genitive of-phrase and if the of-phrase can also be paraphrased by a prenominal possessive then we can generally assume that the noun has the appropriate relational sense" (cf. Barker and Dowty 1993: 51).

The observed tendency is largely supported by the FRED data (see 'HEAD NOUN' column in spreadsheets). However, it is important to note that there are exceptions and that the Barker-Dowty Generalization is violable. Examples (527)–(529), for instance, show three *of*-genitives with non-relational head nouns where the pronominal complement could

just as well be a pre-nominal possessive (*its chimney*, *its door*, *their gates*). Since these examples all encode COMPONENT–WHOLE relations, they do not affect Stefanowitsch's statement that, in the case of non-relational nouns, "a relation of ownership is only encoded by the *s*-genitive" (2003: 428). But this claim, too, is not categorically true. Examples (530) and (531) show two *of*-genitives with non-relational head nouns encoding an ownership relation (note that in (530) the interviewer repeats the speaker's *of*-genitive by switching to the *s*-variant). The fact that the only counter-examples in FRED to contradict Stefanowitsch's claim have nominal modifiers tallies with Anttila's and Fong's results from the BROWN corpus (cf. 2004: 1270): they found no *of*-genitives with 'non-relational head + pronominal modifier', but 63% of all 'non-relational head + nominal modifier' combinations were encoded by *of*-genitives.

(527) ...they used to packet powder, cocoa, coffee, flour, everything you could mention they they packeted there and it 'd been stood for years and **the chimney of it** was smack in the street... (*of*-genitive with non-relational head noun, WigWHJ, LAN_020)

(528) I can't tell you much about -- I 've never been in the army. I 've only been past **the door of it**. {<u Int> This is the building in Ironbridge where they did their training?} (*of*-genitive with non-relational head noun, IBGJ, SAL_032)

(529) Well, they was the powder mills, and on the gates -- I can remember -- on **the gates of them**, it was John Hall and Son. (*of*-genitive with non-relational head noun, HM, KEN_003)

(530) I went there -- **a little private school of Miss Owen**. [...] {<u Int> Coming back to **Miss Owen's private school**...} (*of*-genitive with non-relational head encoding an ownership relation, IBCE, SAL_026)

(531) There weren't many yachts but **rowing boats of Mr Thwaites**, yes. (*of*-genitive with non-relational head encoding an ownership relation, WesBE, WES_005)

13.4 Collocational preferences

Apart from the possible determinants discussed so far, it appears that analyticity can also be motivated collocationally, albeit to a somewhat smaller extent. This is indicated by the higher-than-average occurrence of certain head nouns in *of*-genitives in the corpus (a collocational analysis of non-standard pronouns in possessive function will be presented in ch. 14).

Collocational preferences are not usually considered in the literature on genitive variation in English; they have, however, been described for other languages. Nichols and Bickel (2005), for example, have studied lexically conditioned morphological oppositions in languages with two or more types of possessive marking (WALS: 'Possessive

Classification'). In the Australian language Maran, for instance, possessive prefixes are used with kinship terms, as opposed to free possessive pronouns with other head nouns. In the classification of Nichols and Bickel, English is a language without such oppositions in genitives. I would like to modify this classification and propose the following view: the difference between English and languages like Maran is that, in English, certain lexical items, or groups of semantically related items, can be observed to favour the use of *of*-genitives, but they do not condition the structural output categorically. Lexical motivation is, therefore, a more appropriate term than lexical conditioning.

Based on the corpus data, four groups of head nouns were identified which appear to favour the use of *of*-genitives:

- family and friendship terms, such as *friend* or *uncle*;
- spatial collocates indicating a specific location or position relative to the modifier referent, either in the referent itself (*on the bottom of them*) or its surroundings (*round the front of it*);
- *name*/ *names*, as in *the name of it*;
- so-called owner nouns, i.e. head nouns indicating ownership or leadership, such as *owner*, *captain*, *organiser*, and so forth.

In the following, these four groups will be briefly described. Since there is only a limited number of occurrences for each category in the corpus, further investigations could be useful in order to explore the exact nature of lexical motivation in English genitives.

Friends, mates, pals and uncles

In the data, the vast majority of *of*-genitives involving possessive pronoun modifiers describe a relationship relation (compare 13.1), as in *an aunt of mine* or *another friend of ours*. Together, lexemes like *friend*, *mate*, *pal*, *uncle*, *aunt*, *brother*, etc. account for over 86% of all heads in the analytic variant ('HEAD' column in the spreadsheets). The semantic field can be roughly split into two groups: friendship terms and kin terms. Genitives with head nouns from the friendship group are encoded analytically in a substantial number of cases, as seen in Table 13.5. This is in particular the case with *friend* and *friends*, for which we get almost 33% *of*-genitives (compare examples (532)–(534)). Among the different kin

terms, the most considerable proportion of analytic cases was found for *uncle* and *uncles*, also listed in Table 13.5 (compare examples (535) and (536)). Other kin terms are mainly encoded synthetically. Anttila's and Fong's finding that *of*-genitives are not used with personal pronoun modifiers and head nouns of the 'brother' class (*my brother*, but **the brother of me*; 2004: 1275–1279) seems generally true. Comparable family terms as shown in examples (377) (ch. 11.1) and (573) (ch. 14.2) are extremely rare in the corpus.

Table 13.5: Genitives with head nouns of the 'friends and family' category
(only occurrences with interchangeable *s*- and *of*-genitives)

head noun collocates 'friends and family'	*s*-genitives	*of*-genitives	% *of*-genitives
'friends' group *friend/ (school)mate/ (school)pal/ chum*(*s*)	137 (47/75/12/3)	32 (23/4/4/1)	18.9%
uncle/ uncles	100	16	16.0%

(532) **A friend of ours** used to look after that it used to be beautiful. (*of*-genitive with 'friend' head, TCA_K, DEV_008)

(533) Oh, it was **an old mate of mine**, he worked there and spoke for me. (*of*-genitive with 'friend' head, DurNB, DUR_002)

(534) I remember the day well. I met her -- she was with **a friend of hers**, and her friend was going with **a pal of mine**... (*of*-genitives with 'friend' heads, ICS_EW, CON_009)

(535) Somebody had to go on a bicycle. **One of mi uncles** likely. (*s*-genitive with *uncle*, WesDB, WES_015)

(536) Now **an uncle of mine** had a turnover plough, left and right, you know... (*of*-genitive with *uncle*, CorJLN, CON_011)

A different aspect, which exceeds the scope of this study but should nevertheless be mentioned, is the possibility of encoding interpersonal relations by an *of*-genitive where the modifier is a personal name, as seen in (537). Comparable examples have elsewhere been excluded from the semantic relations encodable by *of*-genitives (e.g. Stefanowitsch 2003: 421, Table 1 **the girl of Kate*).

(537) Every time I went down Little Dawley, **the eldest girl of Lees** -- Hilda -- she used to shout after me... (IBFB, SAL_013)

Spatial collocates

Almost half of all *of*-genitives with pronominal modifiers in FRED have head nouns from the semantic field of spatial expressions. These include *side*/ *sides* (28), *top* (10), *middle*/ *middle part*/ *heart* (7), *front* (4), *back*/ *bottom* (3 each), *the inside*/ *the outside* (2 each), and *centre*/ *centre piece*/ *end*/ *surround* (1 each) (excluding *on the top of* meaning 'on top of', and body parts like *back*). A comparison with synthetic equivalents in the same dataset (20 cases) shows that spatial collocates clearly favour analytic encoding. It can even be observed that the use of a spatial collocate leads speakers to switch from one grammatical structure to the other, as seen in (541).

Especially striking are collocations with *side*. Out of the 39 occurrences with *side* in the corpus, 28 are *of*-genitives and only 11 are *s*-genitives. The FRED results deviate from findings in similar studies. Anttila and Fong (2004: 1269–1270), for example, found that relational expressions, including *side*, *middle*, etc., were usually encoded by the *s*-genitive: 97% *s*- vs. only 3% *of*-genitives.[156] A possible explanation for the difference between their results and the data at hand is that many of their examples were taken from the BROWN corpus, which consists of written standard American English texts. Accordingly, the use of *of*-genitives with spatial collocates could be a spoken feature, or an exclusive dialect feature of British English. More detailed cross-varietal studies are needed to clarify this point.

(538) ...and if you put the first one up then somebody 'd put some the **right hand side of you** and some would put some the, **the left hand side of you** so that they 'd be evenly balanced... (*of*-genitives with spatial collocates, SRLM_WB, SOM_011)

(539) I used to walk at **the side of him**, now and again. (*of*-genitive with spatial collocate, CA, LND_001)

(540) She used to have a little jug **at the side of her** while she was working all the while, she couldn't carry on without that. (*of*-genitive with spatial collocate, NotA85, NTT_010)

(541) ...on wet days he 'd be having bags all round **his shoulders** and round **the front of him**... (*of*-genitive with spatial collocate, IBJD, SAL_019)

(542) ...we used to have a box down **at the side of us** there, and, and you used to put, uh, pieces of cloth in... (*of*-genitive with spatial collocate, YksMR, YKS_011)

(543) Oh I was thankful because I was thinking about the passenger coming over the junction, see, that eight o'clock at night. I thought we shall go straight through **the middle of him**. (*of*-genitive with spatial collocate, IBErDi, SAL_020)

156 Anttila and Fong (2004: 1270–1271) admit to a shortcoming in their descriptive taxonomy, which treats relational nouns as a homogeneous group ignoring systematic differences between subgroups and individual lexical items. At the same time, the authors maintain that "[w]hile individual words do have a certain degree of grammatical autonomy in choosing their linking patterns [...] the possible choices are severely limited by the grammar" (ibid, p. 1280).

'the name of it'

Particularly noticeable in *of*-genitives are collocations with *name* and 3SGn *it*. In the data, we find 12 *name of it*, as compared to 16 *its name* (other pronouns prefer the *s*-genitive). The numbers in Table 13.6 include all *name*-related head nouns, including *Christian name*, *nickname*, *real name*, and so forth.

Table 13.6: Grammatical variation in genitives with head noun *name*

collocate *name*	1SG	2SG/ PL	3SGm	3SGf	**3SGn**	1PL	3PL	TOTAL
s-genitives	28	32	197	54	**16**	4	42	373
of-genitives	0	0	1	0	**12**	0	4	17
% *of*-genitives	0.0	0.0	0.5	0.0	**42.9**	0.0	8.7	4.4

Contrary to the above-mentioned assumption by Anttila and Fong that the two genitives will almost never occur in the same linguistic context, several synthetic and analytic cases with *name* are directly comparable, as shown in the following pairs of examples from the corpus:

(544) There was also another path that passed over the top of which is now (pause) hmm, **I forget its name**... (genitive *its name*, MdxCG, MDX_001e.wav)

(545) She was working, she was working in Brewery Road, the uh old rag place used to be up there, **I forget th' name of it now**... (genitive *the name of it*, JG, LND_004b.wav)

(546) ...the old people used to say that it kept the, it away (gap 'indistinct') **what 's its name**? (genitive *its name*, SRLM_B, SOM_003)

(547) Like Sundays, an' that, I forget what they call it, **what 's the name of it**? (genitive *the name of it*, SRLM_RF, SOM_005)

(548) I forget **what his name was now**, as used to come. (genitive *his name*, IBWB, SAL_004)

(549) I don't know **what the name of him was now**, a good dog he was, you know. (genitive *the name of him*, CAVA_WW, CON_006)

Owner nouns

The last identifiable group of collocates associated with analytic encoding can be summarised under the heading of owner nouns. These head nouns have referents who own, command or have the power to decide over the modifier referent in some way (not including relatives and teachers). Unfortunately, the number of cases in FRED is small, and further tests are needed to support this observation. Nouns in this group are *owner* (5),

skipper (4), *captain* (2), and *foreman*/ *organisers*/ *originator*/ *manager* (1 each), as shown in the examples below. Most of these nouns also occur as collocates in *s*-genitives (24 *boss*, 11 *foreman*, 10 *master*, 7 *skipper*, 6 *captain*, 2 *owner*).

Particularly noticeable is the use of *skipper of* and *captain of* with 3SGf modifiers referring to ships or vessels. A look beyond the FRED data shows that *skipper of her*, in particular, is surprisingly common in English: a simple Google search in July 2010, for example, rendered 2,100 hits worldwide for *skipper of her* as compared to 19,500 *her skipper*. In the data, all cases with 3SGf modifiers and heads from the owner noun group are analytic (4 *skipper*/ 2 *captain*/ 1 *owner of her*, not including 2 *skipper on her*).

(550) ...the captain [who] did have her, or the **owner of her**, was a kind of doctor. (*of*-genitive with owner noun, SRLM_WT, SOM_028)

(551) He had the Irene, that 's the last one launched. See the **captain of her** was half owner and Carlow Reynolds had the other half. (*of*-genitive with owner noun, SRLM_WT, SOM_028)

(552) ...and we weren't there long before they say t' me, All right, You 'd better go aboard the Pevensey Castle! That was one o' Jack Breach's what 'd been taken over, Paddy Mewse was **skipper of her**. (*of*-genitive with owner noun, EJM, SFK_006)

(553) ...old Jack Hodge he used to be the skipper of the Pioneer, the tug boat, what used to tow barges to and fro, paddle boat, tug boat, old John Hodge. He used to be the **skipper of her** and he had this little shop... (*of*-genitive with owner noun, HM, KEN_003)

13.5 Summary

The distribution of *s*- and *of*-genitives in FRED shows that genitive variation is a gradient phenomenon. The variation between the two grammatical structures is only to a certain extent explained by constraints on the head noun, the pronominal modifier and the modifier–head relation. While the present study is not directly comparable to previous studies on genitives with nominal modifiers, several tendencies and constraints described in the literature could be tested successfully.

In accordance with previous studies (including more formal registers and written English texts), the data show a strong overall predominance of *s*-genitives in utterances where both structural variants are possible (98.5% on average). Special attention was given to 3SGn cases for their exceptionally high proportion of *of*-genitives (over 42%), and different explanations were suggested for this form-specific behaviour. Absence of animacy was considered as a main predictor for the high amount of analytic cases with *of its*, as compared to other pronouns with naturally animate referents. The results indicate

that animacy on its own is not decisive. It may, however, interact with other factors, including the opposing force of topicality. Unlike animacy, topicality is likely to favour the realisation of discourse-prominent 3SGn modifiers as part of *s*- rather than *of*-genitives. Similarly, a relatively recent development which also favours the use of *its* over *of it*, is the increasing use of *s*-genitives with inanimate modifiers, including concrete referents. It is easily conceivable how interaction between the different opposing forces – some favouring an analytic encoding, others favouring a synthetic encoding – could lead to an even distribution of the two variants in the 3SGn as observed in the FRED data (Table 13.3; Figures 13.1, 13.2, 13.3). For all other pronouns, the combined influence of [+animacy] and [+topicality] features explains the strong predominance of *s*-genitives.

Additional explanations for the special distribution of genitives with 3SGn modifiers are found in the historical development of the English possessive paradigm. The common use of analytic *of it* could, even nowadays, be historically linked to a systemic gap in 16th century English, caused by the absence of a synthetic neuter modifier (later *its*). In the future, it will certainly be interesting to see whether the distributional difference between 3SGn modifiers and other pronominal modifiers continuous to exist, and in how far the traditional dialect data investigated in this study tally with the speech of younger generations.

Regarding the modifier–head relation, evidence was provided against some absolute constraints from the literature. It was shown that the subjectiveness or objectiveness of the relation is not a reliable predictor for genitive choice. Subjective and objective readings are both encoded by both structural variants. However, while the two readings are predominantly encoded by *s*-genitives (in accordance with the strong overall predominance of *s*-genitives), objective readings are more often encoded analytically than subjective readings. With respect to the relationality of the head noun, the corpus results largely support predictions that non-relational nouns are encoded synthetically. At the same time, there exist counter-examples which show that structural variation can not be categorically excluded for non-relational heads.

Last but not least, it was found that analyticity in genitives can be lexically motivated. Even though collocate head nouns in English do not have the same influence as in other languages, they can motivate speakers to use the *of*-variant. Among the collocates identified above are spatial expressions, especially *side*, friendship terms like *friend*, *mate* and *pal*, and so-called owner nouns, especially *skipper*. Interesting results were also obtained for the collocate *name*: while speakers usually prefer the *s*-variant with this head

noun, as in *my*/ *her*/ *his*/ *their name*, the two genitives vary freely in the 3SGn, where *its name* and *the name of it* are very much in balance. The results for *name* tie in nicely with the general distributional tendency in 3SGn cases.

Chapter 14
Possessive function

In recent decades, the use of personal pronouns in possessive determiner function has been noticed in various descriptions of non-standard varieties in and outside the British Isles. Wales (1996: 167), for example, mentions the use of possessive *us* in Yorkshire and the West Midlands, possessive *thee* in Somerset, and possessive *mi*, *im* and *dem* in AAVE. Wakelin (1986: 34) also mentions the use of possessive *thee*, in Bristol. According to Petyt (1985: 190), possessive *us* in West Yorkshire "not uncommonly occurs for the unstressed *our* e.g. 'we all take us cars to work nowadays' ," but his "experience in school is that this is a highly castigated feature." One of the most frequently commented features in British English dialects is probably the use of possessive /mi/, which in some studies is regarded as an irregular use of pronoun *me*, in others, including the present study, as a pronunciation variant of *my* (cf. Wales 1996; Anderwald 2003).

From a wider perspective, the use of identical pronoun forms for subject, object and possessive functions has been described for different English creoles, for example basilectal 1SG *mi*, 1PL *wi*, 3SG *im* and 3PL *dem* in Jamaican Creole (cf. Sebba 2004: 203; Patrick 2004: 429). Other varieties use oblique forms in object and possessive function, but a distinct form for expressing subject function (cf. Winford and Migge 2004: 508–509, Surinamese creoles). S-form possessives, on the other hand, seem to be relatively rare; possessive *they*, for example, can be found in AAVE (cf. Wolfram 2004: 333).

After analysing the distribution of genitive structures in the previous chapter, this additional chapter looks at the distribution of formal variants in genitive modifier slots. The same terminological distinctions apply as in chapter 13; occurrences were coded as POS 80 in the case of *s*-genitive determiners, and POS 81 in the case of *of*-genitive complements.

Non-standard pronominal modifiers are found in both *s*- and *of*-genitives in the corpus. However, the distribution differs considerably in the two syntactic structures, requiring a more detailed description. In *of*-genitives, the focus will be on the distribution of possessive pronoun forms, as in *an old friend of mine*, versus personal pronoun forms, as in *an old friend of me*. In *s*-genitives, special attention will be given to 1PL pronouns *we* and *us*, as in *we came home for we dinner*, or *us mum would fetch us*, as well as possessive *thee*, as in *hold thee noise*. In order to assess the frequency with which these non-standard

modifiers occur, they will be compared against standard *my*, *your*, *his*, etc. in the same data. Note that colloquial uses of the latter also occur in the corpus but cannot be discussed here. They include: the use of possessive determiners instead of articles and vice versa, as in *I had my bath* or *I said to the wife* (cf. Miller 2004); possessive determiners expressing habituality, as in *I'm off for my dinner*; and 'possessive determiner + personname' occurrences, as in *our Tommy*.

14.1 *S*-genitives: pronouns in possessive determiner function

As a general tendency and in accordance with the grammar of Standard English, the prenominal determiner position in *s*-genitives is usually filled by possessive determiner forms.[157] Other variants are rare but possible: among the 14,367 *s*-genitives in the corpus are 58 personal pronoun forms, which roughly corresponds to 0.4%.[158] Most non-standard examples have O-forms (*him*, *thee*, *us*, *them*); the only S-form is *we*, as shown in examples (558)–(561).

Although personal pronoun forms only account for 0.4% of all *s*-genitive determiners in the corpus, they are used by a substantial number of speakers (SN = 19). A null hypothesis defined as the prenominal modifier slot being filled by no other pronominal forms than standard possessive determiners can be rejected at $p < .01$ (3.8 σ, $p = 10^{-4}$).

(554) {<u Int> How much did he charge?} Very little, more or less he used to make **his own bottles**, you know the doctor used to make **him own bottles** if you wanted medicine. (prenominal possessive *his* in alternation with *him*, WigWO, LAN_023)

14.1.1 Possessive *we* and *us*

Most striking is the recurrent use of three forms: 1PL *we* and *us*, and 2SG *thee* (see 14.1.2). *We* and *us* deserve special attention as the most frequent non-standard forms (50 occurrences). An important question, of course, is whether possessive *we* and *us* represent phonological variants of *our* (similar to /mi/ 'my'). In practice, this can be especially difficult to determine before words beginning with a vowel, where the speaker could be using either '*we* + linking -r' or *our*. Consider the following examples:

157 Pronunciation variants like /mi/ 'my' or /yə/ 'your' are typical of informal dialogue.

158 Excluding *her* (O-form and POSS-form); excluding unclear cases such as *the last one that 's in en name do go on* (CON_005). Including 819 cases with *own*, as in *we 'd all got us own toilets* (NTT_014), or *unless we took we own stuff* (LAN_005); including 1 *mine* and 2 *ours*; spreadsheet s_possessives.xls.

(555) And we had to make **[/wər/]** own amusements. (LanD1P, LAN_006c.wav)

(556) And then Christine would be about seven, and we started going on **[/wər/]** holidays. (LanC5P, LAN_005c.wav)

(557) ...the night before we 'd tie **[/wə/]** rods on **[/wə/]** bikes, and we 'd come home rush **[/wə/]** dinner down, and get there as quick as we could to catch that tide... (Lang1p, LAN_003a.wav)

In the literature, this particular issue remains unsolved. Similar occurrences are sometimes referred to as possessive *we* (cf. Singler 2004: 889, Liberian Settler English) and sometimes as pronunciation variants of *our* (cf. Beal 2004: 118, 'wor' in the Northeast of England). In the present study, it was decided to include only those occurrences in the analysis where a /wi/ pronunciation was recognisable in the recording (less than half of the occurrences). All other cases closer to /wər/ were excluded. Note that the numbers in Table 14.1 differ slightly from those in Tables 13.1–13.3 in the previous chapter because cases with *own* are included this time.

Table 14.1: Distribution of 1PL forms in prenominal determiner position

1PL prenominal determiners (incl. cases with *own*)	*our*	%	*we*	%	*us*	%	*ours*	%	TOTAL	%
	1,149	95.7	22	1.8	28	2.3	2	0.2	1,201	100

The first noticeable tendency in the distribution of possessive *we* is that it appears almost exclusively in data from the North (0 SE, 0 SW, 1 MID, 21 N; not correlated with the number of standard determiners in the different areas). In addition, possessive *we* appears to be attracted by collocates from the semantic field of 'food' or 'meals': 14 out of 22 occurrences have head nouns like *food*, *dinner*, *supper*, *tea*, *meals*, *jam and bread*. Both the geographic area and the collocational context can, therefore, be considered to represent potential determinants of possessive *we*, but further tests are required to quantify their influence.

(558) Same as a Sunday, we used to go t' church at morning, to go home and have **we dinner**. Sunday school at afternoon, and it finished at three o'clock. Or we used to go for a walk, maybe 'round Waverley Park [...] Then come home for **we tea,** and then **we church** again. (prenominal possessive *we*, LanC5P, LAN_005a.wav, trba.)

(559) And I, and I remember our Jimmy saying, Why can't we come home for **we dinner**, mother? (prenominal possessive *we*, LanC5P, LAN_005d.wav)

(560) We used to go to Throstleness Farm, take **we tea**, she used to brew it, put fresh cream in... (prenominal possessive *we*, LanD1P, LAN_006d.wav, trba.)

(561) ...I think we used to take **we food** with us and eat it there on the Trent side. (prenominal possessive *we*, NotA80, NTT_009)

Possessive *us* is well known in the dialectological literature (e.g., Trudgill 1999[2]: 89). It was, for example, recorded in the SED cloze question VIII.8.8 'We have stopped borrowing the neighbour's tractor because we now have (our/ us own)', rendering two *us* areas in the Midlands (Warwickshire and Staffordshire) and North of England (especially Yorkshire). The SED results are shown in Map D8 (Appendix D).

In the FRED recordings, the variant ending in voiceless -*s* sounds exactly like personal pronoun *us*. Nevertheless, some may argue that, here too, we are merely witnessing non-standard pronunciations of *our* – either by adding final -*s* to the possessive determiner, or as a result of a diachronic reduction along the lines of /aʊəz/ > /ʌz/ > /ʌs/. While the second development may seem unlikely at first, *ours* in possessive determiner function was actually found in two utterances in the corpus:

(562) Well, if **ours nightfighters** was up, they switched the searchlights out. (prenominal possessive *ours*, TCA_WH, DEV_009b.wav)

(563) Stayed a lot in **ours Lancaster**. (prenominal possessive *ours*, H3L, LAN_010b.wav)

(564) We, we used to have **us own** in the garden... (prenominal possessive *us*, Lei2, LEI_002a.wav)

(565) ... I used to pal up with him and we used to go in and have **us food** with his parents. (prenominal possessive *us*, YksHS, YKS_006)

(566) It were marvellous when they took the bandage off **us eye**, and you could see the windows. Oh it was marvellous. [...] the day as I 'd got to go out, we waited for **us mum** to fetch us. But she never come. (prenominal possessive *us*, ALI, NTT_012a.wav)

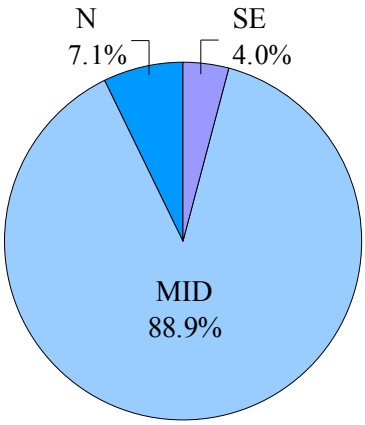
(567) And we have done all **us lives**, and then when she got married we still kept in touch with one another... (prenominal possessive *us*, NotA58, NTT_006)

As a dialect feature, possessive determiner *us* is not particularly frequent (2.3% of all 1PL cases), but the number of occurrences and the number of speakers who used the feature in FRED are significant enough (28 cases; SN=11).

Similar to possessive *we*, possessive *us* is not applied categorically. This is evident from the fact that all of its speakers also used standard *our*, showing considerable internal variation. Unlike possessive *we*, on the other hand, the head nouns which accompany possessive *us* are more varied, so that the collocational context can be neglected as a determinant ('HEAD' column in the spreadsheets; certain collocates require further

investigation, e.g. *mum* and *mother*).

The most noticeable result in terms of geographic distribution is that possessive *us* represents a typical feature of the MID area, with significant differences in frequencies between MID and the rest of the country. A cross-check of the recordings showed homogeneity between the different transcribers. The distribution of possessive *us* is visualised in Figure 14.1, based on weighted frequencies. The additional information to the right of the graph shows the number of standard deviations between non-standard frequencies of different dialect areas, where $\Delta_{(SE,\ SW)}$ simply is the difference between the proportion of *us* in possessive function in the Southeast as compared to *us* in possessive function in the Southwest. At a minimum confidence level of $p < .05$, the MID area is significantly different from all other areas.

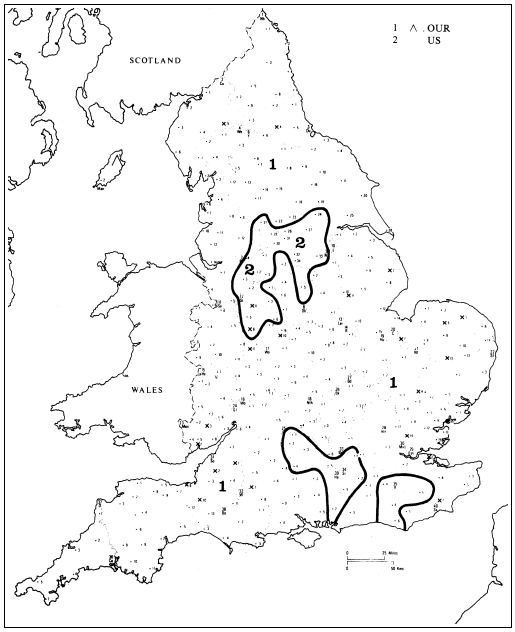


$\Delta_{(SE,\ SW)} = 1.0\ \sigma$ ($p = .62$)
$\Delta_{(SE,\ N)} = 0.6\ \sigma$ ($p = .79$)
$\Delta_{(SW,\ N)} = 1.7\ \sigma$ ($p = .39$)
$\Delta_{(MID,\ SE)} = 4.8\ \sigma$ ($p = .02$)*
$\Delta_{(MID,\ SW)} = 5.2\ \sigma$ ($p = .01$)*
$\Delta_{(MID,\ N)} = 4.6\ \sigma$ ($p = .02$)*

* significant at $p < .05$

Figure 14.1: Areal distribution: prenominal possessive *us*

The corpus results show that possessive *us* is not restricted to the Midlands, since occasional occurrences appear in other parts of the country, but it is predominantly associated with this area. The distinctive areal pattern recognisable in FRED is similar to other accounts, such as the SED. Based on the distribution of 'our/ us eyes', for example, the *Linguistic Atlas of England* (Orton, Sanderson and Widdowson 1978) identifies a clear possessive *us* area in the Midlands, shown in Map 14.1.



Map 14.1: ‘Our/ us eyes’ (LAE M75)

14.1.2 Possessive *thee*

The third most frequent non-standard possessive in the data is *thee*. Although there are only 6 occurrences, possessive *thee* is noteworthy for its archaic form.[159] Here, too, examples were only found in the MID and N areas (SN = 4). Most possessive *thee* appear in quotations of conversations from the past (similar to other 2SG *th-* forms). They are, therfore, not necessarily representative of the speaker’s own linguistic behaviour.

(568) Oh, alright, he said, eh, I ’ll let **thee dad** know if there is (sic) owt (/sic). That 's the way they talk, pardon me if I 'm talking that way you see same as they talk. (prenominal possessive *thee*, H3L, LAN_010d.wav)

(569) They used to wear out in about a month, runnin’ backards and forwards, ‘cos if these hills wun steep, thee ’d got to get **thee feet** in one o’ these steps, and if thee missed one, thee could’st have the tub on top on thee... (prenominal possessive *thee* with other *thee* in the co-text, IBJP, SAL_038)

(570) He said, thee just hold **thee noise** George, he said, thou might happen get kept on regular here. (prenominal possessive *thee* with other *th-* forms in the co-text, WesDX, WES_017)

Besides *we*, *us* and *thee*, the only other non-standard possessives are one case with *him*, shown in (554), and one case with *them’s* which will be mentioned again together with

159 The use of possessive *thee* has, for instance, been observed in Bristol (cf. Wakelin 1986: 34). The use of possessive *thy* will not be discussed.

other rare variants in 14.3. Because of the often discussed origin of English possessive *'s*,[160] it should be mentioned that no *his* genitives were found of the type *the man his house*. This particular construction, which was briefly in use during the Renaissance, is commonly found in modern colloquial German (*dem Mann sein Haus*) and Dutch (*de man zijn huus*).

14.2 *Of*-genitives: pronouns in possessive complement function

Compared to the distribution of pronominal variants in *s*-genitives, the distribution in *of*-genitives presents a very different picture. Here, the data show a relatively even repartition of possessive pronoun forms (*of mine*, *of yours*, etc., 96 cases, 38%) and personal pronoun forms (*of me*, *of him*, etc., 154 cases, 62%). *Of*-genitives of the former type, as in *an old mate of his*, *a relative of mine* or *neighbours of ours*, are often referred to as 'double possessives' or 'double genitives' in the literature. Although they differ structurally from *of*-genitives of the second type (*an old mate of him*), this label is potentially misleading, considering the frequent interchangeability of the two pronominal variants.

Apart from the modifier–head relation, which is an important determinant of pronoun variation (here, for instance, relations of possession or personal relationships), the use of O-form complements in interchangeable cases such as *this house of his/ him* or *a friend of mine/ me* appears to happen more often in the third person than in the first person (consider example (571)). In addition, the data show a complete absence of cases with *of its* and a near-absence of *of theirs*, which contrasts with the preference for possessive pronoun variants in the first person (speakers rather use *a friend of mine* than *a friend of me*). From a phonological perspective, all of these tendencies point in the same direction: avoidance of forms ending in -*s*. Further investigation is, however, required to clarify this point.

Sentences (571)–(575) show some examples with the two principal variants. Sentences (576) and (577) show the only two S-form occurrences found in the data, which takes us to some other rare variants discussed in the next section.

(571) Going ehr, coming back to a to a **friend of mine** that eh -- and I 'm still a very great **friend of him**, and he is **of mine** -- ehr Mr Paradine. (alternation of *of*-genitives with possessive pronouns and personal pronouns, MdxCG, MDX_001h.wav)

(572) But, of course, my wife didn't like tripe -- and that was a favourite **meal of ours**. (*of*-genitive with possessive *ours*, YksSL, YKS_003)

160 Roughly, whether the apostrophe stands for missing -*e*-, having developed from the Old English genitive case ending -*es*, or missing *hi*-, having developed from the possessive pronoun *his*. For a detailed study see Rosenbach (2002).

(573) Oh yes, we all, the **family of us** went to Coalford Chapel, morning, afternoon and night. (*of*-genitive with possessive *us*, IBErDi, SAL_020)

(574) Oh I couldn't tell you exactly where it was, but they used to deliver this **flour of theirs** to the town mill and they were a sort of agents for it you see. (*of*-genitive with possessive *theirs*, TCA_FP, DEV_005)

(575) But eh somebody got hold of their -- somebody got (v 'laughter') hold of **the deeds of them**, they was sold to the Catholic (v 'laughter') to the Catholic Church (v 'laughter'). (of-genitive with possessive *them*, CP, MDX_002, trba.)

(576) I ʿda try and make ʿim walk at the **side of I** you know,... (*of*-genitive with S-form, TrbrGR, WIL_001c.wav)

(577) ...but I, I had a pair of boots; sometimes the **bottom of they** were out. (*of*-genitive with S-form, SRLM_RF, SOM_005f.wav)

14.3 Rare variants

In the corpus, there are several rare morphological variants which deserve mentioning. First of all, there are two examples of possessive pronouns in *-n*, which may have developed historically by analogy to *mine* and *thine* (cf. Montgomery 2004, Appalachian English). The fact that both examples from the corpus appear in interviews from Wiltshire supports Trudgill's description of a regularised pattern in *-n* in dialects of the South of England (1999[2]: 90).

(578) And now my age, uh, my age were about July, and I think **hern** were on the following January. (dialectal possessive *hern*, WflsWGP, WIL_010a.wav)

(579) He do come round ʿere an' fly low an' say that man 's got his corn in so we shall have to get on with **ourn**. (dialectal possessive *ourn*, WflsWH, WIL_011)

Extremely rare is the use of *mine* as a prenominal determiner. A possible explanation can be found in the historical development of 1SG possessives, given that *my* originally developed as a reduced form of *mine*, and given that *mine* continued to exist alongside *my* for some time after the shorter form emerged – especially before words beginning with a vowel. This is also the case in (580).

(580) So, on this particular Saturday, I come home, my mother had been washing in the wash-house, you know. Eh only **mine** and her washin'. And eh, she hadn't done a thing in the place. (prenominal possessive *mine*, FM, LND_003e.wav)

A different type of variants is characterised by additional possessive *-s* on forms which do not have this marker in Standard English. The resulting non-standard determiners *mi-s* (1 case), *them-s* (1) and *our-s* (2) can be seen in (581) and (582), as well as (562) and (563).

In one utterance, *-s* was added to the possessive pronoun complement, resulting in *mine-s* (583).[161] Note that prenominal *he-s* or *he's*, which also occurs in the transcripts, is considered to reflect the stretching of the vowel in *his*, as in *Another man 'd come, stick he 's glass on top* (LAN_020).

(581) But I 'd a granny and that were **mi 's mother's aunt** and we called her Nanny. (prenominal possessive *mi's*, LanW1P, LAN_004d.wav)

(582) It was, you know, an individual, a family firm and then during the war these people died, well in came one of **them's daughters**... (prenominal possessive *them's*, TrbrRS, WIL_020)

(583) So, him and a **a friend of mine's**, my fathe- eh my friend's father, a Mr Dell, eh, they picked up this here figure of a man, pointing to the polling-booth. (possessive pronoun *mine's*, MdxCG, MDX_001d.wav)

Non-standard markings in *-s* are hardly ever mentioned in the literature, but they are not unheard of. In his early *Dictionary of Archaic and Provincial Words*, for example, Halliwell (1887, Vol.I, xxxii) mentioned the use – in his view "perversion" – of genitive *she's* in the Worcestershire variety, as in *she's bonnet*. Elsewhere, the use of possessive *theys* has been observed in Bahamian English, as in *it's theys boat* (Reaser and Torbert 2004: 399).[162] Forms similar to the ones found in FRED have been described for Norfolk-Island English by Mühlhäusler (2004: 798), including 1SG *mais*, 2SG *yus*, 3PL *dems*. The regularisation of *mine* to *mine's*/ *mines* has been observed in other varieties than British English. Wolfram (2004: 333), for example, has described this particular feature as "quite robust in most varieties of AAVE, though it appears more typical of preadolescent speakers than older speakers", and Miller (1993: 108) has mentioned its occurrence in Scottish English. Since there is only one example in FRED, and no comparable cases were found in the SED, we must assume that occurrences of *mine-s* in British English are extremely rare.

Overall, the rarer variants in the corpus indicate that some interference is to be expected – either along the lines of analogy formation or regularisation – in a language which uses different means for marking the possessive on nominal and pronominal NPs. A plausible explanation for the mentioned variants in *-s*, for instance, is analogy with nominals. To insert an anecdotal comment for once, I myself recently heard an Australian friend refer to

161 In the previous chapter, *our-s, mi-s* and *mine-s* were included in Table 13.2, *them-s* under *them* in Table 13.1. Note that the speaker who used prenominal *ours* also used various other non-standard features, including O-form subjects, generalised *was* and other regularisations.

162 The authors also mention the categorical absence of possessive *'s* on nominals, and the use of postverbal *theys* in sentences like *what theys is theys*.

his and his wife's wedding as 'Sara and I's wedding', adding clitic '*s* to the coordinated pronoun the way it is usually done with coordinated nominals (e.g., *That was [my mother and father]'s shop* (SAL_026); *[your daughter and son-in-law]'s affairs* (SFK_010)).

Furthermore, the use of identical possessive markers on nouns and pronouns is a typological possibility which is customary in the standard varieties of other languages. In Finnish, for example, the nominal possessive marker -*n* is added to personal pronouns to form the genitive case: *Kallen auto* 'Kalle's car'; *minä* 'I' > *minun* 'my'; *sinä* 'you' > *sinun* 'your', and so forth.

14.4 Summary

The corpus results show that variation between different pronominal categories is found in both genitive structures. The nature and determinants of this variation, however, differ between *s*- and *of*-genitives. In *s*-genitives, 1PL *we* and *us* merit special attention. The two pronouns differ with regard to their areal distribution, possessive *we* representing a typically Northern feature, possessive *us* a typical feature of the Midlands (both occasionally appearing elsewhere). It could be argued, of course, that possessive *we* and *us* represent but different phonological realisations of standard *our*, which would mean that a Northern realisation closer to /wər/ or /wiər/ contrasts with a Midlands realisation closer to /ʌs/. However, an alternative interpretation suggests itself when listening to the recordings: it could well be that one feature is phonological, consisting in the erosion of *our* to /wər/ or /wiər/, whereas the other is morphological/ lexical in nature, i.e. O-form *us* substituting standard possessive *our*. The latter view is supported by the fact that it is often difficult to keep apart *we* and *our* in the soundfiles, whereas *our* and *us* are always distinct.

Of-genitives show a much more even distribution of pronoun forms. Interchangeability between personal pronoun forms (O-forms) and possessive pronoun forms is given in certain modifier–head relations, such as those indicating possession (*this house of his/ him*) and personal relationships (*a friend of his/ him*).

The last section of this chapter was dedicated to some rarer variants with additional possessive markers -*n* or -*s* which can be explained as analogy formations or regularisations. These variants, which are too rare for an in-depth analysis in FRED, may be worth investigating on an even larger dataset, or in an even more informal register than the interviews at hand.

Chapter 15
Demonstrative function

15.1 *They* and *them* demonstratives

Among the different personal pronouns in English, the 3PL forms *they* and *them* occasionally fill the position of a demonstrative pronoun or demonstrative determiner, replacing standard *these* and *those* in phrases like *them days* or *they were good times*. While this use of subject and object forms is non-standard in English, the existence of identical or derivationally related 3PL personal pronouns and demonstratives is typologically attested in a great number of languages (cf. Bhat 2005; see Map D10, Appendix).

When used as demonstratives, non-standard *they* and *them* are necessarily compatible, semantically and functionally, with the meaning assigned to the demonstrative slot by the corresponding construction (cf. Stefanowitsch and Gries 2003: 213). The phenomenon is well known in the linguistic literature. Demonstrative *them*, in particular, has been described for many varieties in and outside the British Isles. It appears among the top non-standard phenomena of American and Australian varieties, and it is one of the three most prominent dialectal pronoun features in English worldwide (besides subject *me* in coordination and distinct 2PL forms; cf. Kortmann and Szmrecsanyi 2004: 1163/ 1198). The *Handbook of Varieties of English* (Kortmann et al. 2004) lists the use of *them* instead of *those* among the top twenty vernacular features of British English, found in seven out of eight varieties. Coupland (1988: 35) also lists demonstrative *them* as one of the most widespread dialect features of British English; and Cheshire and Edwards (1998: 64) found that demonstrative *them* was reported by more than 90% of the schools which took part in their extensive survey of British dialect grammar (questionnaire-based). According to Cheshire and Edwards, similarly widespread features do not represent social features.

Descriptions of (distal) demonstrative *them* in England include the Southeast (cf. Anderwald 2004: 179), East Anglia (Trudgill 2004a: 148), the Southwest (Wagner 2004b: 164), and the North (Beal 2005: 119).[163] Filppula (2004) has investigated the use of

163 Beal also mentions the less frequent use of demonstrative *they* in the Northeast. In the SED, demonstrative *them* is documented in various questions and in the incidental material, for example Question IX.10.5 *them ones*; Q.IX.10.6 *them over there*; Q.IX.10 incidental material 25,4 *them two*, *them over there*; Q. IX.1. incidental material 21,6,iJ *them plates*. The OED describes the phenomenon as "now only dialectal or illiterate" ('Them', section III.5).

demonstrative *them* in the Irish English data of the NITCS corpus. His statement that "*them* in this function is one of the most commonly occurring features of non-standard British English dialects, both urban and rural" (p. 92) is supported by the present study. It will be shown in the following that *they* and *them* are highly frequent in demonstrative positions (DEM). It will also be shown that, while demonstrative *they* and demonstrative *them* occur in the same syntactic contexts, they differ considerably in frequency and areal distribution: *them* is a highly frequent supraregional variant, whereas *they* appears almost exclusively in the SW. Last but not least, it will be argued that certain constructions, and certain collexemes within these constructions, have a positive influence on the use of non-standard demonstratives. This will be demonstrated with the aid of collostructional analyses after Stefanowitsch and Gries.

The demonstrative system described in the following is dual in nature. There is little evidence in FRED for a threefold 'proximal–distal–remote' system as has elsewhere been described for traditional British English dialects (e.g., Trudgill 1999[2]: 86, remote *yon*/ *thon*/ *thik*; Harris 1991: 22, *this–that–thicky*).[164] Furthermore, the analyses presented below make no proximal–distal distinction, since (a) it is usually impossible to establish whether the non-standard demonstrative in question replaces *these* or *those*, and (b) *these* and *those* vary themselves regarding this distinction. In the FRED corpus, *these*, for instance, is repeatedly used as a distal demonstrative in sentences like *there were no cages in these days* (SAL_034).

Regarding the coding procedure, prenominal demonstrative determiners as in *them kids* were separated from non-demonstrative uses of other pronouns in qualified NPs such as *us kids* (coded POS 90 and POS 62 respectively). Cases with *all them* were included in the analyses if replacing 'all those/ these/ this' rather than 'all of them', for example *I knowed all them what was killed* (KEN_003); *all them were houses* (LAN_020); *the sailors and all them used to come out* (KEN_005). Non-standard, non-demonstrative *they* and *them* were discussed under pronoun exchange in chapter 8.

164 Wagner (2004a: 59) comes to a similar result: "Although a close–distant–remote system has been postulated for Southwestern dialects in some modern studies, this assumption is supported neither by traditional accounts nor by data from the corpora, as examples of a threefold distinction are non-existent or at least difficult to find. Judging from the examples, the traditional system has declined, and the form *thick(y)* has all but died out, with a total of some 20 forms in the corpus." Note that listings of regional variants in *th-*, such as *thicky, thucker*, etc., can be found in the SED documentation and Wakelin (1986).

Hey! Teachers! Leave them kids alone!
(R.Waters/ Pink Floyd "Another Brick in the Wall")

15.2 Syntactic environments

Demonstrative *they* and *them* occur in different syntactic environments. In simplex cases where the pronoun stands on its own, the difference between a demonstrative deictic and a subject pronoun can be very subtle. The functional attribution does, to a certain extent, depend on the hearer's perception: consider *them were yours* 'these were yours'/ 'they were yours'. Determining the relative frequencies of simplex demonstrative *they* and *them* as compared to *these* and *those* is, therefore, not a straightforward task. In ambiguous cases, a substitution test was used to establish demonstrative status: if *them* could be sensibly replaced with *these* or *those*, it was classified as demonstrative, otherwise according to its function in the sentence. In *My mother and all them were down there* (SFK_005), for example, *all them* was classified as a coordinated subject. In examples (584)–(587), on the other hand, simplex *they* and *them* were attributed demonstrative function. Example (587) illustrates the possibility of demonstrative *those* and *them* occurring in identical contexts in the same sentence.

(584) I worked all day Saturday, you see. And that was from eight o'clock at morning till eight o'clock at night then, you know, when -- That them **them** was the shop hours then. (simplex demonstrative *them*, H3L, LAN_010g.wav)

(585) ...but **they** was good times really, better than what we 're having now. (simplex demonstrative *they*, WigWHJ, LAN_020)

(586) Richard Baxter and Sarah Baxter, **them** were the names,... (simplex demonstrative *them*, IBCH, SAL_034)

(587) **Those** used to be my happy days -- I -- **them** used to be my happy days. (demonstrative *them* and *those* in same sentence, IBMrsF, SAL_028)

In other syntactic environments, demonstrative status is easier to ascertain: the demonstrative is followed by a qualifier noun or numeral (*them kids*/ *them three*), by a relative clause (*they as could afford it*), or by an adjunct (*them in front of you had the best chance*). Most frequently, demonstrative *they* and *them* combine with qualifier NPs, which can be broken down into 4 groups: nouns (*them kids*), numerals (*them three*), combinations thereof (*them three kids*), and personal names (*them Giddenses*). The different types of qualifiers all allow for intermitted adjectives, as shown in examples (592) and (595).

15.3 Quantitative distribution in the corpus

The first basic difference between *they* and *them* in demonstrative function is readily recognisable from their absolute frequencies: the comparison between 159 *they* and 1414 *them* shows that the latter variant is far more common. Demonstrative *them* is even highly common compared to the two standard variants, of which there are 2698 occurrences overall (1785 *these*, 913 *those*). Together, *they* and *them* account for over 36.8% of all DEM in the corpus – a high non-standard proportion in a highly frequent syntactic function.

In 'DEM + N' phrases, the more frequent variant *them* alone accounts for 32.6% of all occurrences. In addition to the variation between standard and non-standard demonstratives, various *they* demonstratives form part of NPs in object or prepositional complement function (approx. 23%), and various *them* demonstratives form part of NPs in subject or subject complement function (approx. 14%).[165]

(588) See, but they was, **they two** was working together. (DEM '*they* + numeral', CAVA_DB, CON_001b.wav)

(589) They four **they four chaps** from London, they looked at me as though they gonna shot me. (DEM '*they* + numeral + noun', SRLM_RM, SOM_014a.wav)

(590) Like a barber's, you know, like you 're shaving, uh, cutting anybody's hair. They was **them sort of clippers**. (DEM '*them* + noun', KentEDG, KEN_007a.wav)

(591) ...or whatever they do think of all **they Giddenses** and all **they cronies** and all that what used to get up this chapel... (DEM '*they* + personname' and '*they* + noun', WflsWGP, WIL_010)

(592) The only thing I drove is one of **they old fashioned old tractors**. (DEM '*they* + noun', WflsWH, WIL_011)

(593) And out there, and uh, **them Noy** had one [horse] up there, an old hairy-leg one, kick like the devil... (DEM '*them* + personname', CAVA_DB, CON_001)

(594) We 'd say, Are you going to change it, we 're fed up with **them chops and sausage**, can't you give us a bit of steak? (DEM '*them* + noun', NotA8, NTT_013)

(595) ...and then we used to decorate all this archway, every door had a archway, and every door had **them little fairy lights** on, then they were switched on at night. (DEM '*them* + noun', LAN002, LAN_002a.wav)

(596) It was a real needle match when **them two** met. (DEM '*them* + numeral', IBGJ, SAL_032)

(597) ...out we come, yeah, God Jesus, **them two lads them two brothers** have been in trouble again... (DEM '*them* + numeral + noun', NotA8, NTT_013c.wav)

165 Analytically qualified cases such as *two of them* were not included unless followed by a noun, as in *two of them kids*.

The speakers' use of demonstrative variants is characterised by internal variation. Occasionally, demonstrative *they* and demonstrative *them* are found in the same sentence or text passage and even with the same referent, showing non-persistent usage of non-standard variants. Elsewhere, speakers can be seen to vacillate between standard and non-standard variants, as shown in some of the following examples (also (587) above):

(598) **They** was, **Them crisps** was cheese. (DEM '*them* + noun', HM, KEN_003)

(599) ...they 'd got all new machines all **these multifeeds** and there was all old fellows who 'd been on loopwheels and they 'd brought them on to these machines and and they 'd they still you know with them getting on, they couldn't grasp they were making more bad work [...] It 's impossible, he says, It 's more than what they 're getting off on **them multifeeds** [...] Well then, with **these here multifeeds** you could see what was coming off this... (alternation of DEM *them* and *these* with same referent, A15, NTT_014d.wav–NTT_014f.wav)

(600) ...there was a stud at Worpley kept Clydesdale stallions, where you took your mares there, **they** were, **them** were Clydesdales,... (vacillation between DEM *they* and *them*, YksWG, YKS_009)

(601) If you open **they gates**, then think you were goin' to get **them pigs** out... (DEM *they* and *them* in same sentence, WflsWGP, WIL_010)

(602) ...right round by the western works there, where **them that** old powder mill is. (DEM *them* corrected to *that*, HM, KEN_003c.wav)

The highest frequency for demonstrative *them* was obtained in 'DEM + relative clause' occurrences. Over 55% of all demonstratives followed by a relative clause have *them* (together with *they* over 60%). Various of these examples also have non-standard relative pronouns as seen in (604) and (605). The corpus findings clearly contradict studies which rule out the use of non-demonstrative forms with appositions (e.g. Quirk et al. 1985: 352–353, **They who work hard deserve some reward*).

(603) And **they that** would, any that have it, they wasn't allowed out. (DEM '*they* + relative clause', referring to cattle with foot and mouth disease, SRLM_GG, SOM_013)

(604) ...only **those as** was in work, **them as** was working at Celanese or Players... (DEM '*them* + relative clause', preceded by standard *those* in same context, NotA30, NTT_003)

(605) ...you used to pay them for it, **them** you see **what** 's got it for you. (DEM '*them* + relative clause', NotA58, NTT_006)

The relative clause context is closely followed by 'DEM + adjunct', where *them* accounts for 41% of all cases (together with *they* almost 48%). In the SE area, *them* even predominates over *these* and *those* in this context.

(606) Oh, **they at the, at the factories**, like Morlands 'd be, you know, they did do the sheepskin... (DEM '*they* + adjunct', SRLM_WB, SOM_010)

(607) ...and then there were Snooks and all (reg sic=thy) **they** (/reg) **around there** had the ground and never kept nothin'. (DEM '*they* + adjunct', WflsWGP, WIL_010)

(608) Ellers Farm was their own, **them at Brow Head** [sheep] was their own. (DEM '*them* + adjunct', WesBR, WES_008)

(609) ...but when you went to use the toilet, especially **them with no bucket**, you got to first of all disturb any rats that were in amongst that filth -- that was one feature that I always watched for, or else if you didn't you 'd get your penny nibbled, and that 's been known to happen. (DEM '*them* + adjunct', IBME, SAL_023)

A special subcategory of adjuncts are the locatives *here* and *there*, used as reinforcers in sentences like *these here big quilts* or *they used to pay one another with them there* (cf. Trudgill 2004a: 148). The use of demonstrative *they here* and *them there* before nouns is well known in the literature and is usually considered colloquial. Interestingly, Shields (1994: 311) has noted that "Indo-European originally marked the distinction between emphatic and nonemphatic personal pronouns by affixing deictic particles with 'here and now' deixis as markers of the emphatic." According to Shields,

> the question can be raised as to why so many different 'here and now' deictics are attested as extensions of pronominal forms. To my mind, the answer lies in the tendency, noted earlier, for deictic particles to weaken in deictic force through time. As certain particles started losing some of their 'emphatic' potential, they became less productive as other stronger forms took their place. (Shields 1994: 312)

Unfortunately, a detailed investigation of this particular feature is not possible with the data at hand. The only occurrences in FRED consist of 4 *them there*; no examples were found with *them here*, *they here* or *they there*. The only more frequent combination with a standard demonstrative is *these here* (73 cases). Individual occurrences with *this* and *that* can also be found in the corpus, for example *This here bloke come to him to give him one* (NTT_012), or *I said, That there bathroom cabinet!* (DEV_010).

(610) They used to be windows but they used to all be **them there** little tiny ones like that. (DEM *them there*, NotA58, NTT_006)

(611) I 've a-knowed **them there** pigs up in our ground down here... (DEM *them there*, WflsWGP, WIL_010)

(612) I never used to use the bunk, used to have like one of **these here** big quilts, that women used to make with patchwork, you know. (DEM *these here*, SRLM_WT, SOM_028)

15.4 Areal distribution

Not only is *them* the far more common non-standard variant in DEM positions (over 33% of all DEM in the corpus), it is also radically different from *they* regarding its distribution across the four dialect areas. Demonstrative *them* is a high-frequency supraregional feature, whereas demonstrative *they* is much rarer (3.7%) and appears almost exclusively in the SW. Accordingly, *them* is less frequent in the SW than in other parts of the country.

Perhaps surprisingly, this very clear distributional distinction has not yet been given much attention in the literature.[166] The results are all the more interesting considering that they coincide with the SW-specific use of non-standard S-forms in connection with a different grammatical phenomenon, viz. pronoun exchange (ch. 8).

In FRED, the same distributional tendency is observable in different DEM positions, including simplex cases. Figure 15.1 shows the accumulated distribution of the four lexical variants in 'DEM + noun/ adjunct/ relative clause' contexts, visualising the higher frequency and supraregionality of demonstrative *them* as compared to the lower frequency and geographical restriction of demonstrative *they*.

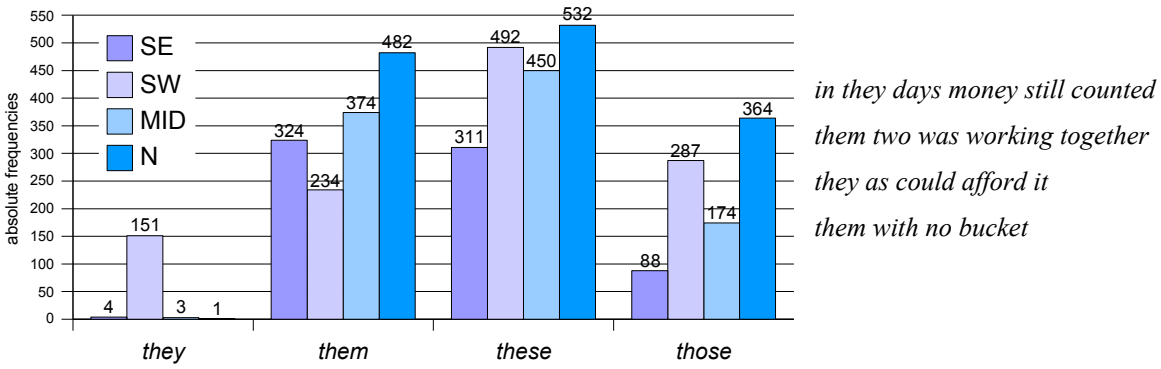


Figure 15.1: Areal distribution: *they*, *them*, *these*, *those* in DEM positions

166 Wakelin (1984: 82) comes to the same conclusion (unquantified).

15.5 Applying collexeme analysis to dialect data

They and *them* demonstratives occur particularly often in certain syntactic constructions, and they seem to combine particularly often with certain nouns. But are the observed combinations more frequent than we would expect by chance? In order to answer this question, it was decided to apply the collostructional approach, which offers powerful tools for quantifying the previous observations. The analyses presented below serve as an example for how the investigation of dialect phenomena can benefit from such an approach.[167]

The main goal in this particular analysis is to identify distinctive elements in DEM positions. I follow the methodology developed by Stefan Th. Gries and Anatol Stefanowitsch (2003, 2004, 2010; Hilpert 2006),[168] which has been applied by the authors themselves to a variety of syntactic environments, including verbs in ditransitive constructions, semantically motivated association patterns in *will*- vs. *going to*-futures, and grammatical alternation between the English *s*- and *of*-genitive.

In this study, the Coll.analysis software (Gries 2007) was used for two different types of analyses: *distinctive collexeme analysis* was used to test whether *they* and *them* demonstratives occur particularly often in certain syntactic constructions; *covarying collexeme analysis* was used to identify distinctive nouns in 'DEM + N' phrases. The software does not yet allow for additional variables such as register or dialect area, but it may do so in the future.

15.5.1 Distinctive collexeme test

Distinctive collexeme analysis helps identify those words that best distinguish between semantically near-equivalent constructions. It is generally used to compare the frequency of a word in one construction C to its frequency in the corresponding slots in near-equivalents of C (cf. Gries and Stefanowitsch 2010). In the present study, this approach was used to compare the frequencies of the four lexical DEM variants *they*, *them*, *these* and *those* in the four environments distinguished above:

167 The predictive superiority of collostructional analyses over raw frequency counts has been described by Gries, Hampe and Schönefeld (2005).

168 The idea to apply the collostructional method to the FRED data came up after a talk by Martin Hilpert. I am very grateful to Martin for introducing me to the Coll.analysis software.

- simplex;
- 'DEM + N' (including nouns, names, numerals, intermitted adjectives);
- 'DEM + relative clause';
- 'DEM + adjunct'.

Table 15.1 shows the results of a multiple distinctive collexeme analysis with the four lexical variants and the four construction types. The software computes one *p*-value for each variant in comparison with each other variant in each construction, using exact binomial tests based on actual co-occurrences in the corpus (column 2) vs. expected co-occurrences (column 3).[169] The results are then log-transformed so that highly positive and highly negative values indicate a large degree of attraction (+) or repulsion (–) (0 indicating random co-occurrence). The resulting values for collocational strength (column 4) indicate how strongly the observed frequency of each variant differs from its expected frequency. In addition, the last column shows which construction is most strongly associated with which variant (my addition; Coll.analysis displays the largest deviation).

Table 15.1: Multiple distinctive collexeme analysis: lexical variants in DEM contexts

DEM	construction types observed frequencies				construction types expected frequencies				collocational strength				strongest association
	Simplex	Noun	Rel.cl.	Adjunct	S	N	R	A	S	N	R	A	
they	12	149	6	4	15	149	5	2	–0.61	0.26	0.48	0.73	A
them	35	1,317	72	25	128	1,262	40	19	–23.12	**5.37	**5.67	1.00	R
these	225	1,753	11	21	178	1,751	55	26	**3.74	0.34	–12.59	–0.74	S
those	142	862	40	11	93	919	29	14	**6.35	–6.41	*1.54	–0.55	S

S = simplex DEM; N = 'DEM + N'; Rel.cl. = 'DEM + relative clause'; A = 'DEM + adjunct'.
* attraction significant at $p < .05$, ** attraction significant at $p < .001$.

The results agree – perhaps to a surprising extent – with previous assumptions: the two standard variants *these* and *those* are most strongly associated with simplex cases

169 Expected co-occurrences based on a comparison of the variants' distribution in the corpus. Note that regular collexeme analysis uses the Fisher Yates Exact test to compute the association strength between elements; for tests with more than two alternatives the software uses an exact binomial test that is not sensitive to low frequencies.

(collocational strength > 3, ≙ $p < .001$) whereas repulsion is indicated for *they* and *them* in simplex cases (although instances do occur in the corpus). The two non-standard variants are generally associated with non-simplex cases (collocational strength > 1.3, ≙ $p < .05$), demonstrative *them* particularly strongly with 'DEM + N' and 'DEM + relative clause' constructions. For the distribution of the four DEM variants in FRED, and quite possibly in other data of spoken English, too, this means that the standard demonstratives *these* and *those* are most likely to be replaced by non-standard *they* and *them* before nouns or noun-like expressions, and when followed by an adjunct or relative clause.

15.5.2 Covarying collexeme test

As a second test, covarying collexeme analysis was used to assess the association of demonstrative variants with specific head nouns in 'DEM + N' phrases. Similar to traditional collocational methods, this approach investigates dependencies between words in two different slots in a given construction, but at the same time it pays closer attention to syntactic and semantic structure (cf. Gries and Stefanowitsch 2010). In the analysis presented below, this is reflected in the fact that the tested association is not between the demonstrative determiner and its first collocate to the right, but between the demonstrative determiner and the main head noun of the phrase. For an expression like *in them bloody days* the input would be DEM = *them* and COLL = *days*:

in them bloody days:	DEM = *them*; COLL = *days*;
they iron-wheeled ones:	DEM = *they*; COLL = *ones*;
these type of things:	DEM = *these*; COLL = *type*;
those tin bath tubs:	DEM = *those*; COLL = *bath tubs*; etc.

Remarkably clear results were obtained for the attraction of demonstrative *they* and *them* to certain nouns, in particular the highly frequent time noun *days*. Collocations like *they days* and *them days*, but also *they/ them times*, are well-known dialectal features which have come to be recognised as set expressions in colloquial English. *Them days*, in fact, is so frequent that it may well be on its way to freezing – both in its prepositional and adverbial use:

(613) {<u Int> And were, did you get payment for these extra um times that you had to sing?} Oh no, no payment **in they days**. (TCA_FP, DEV_005b.wav)

(614) No, well time were nothing, was it, **in they times**. (SRLM_HW, SOM_030)

(615) So anyhow they never had no -- never had no glasses nor nothing **in them days**, you know. (CAVA_WW, CON_006d.wav)

(616) Nah, there was no trade unions **in them bloody days**, no unions at all,... (NotA8, NTT_013)

(617) He was a great old chap. Never used to ride **in them times o' days**, always walked, you know... (MdxCG, MDX_001)

(618) Course, **they days**, same as I said, 't was all driving, you know,... (SRLM_HG, SOM_032)

(619) Yoah-aye, yeah, they were murder then, the husbands **them days**, boy. (JG, LND_004)

Among all collexemes in 'DEM + N' phrases, the strongest attraction and repulsion values were obtained for *days* (see Tables C3–C6, Appendix). Table 15.2 shows the analysis results. In FRED, there are 1,347 occurrences of 'DEM + *days*' alone, of which 837 have *they* or *them* demonstratives (62.1%). Among all collexemes, *days* is the one most strongly attracted to both non-standard variants, with positive collocational strengths of 9.23 for *they* and 107.89 for *them*. A strong attraction also exists between *days* and *those*, with a collocational strength of 66.39. The noun is, however, strongly repulsed by *these* (usually proximal), as can be expected in interviews where speakers mainly talk about the past (collocational strength = –indefinite). Based on the collocational strength values, the preference for different demonstrative determiners with head noun *days* can be summarised as *them* > *those* > *they* > *these*.

Interestingly, a similar tendency can be observed for *times*, although this collexeme is generally less frequent than *days* and more strongly attracted to *they* (frequency in C = 23; collocational strength *those* = +1.41, *they* = +2.06, *them* = +0.31, *these* = –3.35).

Table 15.2: Measured attraction for different DEM variants and collexeme *days* (only cases with an identifiable noun-slot lexeme; not including 'DEM + person name'; irregular plurals regularised, e.g. 'them boat': COLL = *boats*)

'DEM + N' construction (C)					
DEM	observed freq. DEM in C	observed freq. *days* in C	observed freq. DEM + *days* in C	expected freq. DEM + *days* in C	collocational strength
them	1,309	1,347	752	438	+ 107.89
those	857	1,347	504	286	+ 66.39
they	147	1,347	85	49	+ 9.23
these	1,713	1,347	6	573	– infinite

The areal distribution of the different variants in 'DEM + *days*' occurrences tallies with the distribution described in 15.4: *them days* is the preferred non-standard variant, with high frequencies in all dialect areas, whereas *they days* is used exclusively in the SW. Together, *them* and *they* account for over half of all demonstrative occurrences with head noun *days* in each individual area: SE 74.9%, SW 58.4%, MID 69.4%, N 56.7%.

15.6 Summary

The use of non-standard demonstrative *they* and *them* was analysed from different perspectives, combining simple quantitative comparisons of different DEM variants with a collostructional approach. The most important result concerns the frequency and areal distribution of the two variants: demonstrative *them* represents a high-frequency supraregional feature whereas demonstrative *they* is much rarer and appears almost exclusively in the SW. Any description of *they* and *them* demonstratives should take this fundamental difference into account.

Collostructional methods have elsewhere been applied to spoken corpus data, but not specifically dialect data. The analyses presented in this study illustrate their applicability to dialect corpora like FRED. Distinctive collexeme analysis was used to prove that non-standard demonstratives are positively influenced by certain syntactic environments. It could furthermore be shown, by using covarying collexeme analysis, that in 'DEM + N' phrases certain head nouns favour non-standard occurrences. In particular the collexeme *days* most strongly attracts demonstrative *them*, although *days* occurs with all four lexical variants: *them days* > *those days* > *they days* > *these days*.

The corpus findings are summarised in the Context hierarchy of non-standard demonstratives shown in (620). The use of non-standard *they* and *them* in spontaneous conversation presents itself as a gradient phenomenon which is strongly influenced by the immediate syntactic and collocational environment. Non-standard demonstratives are most likely to occur with collexeme *days* and where DEM is followed by a relative clause. They are slightly less frequent before adjuncts, and least frequent, although still very common, in 'DEM + N' phrases with head nouns other than *days*.

(620) Context hierarchy of non-standard demonstratives, based on FRED

DEM + *days*	>	DEM + rel. cl.	>	DEM + adjunct	>	DEM + other NP
62% ns		60% ns		48% ns		23% ns

Part III

Synopsis and discussion

Chapter 16
The major empirical results

In this study, a variety of phenomena were analysed which revolve around the interchangeability of pronominal expressions in contrast to the prescriptive rules of written Standard English. While the individual phenomena differ regarding their distribution and variation frequencies, they all point in one direction: pronoun variation is a regular characteristic of spontaneous spoken dialogue. The empirical results are no doubt incompatible with the prescriptivist concept of discrete and invariant linguistic units. In the following, the results obtained for different pronominal forms and syntactic functions will be summarised for the sake of a few more general observations. The synopsis will bring to light the most salient patterns of pronoun variation and will help establish connections between empirically substantiated tendencies and general linguistic principles.

16.1 Areal distribution of non-standard pronouns

The preceding chapters offered a comprehensive investigation of a great variety of phenomena involving variation in the use of personal pronouns. Some of these phenomena are among the most prominent dialect features in English worldwide, such as the use of subject *me* in coordination or the use of demonstrative *them* (8.3, 15.1). Others, for example resumptive pronouns, ethic PPs or the use of existential *it*, are much less common in English in general and exhibit low relative frequencies in the data at hand. Nevertheless, there are tendencies which apply to both the frequent and the rare phenomena in the corpus. One of these tendencies is the general supraregionality of the phenomena under discussion, which characterises most of them as features of spontaneous spoken English rather than a specific dialect. Whether the phenomena form part of a spoken standard very much depends on the definition of the term (cf. Van Marle 1997; Weiß 2004; Kortmann 2002). What is certain, is that the observed supraregionality is partly in conflict with previous accounts. Even the very few phenomena which show regional clustering in the corpus (Table 16.1) are characterised by gradient differences between the four dialect areas. The results, therefore, provide strong empirical support for a definition of regional varieties in terms of qualitative *and* quantitative distinctions. In the words of Chambers and Trudgill:

> Lects may differ quantitatively when a variable is involved. That is, lects may be distinguished not only by the presence or absence of a variable, but also by the frequency with which a particular variant occurs. (Chambers and Trudgill 1998[2]: 129)

For the great majority of phenomena investigated in this study, the ubiquitous use indicates a natural structural development that may have happened independently in different places. As Chambers (2004) has pointed out, diffusion by a certain group of speakers is implausible in the case of ubiquitous phenomena because of geographic spread. According to Chambers,

> [i]t is also implausible linguistically, because these features occur not only in working-class and rural vernaculars but also in child language, pidgins, creoles and interlanguage varieties. Therefore, they appear to be natural outgrowths, so to speak, of the language faculty. (Chambers 2004: 128)

Chambers' arguments are easily transferred to this study, especially since pronoun variation has not only been evidenced to play a role in most non-standard varieties of English but also in language acquisition and other languages (see 17.1). In addition, the supraregional appearance of non-standard cases distinguishes the phenomena under investigation from phonological phenomena, in line with the well-known fact that "grammatical and syntactic features [...] may unite areas showing a great deal of phonological differentiation" (Ihalainen 1994: 248). The corpus results show that a geographical demarcation of dialects which is based on phonological criteria (Map 4.1) can differ considerably from the grammatical variation observed in the same space.

A summary of the different empirical results is given in Table 16.1 in the form of a feature catalogue which contains the investigated phenomena and their areal distribution in the FRED corpus. Most phenomena are distributed relatively evenly throughout the country, with very few exceptions. Non-standard O-forms, for example, appear in all four dialect areas irrespective of their function or the surrounding syntactic structure. One minor exception is possessive *us*, which is typical for the Midlands (see ch. 14).

Most noticeable among those phenomena that were identified as regional – i.e. phenomena with statistically significant differences in frequencies between the four dialect areas – is the use of non-standard S-forms in the Southwest. This includes both the use of demonstrative *they* (as compared to supraregional demonstrative *them*, see ch. 15) and the

heavy clustering of S-forms in object and prepositional complement functions. The latter constitute the only regionally marked component of the complex phenomenon known as pronoun exchange (ch. 8).

Table 16.1: Catalogue of non-standard pronouns and their areal distribution in FRED

phenomenon	**areal distribution**	**chapter**
gendered pronouns	***South–North continuum***	chapter 7
pronoun exchange	O-form subjects: ***supraregional*** S-form objects/ prep.compl.: ***heavy clustering in SW*** O-form reflexives: ***supraregional*** independent self-forms: ***supraregional*** (mostly *myself*) O-form subjects in q-tags: ***typ. SW***	chapter 8
case variation after *as*, *like*, *than*	***supraregional***	chapter 9.2
gerund-participles, O-case modifiers	***supraregional default***	chapter 10
adverbial participles, overt O-case modifiers	***supraregional default***	chapter 10.5
qualified pronouns	***supraregional***	chapter 11
'pronoun + *all*' plurals	***supraregional***	chapter 11.1
second person plurals	***supraregional***	chapter 11.2
pleonastic pronouns	***all supraregional*** (overt *you* imperatives, pleonastic direct/ indirect objects, ethic PPs*, disjunctives, resumptives*)	chapter 12
possessive function	***supraregional*** ***us* determiners *typ. MID*** ***we* determiners *typ. N* (phon.variant?)**	chapter 14
demonstrative function	***them: supraregional*** ***they: almost exclusively SW***	chapter 15
singular *us* *they* with singular NP	***supraregional**** ***South****	chapter 6.1.1
dummy *it* in *walk it*	***supraregional****	chapter 6.2
existential *it* for *there*	***supraregional****	chapter 6.3
generic question tags	***supraregional*** ***innit: supraregional****	chapter 6.4

*rare feature

Taking into account the low mobility of the interviewees, the regional restriction of non-standard S-forms seems to agree with an exemplar-based conception of persistence, i.e.

language acquisition by exposure. Together with the observed variation in gender (gendered pronouns, South–North continuum), the few but very salient non-standard S-form features identify the Southwest of England as a special area: it shares all of the features found elsewhere but also has some grammatical idiosyncrasies of its own. This may be the reason why – while other geographic areas are associated with other grammatical phenomena[170] – the Southwest of England has long been notorious for its use of non-standard pronouns.

A different tendency observed in this study is the fact that most of the investigated phenomena represent so-called generalised features, meaning features that are used frequently and by a substantial number of speakers. In FRED, the vast majority of speakers use at least some non-standard pronouns. Despite the largest possible reduction of potentially influential sociolinguistic and stylistic factors (ch. 4), however, the current dataset still exhibits a considerable amount of unpredictable variation across and within the different interviews, down to the same sentence.[171] In other words, we are faced with inherent variability and internal variability in the Labovian sense (inter- and intra-personal; cf. Labov 1971; Trudgill 2003; Adger 2006). Consistent micro-parametric variation is not given, since individual speakers make inconsistent morphosyntactic choices.[172] Variation among speakers is of course difficult to grasp systematically. A straightforward, simplified solution was presented in the form of underlying speaker distribution patterns, as described in 5.2. The identified patterns, containing information on the number of speakers who used a certain feature and the mean speaker–case ratio in the corpus, were highlighted in connection with different phenomena in the analyses in order to rule out idiolectal language use and in order to accomplish the identification of individual outliers (e.g. in 7.2.8 and 8.3.7).

The goal of this study was to contribute to a better understanding of the distributional properties of pronoun variation. The empirically verifiable differences between generalised

170 In the North of England, for example, the Northern Subject Rule and a regularised pattern of reflexive pronoun forms (cf. Beal 2004; Pietsch 2005).

171 In his investigation of persistence phenomena, Szmrecsanyi (2006: 197) noted that "[m]ore often than not, FRED exhibited the lowest level of persistence [...] It is likely that this has less to do with the data sampled (dialect speech), but rather with the much higher mean age of speakers in the corpus." Szmrecsanyi (ibid.) also found that "[o]n aggregate, age seems to have a weakening effect on persistence." In addition, there appears to be a reduced chance for allo-repetition, or comprehension-to-production priming, in longer stretches of monologue (cf. Tannen 2007[2]; Cleland and Pickering 2003).

172 In a supplementary questionnaire which could not be included in this study, speakers were observed to contradict their own uses in their grammaticality judgements. Some of them, for example, used O-form subjects but judged them acceptable when presented with the same sentence later on.

phenomena, on the one hand, and geographically restricted or rare phenomena, on the other hand, make a case for the importance of cross-dialectal studies in the investigation of grammatical variation.

16.2 Frequencies

Considering the prevailing supraregionality illustrated above, the main difference between the individual phenomena in quantitative terms lies in differing degrees of variation. Overall, it could be established that semantically equivalent variants are mostly used variably. In spontaneous speech, in other words, the complementary distribution of semantically equivalent or near-equivalent expressions is the exception rather than the rule. In many cases, the data show straightforward 'complementarity failures' (Kiparsky 2002: 182). The corpus results hence serve to identify those grammatical rules which are not functionally relevant in actual conversations but represent a mere formality (e.g., positive–negative verb sequences in question tags, or question tag agreement in general).

If we boil down the results to concrete figures, we get a wide range of non-standard frequencies from 0.01% to almost 100%, with an average of 32.5% and a median of 16%. The frequency range goes from rare phenomena such as S-form objects (0.01%), O-form subjects (0.3%) or gendered pronouns with inanimate mass reference (0.3%), to almost-categorical occurrences which still contradict the prescriptive standard, such as non-S subjects in no-verb utterances (97.9%), non-S subjects of adverbial participles (94.4%) or O-form modifiers in nominal gerund-participles (99%).

The two syntactic contexts most resistant to non-standard inclusions are by-subjects and 'subject before finite V'. In syntactic functions with an extremely high non-standard proportion, on the other hand, prescriptive claims are hard to maintain. This is, for example, the case in subject complement function (84.1% non-S forms in the data), or in pronouns after comparative *like*, *as* and *than* (89.3% O-forms).

Table 16.2 shows a breakdown of non-standard phenomena into frequency ranges. Although non-standard frequencies in the corpus cover the whole range from 0.01% to almost 100%, most phenomena are either rare (featuring in the < 10% and < 20% groups), or highly frequent, with relative frequencies > 90%. In this regard, it is interesting to observe that the perceived salience of certain phenomena, and the attention they are given in the dialectological literature, are not necessarily correlated with their empirical frequency in actual performance data. Indeed, some of the most widely discussed

phenomena, such as pronoun exchange or gendered pronouns, range among the least frequent features in the corpus. S-form objects, in particular, have the lowest relative frequency of all non-standard pronouns in the data, but they are still being discussed in the literature as a highly salient grammatical feature of dialects spoken in the Southwest of England (salience is of course given due to the exceptional regional restriction of non-standard S-forms, as mentioned in 16.1).

Table 16.2: Breakdown of non-standard phenomena by relative frequency range

Relative non-standard frequencies (average 32.5%, median 16%)	
0%–5%	non-O objects non-S subjects independent self-forms (objects, prepositional complements, subject complements) self-forms after comparative *like*, *as*, *than* O-form reflexives non-O analytically qualified pronouns non-O prepositional complements non-O in ECM non-POSS determiners in *s*-genitives gendered pronouns, inanimate mass referents and inanimate count referents demonstrative *they*
5%–10%	independent self-form subjects in coordination non-O in for-to constructions S-forms after comparative *like*, *as*, *than* S-forms in qualified object NPs
10%–20%	independent self-form subjects, no V, simplex and in coordination non-O objects and prepositional complements in coordination O-forms in qualified subject NPs generic question tags
20%–50%	independent self-form subject complements in coordination non-standard case in 'all + pronoun' NPs gendered pronouns, animate non-human referents non-DEM demonstrative determiners demonstrative *them*
50%–80%	O-form subjects in coordination non-S disjunctive subjects (simplex and coordinated, pre- and post-positioned) non-standard case in 'all + pronoun' NPs in coordination non-POSS possessives in *of*-genitives
80%–90%	non-S subject complements O-forms after comparative *like*, *as*, *than*
> 90%	non-S subject complements in coordination, with and without V non-S overt subject pronouns in adverbial participles non-S subjects/ subject complements, no V O-form modifiers in gerund-participles non-S disjunctive subjects, pre-positioned

Based on the insight that all of the investigated phenomena are gradient phenomena, several categorical claims were refuted in the analyses, including the categorical rejection of: gendered reference for mass referents (7.2.2, (104): *Cause if he 's all mixed up, then you got to bag en* (cowfeed)); S-forms in prepositional complement function (8.2.1, (121): *he never interfered with I*); semantic identity of *of*- and *s*-genitives (13.1, (509): *we used to do their backs [...] we did do the backs of 'em*); and *of*-genitives with non-relational head nouns (13.3, (529): *the gates of them*). At the same time, however, claims predicting the categorical use of non-standard forms were also refuted.

16.3 Case assignment

In the individual analyses, different systematic tendencies could be identified and translated into preferential hierarchies. Some of the most important feature-specific findings are summarised in the Context hierarchy of non-standard object forms (8.3.6, (183)), the Context hierarchy of non-standard demonstratives (15.6, (620)), and the projection of gendered pronouns onto the Animacy Hierarchy (7.2.3, Figure 7.1). However, it appears that the strongest overarching link between the micro-level variation in specific syntactic functions and the macro-level variation of pronoun forms in the corpus is the great functional versatility of O-forms, as summarised in the Functional Diversity Hierarchy in (40) and shown in the function matrices in Tables C1 and C2.

All in all, the results do not go so far as to suggest the complete abolition of pronoun agreement in English, or the absence of structural case assignment. Although it can not be ruled out that English could, in the far future, develop into a different typologically attested language type (neutral, no case distinctions), there is enough evidence still that certain syntactic positions require morphological realisations in line with a nominative–accusative system. Object forms, for example, clearly dominate in direct object position, and locally-bound reflexives are usually overtly marked as self-forms. The most stable position for S-forms is as subject before a tensed verb. Currently, structural case assignment is under no threat in this position. If pronoun agreement, hypothetically speaking, were ever to be abolished in favour of an O-form paradigm, the 'subject before tensed V' position could well become the last S-form resort.[173]

173 In the Noun Phrase Accessibility Hierarchy of Keenan and Comrie (1977), the subject position is the most frequently relativised grammatical function. It is the entrance gate for innovations and the last resort for relic forms. Harris (1981: 19) predicts that "*I* will in due course be reserved solely and exclusively to contexts where it is directly bound to a main or auxiliary verb form within a finite verb phrase as its subject (exactly like *je* in French); in all other contexts, *me* will be appropriate (cf. French *moi*)."

In addition, certain pronoun uses could be argued to have inherent case. One example are the benefactive objects discussed in 12.3, which can be compared to benefactive arguments with theta-related dative case in languages like Icelandic (cf. Woolford 2006: 123). In other uses, pronoun choice was shown to be closely linked to specific lexical items. This type of correlation, which is not to be confused with lexical case assignment, was illustrated by collocational effects on genitive choice in chapter 13.4 and collocational effects on demonstrative determiner choice in chapter 15.5.

Among the numerous grammatical functions investigated in the empirical part, there are some syntactic environments which stand out with respect to two very specific properties. Firstly, these environments exhibit an above-average degree of variation in pronoun case; secondly, the quantitative distribution in these environments points towards the use of default case, which is usually object case. In terms of markedness, "the opposition between two or more categories is suppressed, and it is the unmarked member which appears" (Greenberg 1966: 29) – the unmarked variant being O-form pronouns. In the different analysis chapters, such environments were repeatedly referred to as default case environments, following the approach of Schütze (2001). It is here where the discrepancy between empirical variation, on the one hand, and prescriptive complementarity in case assignment, on the other hand, becomes particularly noticeable, up to a point where it seriously challenges studies which refuse to consider variation.

The use of default case has been attributed to ambiguity in case assignment which goes hand in hand with linguistic uncertainty and hypercorrection. The following syntactic functions are affected most strongly:

- subjects in sentences with V-elision (ch. 8.3.4);
- disjunctive subjects (ch. 12.5);
- overt subjects in adverbial participles (ch. 10.5);
- ECM and for–to constructions (ch. 8.2.2);
- subject complements (ch. 8.3.1);
- qualified pronouns (ch. 11);
- pronouns in comparative PPs (ch. 9.2).

In addition, coordination – the "great bugbear of prescriptivists" (Denison 1998: 109) – is renowned for causing ambiguity in case assignment (for instance, Quirk et al. 1985;

Emonds 1986; Parker, Riley and Meyer 1988; Dalrymple and Kaplan 2000; Kortmann and Szmrecsanyi 2004; De Vries 2005; Grano 2006).[174] The positive impact of coordination on non-standard case was confirmed in various places (e.g., 8.3.2, 10.3.1, 12.5.3).

Based on the results of this study, the concept of default case requires specification in two respects. On a qualitative level, it was shown that variation in case is not restricted to subject–object alternation, since it also involves the use of independent self-forms as a third pronominal category or formal variant (see pronoun exchange, ch. 8). In default environments, we usually find both S-forms and self-forms as less frequent variants. The exact reasons for the appearance of these secondary forms remain to be investigated. Emphasis alone is not intuitively satisfying.

The second specification concerns the quantitative distribution of morphological variants in default environments. While it is uncommon in the literature to include concrete frequencies in the description of default case, the results obtained for different syntactic functions in the current study lie so clearly within the same range that frequency itself can be regarded as an integral part of default use. The results show that O-form proportions can be expected to be > 60% (mostly 70%–80%),[175] and that the remaining occurrences are filled by either S-forms or independent self-forms, often at equal amounts. Considering this striking regularity, it is conceivable that information acquired during language acquisition not only includes the syntactic environments that allow for variation in case, but also the probability with which the different case forms can be expected to appear (along the same line of thought, cf. Adger 2006: 509).

174 See Johannessen (1998) for a typology of coordination.

175 Consider the respective sections in Part II of this study. For example, in disjunctive subjects: simplex 60%, coord. 60.5%; subjects with V elision: simplex 79.5%, coord. 77.5%; subject complements: simplex 83%, coord. 70.9%, qualified subject NPs (with qualifier noun or numeral) 69.9%, qualified object NPs 90%; comparative pronominal PPs 89.3% (the latter do not form part of Schütze's default environments).

Chapter 17
Practical and theoretical implications

What are the practical and theoretical implications of the obtained results from a wider perspective? How can the grammatical variation observed for individual phenomena be translated into more general patterns? And what are the connections between these patterns and basic linguistic principles, such as avoidance of ambiguity, or iconicity and economy?

17.1 Prioratisation of morphosyntactic categories

The phenomena investigated in this study fall into four groups, depending on the morphosyntactic category affected by variation in pronoun choice: person, number, gender and case. Based on the overall results, a simple but clear prioratisation emerges which is summarised in the Hierarchy of morphosyntactic categories in (621). It shows the relative amount of variation found in each category, reflecting the overall importance of each category for the correct processing of personal pronouns in English.

(621) Hierarchy of morphosyntactic categories for personal pronoun processing

person > number > gender > case

The presented results, in combination with concurrent observations from the literature, allow the assumption that (621) can be expected to apply to other data of spoken English and probably in writing, too. This means that we can expect to find fewer phenomena with variation in person and number (compare ch. 6), as compared to case (ch. 8–15). From a diachronic perspective, the strong potential for variation in case in Modern English presents a natural continuation of the progressive levelling of case distinctions in the language, the only present-day survivors being those pronouns that have retained distinctive subject and object forms. The status of gender in (621) corresponds to the current findings. In other sets of data – present and future – it will depend on the persistence, or disappearance, of the only gender-related phenomenon, gendered pronouns.

From a wider perspective, it is interesting to see that (621) corresponds to different

proposals in the linguistic literature of the last two or three decades, which in the modelling of morphosyntactic feature hierarchies have attributed different phi-features a different grammatical status (e.g. Noyer 1992; Harley and Ritter 2002a; Carminati 2005).[176]

At the same time, the observed prioratisation in (621) correlates with earlier observations from sociolinguistic and language acquisition studies which attest the early acquisition of person distinctions as compared to the relatively late acquisition of case distinctions among English-speaking children (cf. Emonds 1986).[177] During his investigations of the verbal deprivation hypothesis, for example, Labov (1970: 172) found that the use of non-standard case among African-American children in urban ghettos (e.g., *Me got juice*) showed "only that [the child had] not learned the formal rules for the use of the subjective form *I* and oblique form *me*." Labov's description continues as follows:

> We have in fact encountered many children who do not have these formal rules in order at the ages of four, five, six, or even eight. It is extremely difficult to construct a minimal pair to show that the difference between *he* and *him*, or *she* and *her* carries cognitive meaning. In almost every case, it is the context which tells us who is the agent and who is acted upon. [...] it is evident that the children concerned do understand the difference in meaning between *she* and *her* when another person uses the forms. All that remains, then, is that the children themselves do not use the two forms. (Labov 1970: 172–173)

Furthermore, the hierarchy in (621) correlates with some major typological trends in the realisation of pronominal features described in (i) to (iv):

176 Compare Noyer's Universal Feature Hierarchy: person > number > gender > class. In his typological study, Noyer (1992) elaborates a theory of morphological filters which function as co-occurrence restrictions on specific combinations of morphosyntactic features. These filters are automatically assumed by children in early language acquisition unless positive evidence is obtained that a given category remains distinct (p. 9). Whenever a combination of morphosyntactic features violates a prevailing filter, this causes the deletion of morphological features active in the language (p. 44). The feature hierarchy determines the direction of this deletion: the higher categories are the ones that resist deletion most strongly. With respect to a filter which bars gender distinction in 1SG verbs in Arabic, for example, this means that in 1SG verbs the value of the gender feature is deleted in the morphological derivation while the person feature is preserved. As a result, we obtain a 1SG verb (ʔ-aktub-u 'I write') which is distinct in person from the corresponding 3SGf verb (t-aktub-u 'she writes'), following the hierarchy person > gender. Noyer's feature hierarchy springs from an apparently unrelated field of study but attributes the same relative importance to the different morphological categories as (621).

177 Similar in Dutch. Tieken-Boon van Ostade (1994: 226), for example, describes the use of O-form subjects and S-form prepositional complements among Dutch children as follows: "It is only after continued exposure to standard adult usage and as the result of persistent correction that the grammatically correct forms begin to appear more regularly. This seems true for gender distinctions in pronominals, too. The set of pronominals in which case, number and gender are distinguished therefore has to be actively acquired as part of the process of first language acquisition."

(i) Person has the status of a typologically omnipresent feature.

According to Greenberg's Universal 42, "all languages have pronominal categories involving at least three persons and two numbers" (1963: 90); compare Siewierska (2004); Trudgill (2009); Mühlhäusler and Harré (1990), for exceptions; Forchheimer (1953), for a typology of personal pronoun systems; Ingram (1978), for a categorisation of person systems on the basis of abstract semantic and deictic information.

(ii) The overt marking of number distinctions is widespread, with some exceptions.

Compare Corbett (2000); Greenberg's Universal 36 states that "[i]f a language has the category of gender, it always has the category of number" (1963: 90); in the feature geometry of Harley and Ritter (2002a), languages without number distinctions lack an 'individuation' node.

(iii) Gender varies more widely than either person or number.

According to Harley and Ritter (2002a: 514), "gender (or class) features vary more widely in the worlds' languages than either person or number. For example, while all languages seem to have at most four persons and four numbers, the set of gender/ class systems seems much less constrained. Some languages have no gender marking whatsoever [...]. Other languages have two or three genders. The limiting case is probably presented by the Bantu languages, which have upwards of ten distinct genders or classes of nouns (see Corbett 1991)." In her contribution to WALS, Siewierska (2005) finds that (i) about 30% of the languages in her sample (378 languages) mark gender in independent pronouns (predominantly in Africa, Eurasia and Europe), (ii) most gender contrasts in personal pronouns are sex-based, (iii) gender oppositions are characteristic of the third rather than the first or second person (similar Mühlhäusler and Harré 1990), and (iv) gender is more typical of the singular than non-singular. The corresponding WALS map is shown in Map D11. Note that in Ido, an auxiliary language specifically designed as a universal alternative to English, the gender-neutral pronoun *lu* can mean 'he', 'she' or 'it'.

(iv) Case varies most widely in the world's languages.

The overt morphological marking of case distinctions has been described as "the most highly variable grammatical phenomenon associated with the grammatical relations of 'subject' and 'object' " (Croft 1990: 152). Compare Blake (2001), for a global perspective on case systems and case marking; Sapir (1921), for the levelling of the seven cases in Indo-European nouns, as well as the instability of the *who/ whom* distinction in relative pronouns; Ingram (1978), for a case-based classification of language types.

Note that some constructed international languages do not mark case distinctions. For instance, Otto Jespersen's Novial relies on word order: *me observa vu* 'I observe you' vs. *vu observa me* 'you observe me' (http://interlanguages.net/AIL.html). However, Jespersen (1928, Part II, Case) conceded: "Still, many people would prefer a mark of the accusative to be used in those rare cases in which ambiguity might be feared, and the best ending seems to be *-m*."

<u>Regarding reflexivity</u>: in many languages reflexivity stays unmarked. While the concept itself appears to be universal, not all languages have special reflexive pronouns. Some languages "simply use personal pronouns in their place" (Kiparsky 2002: 203); others only have reflexive markers in the 3rd person. The use of polysemous reflexive pronouns is well-known from other languages: Asian *ziji*, *zibun* and *caki*, for example, are devoid of φ-features. These pronouns can be used as subjects or objects (direct and indirect) and their antecedents are usually co-arguments of the predicate of the matrix clause, similar to so-called long-distance reflexives in English (cf. Huang 2000: 191).

The obvious explanation for (621), as well as for the overarching typological trends, lies in the role that the different morphosyntactic categories play in the expression and interpretation of referential and coreferential relations. Besides its descriptive and predictive potential, the Hierarchy of morphosyntactic categories has fundamental implications for a conjunctive analysis of pronominal expressions. In English, the specification of person and number features is considered vital for a correct resolution of both independent and anaphoric pronouns, whereas the importance of gender and case features is radically reduced by the availability of supplementary information in the discourse. For the correct interpretation of personal pronouns in actual conversations such as the interviews used in this study, overt information is usually required, first and foremost, about the speech act role of the respective antecedent or extra-linguistic referent. Person, therefore, turns out to be the most robust category of all, given its vital importance for the marking of speech act roles: speaker (1) vs. hearer (2) vs. non-participant (3) (in Harley and Ritter 2002a, [±participant] and [±addressee]).

Case, on the other hand, represents the least robust category. In English, its original purpose of establishing syntactic dependencies has become redundant to a large extent due to the fixed word order. In discourse, syntactic dependencies can also be inferred from the verb's semantics (cf. König and Siemund 1997), mechanisms for the track-keeping of referents (cf. Bosch 1985), the information structure, prosodic elements and the overall context. According to Greenbaum and Quirk (1990: 108), all pronouns have one thing in common: "their referential meaning is determined purely by the grammar of English and the linguistic or situational context in which they occur." This is why, in many cases, overt pronoun–antecedent or pronoun–referent agreement becomes redundant. In a sentence such as *I saw me*, for example, the use of a 1SG pronominal object plainly satisfies the morphological requirements for a reflexive interpretation. The same may or may not apply to the same sentence in the 3SG, *He saw him*, depending on the contextual information available to the discourse participants. Against initial expectations, the present study returned no significant differences in the treatment of first/ second vs. third person pronouns regarding variation in case (e.g., pronoun exchange in ch. 8; snake sentences in 9.3).

17.2 Determinants of pronoun variation: a combined approach

The present study is an example for how grammatical variation can be captured by a combined approach. In the corpus-driven inquiry, the data were approached in an un-preconceived fashion, the first goal being the objective quantification of non-standard phenomena. Subsequently, each phenomenon was subjected to a variety of investigative thoughts and ideas, exploring linguistic approaches that offer intuitively adequate explanations for the empirical observations.

Overall, the obtained results point towards a linguistic system which follows the standard agreement rules in part. With respect to case – the morphosyntactic category with the highest degree of variation – structural case assignment can be argued to operate in many places. However, the standard system frequently interacts with other factors, both linguistic and extra-linguistic, which motivate and induce formal variation in their own systematic ways. In the actual data, this interaction is reflected in numerous utterances where the pronoun's syntactic function can no longer be deduced from the pronoun's form, but, instead, has to be inferred from the pronoun's position in the sentence and the overall context. The inference mechanisms obviously work: nowhere in the investigated conversations could missing or faulty pronoun agreement be shown to affect the communication process in any negative way.

The importance of both word order and contextual information has been notoriously underestimated in studies insisting on the importance of correct case assignment and holding on to the ideas of isomorphism and complementary distribution (e.g., the much disputed binding constraints on pronouns and anaphors proposed by Chomsky 1981).[178] However, these notions are loosening up even in more recent generative studies. In his study of anaphora, for example, Safir (2004: 6) recognises that failures of complementarity "arise in contexts where independent principles determine that no interpretation is available, or where thematic assignments are not comparable, or where forms that normally compete do not compete for principled reasons (and hence both appear)."[179] In a different field, Wierzbicka (1996: 21) has stressed the importance of cognitive factors, noting that "a child first makes sense of situations and human intentions and then of what is said. This means that language is not independent of the rest of cognition."[180]

178 The 'no variation' principle lies at the very heart of standardisation (cf. Stein 1997; Milroy and Milroy 1993).

179 Different generative explanations for non-standard pronoun usage by Emonds (1986) were mentioned in the analyses. Compare Branco and Marrafa (1997) for a modification of binding principles; Kiss (2001) for a comparison of binding principle A in English and German.

180 For neurolinguistic studies on sentence-internal parsing see Bornkessel et al. (2003a) and (2003b). For

A fixed word order and the availability of contextual information are fundamental prerequisites which facilitate grammatical variation in the first place.[181] In addition, a variety of determinants can be identified which motivate and influence variation in specific syntactic contexts. In the empirical analyses in Part II of this study, these determinants were repeatedly discussed in connection with the different phenomena. We will now briefly reconsider them in retrospective.[182] Figure 17.1 assembles those determinants which emerged as the most influential in the individual analyses, i.e. those determinants with the greatest explanatory potential for the non-standard uses found in the corpus.

determinants of pronoun choice

syntactic		extra-syntactic
structural case	specific constructions/ missing	> diatopic variation
(standard case)	case assignment	> speaker variables
	favouring non-standard/	> 'Avoid Ambiguity'
	default case	> emphasis
		> hypercorrection
		> avoidance
		> frequency effects
		> collocational effects
		> processing effort
		> pragmatic aspects

Figure 17.1: Determinants of pronoun choice in English

The role of structural and default case assignment was discussed in 16.3, so that the rest of this section can be dedicated to the more complex array of extra-syntactic determinants depicted in the right column in Figure 17.1. Each of these 'performance mechanisms' (Adger 2006: 506; Dressler 1985: 322) contributed, at some point, to the interpretation and explanation of phenomena analysed in the empirical part of this study (often in conjunction

discourse grammar approaches see Cantrall (1974); Kuno (1987); Zribi-Hertz (1989); Parker, Riley and Meyer (1990); Baker (1995).

181 The historical development of the English pronominal paradigm is not to be considered as a determinant per se, even if it paved the way for the variation found today. Historical precursors were considered in all analysis chapters.

182 Changes in the conditioning factors themselves are ultimately reflected in gradual shifts in the distribution of linguistic variants (cf. Labov 1966: 16).

with other factors). It is, however, important to note that the resulting variation is perceived as deterministic only to a certain extent, due to the non-categorical influence of the determinants and the additional impact of inter- and intra-personal variability.

Diatopic variation

As was shown in 16.1, most non-standard uses of personal pronouns and self-forms represent general features of spoken English. Geographic restrictions are rare. The only two distinctly regional phenomena in the corpus involve non-standard S-forms: in object functions (ch. 8) and as demonstrative *they* (ch. 15). Both of these phenomena are typical for the Southwest of England.

Speaker variables

In different chapters, attention was drawn to inter- and intra-speaker variation, for example in connection with the identification of speaker distribution patterns (5.2) and the identification of outliers (7.2.8). While most speakers in the corpus use at least some non-standard pronouns, internal variation prevails over categorical usage. Regarding the role of speaker sex, it is worth mentioning that no basic difference could be established between male and female linguistic behaviour in the use of non-standard pronouns. The FRED results are similar in this respect to the findings of Tieken-Boon van Ostade (1994), but different from other studies such as Trudgill (1974), Milroy and Milroy (1993) or Nevalainen (2000). In the present study, apparent differences in the use of gendered pronouns were attributed to differences in the speakers' personal involvement or attachment to certain types of referents, rather than to the speaker sex itself (see 7.2.7).

'Avoid Ambiguity'

We can of course ask, if English can do without overt case distinctions in nouns, why shouldn't it be able to cope without case distinctions in pronouns? The fact is that English is already coping with a high amount of variation in this area. Therefore, the better question is, is variation ever actively avoided? In some utterances, the answer to this question appears to be yes, and the explanation appears to lie in the preservation of intelligibility, or avoidance of ambiguity.

It is not a new insight that disambiguation plays a role in pronoun choice, occasionally leading to the development of new forms and even new pronominal paradigms (for

instance, the emergence of Old English forms in *-self*). Variation, especially in pronoun case, is generally deemed less problematic if it does not cause ambiguity (even in prescriptive grammars, cf. Vallins 1952: 73). However, avoidance of ambiguity is not usually regarded as an integral part of case assignment, which is what was proposed in section 9.3.3:

(266, rep.) 'Avoid Ambiguity' Principle
Use unambiguous case if disambiguation is required;
if no disambiguation is required mark case as usual.

The results of this study strongly suggest that, firstly, there is a reduced need for disambiguation in oral conversation which allows for greater functional variability of the investigated word forms, and, secondly, the 'Avoid ambiguity' principle only applies if there is a serious risk of ambiguity threatening the communication process. In a language like English, which leaves considerable room for ambiguous pronoun usage in general,[183] the ambiguity threshold is relatively high. The threshold is also certainly higher than suggested in the typical minimal-pair examples used to illustrate disambiguation by complementary case marking in the literature.

With respect to other linguistic principles, 'Avoid Ambiguity' counteracts, simultaneously, (i) the structuralist principle of 'one form – one meaning' in those instances where disambiguation is not required, and (ii) morphological economy in those utterances where disambiguation is required, calling for distinct case forms (compare Zipf's Principle of Least Effort and its antithetical Force of Diversification, cf. Zipf 1949). From a theoretical perspective, the interaction between these three opposing principles presents an alternative to the universal principle of Coargument Disjoint Reference in anaphor resolution (cf. Farmer and Harnish 1987; Kiparsky 2002). The latter requires that arguments of the same predicate do not overlap in reference unless they are specified as non-obviative.[184]

183 Consider, for example, the lack of gender specification in articles, nouns and possessives (*my friend*/ *the friend*); possessive *your*/ *yours* (SG and PL), *her* (possessive and object form); the missing distinction between anaphoric and non-anaphoric possessives, which can lead to ambiguity in sentences like *John*$_i$ *gave Ian*$_j$ *his*$_{i/j}$ *book*; or the use of third person pronouns for endophoric speech act participants as well as third person referents in sentences like *John*$_x$ *told me that he*$_{x,y}$ *had killed the tiger*. As was mentioned in 17.1, no significant difference was found in the treatment of first/ second vs. third person pronouns regarding variation in case. In the data, the Avoid Ambiguity principle does not seem to apply to third person cases any more than first or second person cases. The picture is slightly different in languages which allow O-form reflexives everywhere but in the third person, e.g. Dutch *Ik waste me*/ *Jij waste je*/ **Hij waste hem* 'I washed me'/ 'You washed you'/ *'He washed him' (cf. Gelderen 2000b; Faltz 1985).

Regarding (i), it is interesting to note that the 'one form – one meaning' claim has also been abolished in lexical-functional grammar. In the more recent literature, two rivalling approaches have tried to explain the variability found in the surface realisation of morphological features by assuming that the underlying features are either underspecified or indeterminate (most studies focus on N–V agreement, phi-features not including case). The underspecification approach argues for a minimal specification of lexical items, leaving room for underspecified features to be realised in different ways. For example, variability in the form of the past tense copula *be* after *you* in some dialects (*you was*/ *were*) can be explained by an analysis that assumes the following feature correlations: (a) [+singular] *was*, (b) [–singular] *were*, (c) [+participant] *was*, (d) [+addressee] *were*, (e) [–addressee] *was*; thus allowing for *be* after *you* ([+singular], [+participant] and [+addressee]) to be realised as either *was* or *were* (cf. Adger 2006). In comparison, the feature indeterminacy approach proposes to represent "indeterminate values by sets containing the alternative possibilities" (Dalrymple and Kaplan 2000: 796). For example, a particular pronoun which needs to fulfill different case requirements in one sentence can have a case feature set, as seen in the German sentence *Ich habe gegessen, was* {NOM, ACC} *übrig war* 'I ate what was left' (ibid., p. 764). The two lexical-functional approaches explain why variability in morphological realisations exists in the first place. However, a great variety of factors can still influence the final step to phonological realisation, including the extra-syntactic determinants listed in Figure 17.1.

Emphasis

In various chapters of Part II, emphasis was mentioned as a possible explanation for non-standard uses, including independent self-forms and other sub-features of PE, non-standard case in comparative constructions and coordinated NPs, overt pronominal subjects in adverbial participles, overt *you* imperatives, and pronoun doubling in disjunctives and resumptives (where the doubling itself is iconic of intensity). In various syntactic functions, S-forms and self-forms appear to be linked more closely to emphasis than O-forms, an observation which is readily explainable by the latter's default status. In discussions of PE, for example, S-form objects are often associated with emphasis whereas O-form subjects are regarded as the unemphatic option (cf. Wright 1905: 270; Kruisinga 1905: 35–36). The advantage of the present study is that recordings were available for

184 Similar to the so-called Role Conflict Hypothesis which describes the use of self-forms to mark unexpected coreference (cf. Stern 2004).

most transcripts, showing that non-standard pronoun choices and word stress are only correlated occasionally. The nature and extent of this correlation is certainly worth investigating in more detail, ideally in similar combined databases (text and sound).

Hypercorrection

Cheshire and Stein (1997b: 7) have noted, with reference to van Marle in the same volume, that "we can observe natural processes of linguistic change in dialects, though not in the standardized varieties [...] In the standardized varieties the normative pressures associated with codification tend towards a 'logical' elimination of variation, which prevents these natural developments from surfacing." Hypercorrection is an interesting side effect of these normative pressures which probably comes into play more often than is generally acknowledged. This is partly due to its social aspects, i.e. the interpretation of certain linguistic expressions as indicators for social class and education. According to Emonds (1986: 116), hypercorrections sound prestigious to people since they are what people "unconsciously realize sets them apart from [normal usage]." Furthermore, it has been noted by Krug (1998: 301) that "it is not in a language's interest to lose its stylistic and social markers." One typical example for socially conditioned hypercorrection in pronouns is the continued promotion of *and I*, even in object function, which can be described as an example of 'frozen pragmatics' (cf. Levinson 1987: 420).[185]

Avoidance

Another side effect of normative pressures is avoidance. Uncertainty can be caused among speakers by discrepancies between standard rules and actual language use, ultimately leading to the avoidance of specific constructions (as described in 9.2; cf. Emonds 1986: 119, 'Avoid the construction' chapter). Both hypercorrection and avoidance have a long history as coping mechanisms, even if they are often ignored as determinants of variation. In her historical data, for example, Tieken-Boon van Ostade (1994: 231) observed that 18th century writers already tended to choose either *myself* or *self* in coordinated constructions with more than two legs, probably in order to avoid choosing between *me* and *I*: "the higher the number of coordinated nominals, the more the speaker or writer tends to lose control over the syntax of the sentence as a whole, and the sooner he or she

185 Compare Weiß (2004: 653–654): "It is known to be a naive illusion that prescriptive knowledge can be distinguished from linguistic intuitions proper by normal speakers in any case [...] There is always the possibility that explicitly learned conscious knowledge becomes unconscious implicit knowledge or even that such rules are secondarily and implicitly learned (e.g., via imitation)."

will deploy the avoidance strategy at their disposal in this context." From this perspective, the use of self-forms in coordination is epiphenomenal, since it is attributed to the avoidance behaviour triggered by coordination rather than coordination itself.

The influence of hypercorrection, uncertainty and avoidance is of course extremely difficult to determine, especially if no direct questioning of the speakers is possible. In the present study, therefore, these determinants had to remain unquantified.

Frequency effects

Frequency itself was not considered as a determinant. However, the numerical results for non-standard question tags in section 6.4 strongly suggest frequency as a cognitive motivator for lexicalisation, especially regarding the reduced tags *innit* and *wunnit*. A different phenomenon is the use of *and I* in coordinated NPs with object or prepositional complement function (see 8.3.1 and 14.3). *And I* appears to have developed from a socially motivated hypercorrection phenomenon to a frozen form rivalling *and me* (cf. Redfern 1994: 189; Honey 1995: 7). Similar phenomena can become self-perpetuating: the perception of *and I* as a linguistic unit results in a high string frequency, and the high string frequency reinforces the perception of the string as a linguistic unit. In the words of Krug (1998):

> Frequency of word-form sequences in natural spoken language appears to be the dominant factor that determines the intuition about how closely connected two adjacent words are. The degree of closeness (or connectedness, cohesion, bondedness, bonding) ranges from frozen forms via collocations to chance combinations. (Krug 1998: 301)

An interesting observation awaiting further elaboration was made after completing the analyses. It was not included in Part II but deserves to be mentioned. The analyses show a general correlation between higher frequencies of individual syntactic functions (subject functions) and standard pronoun occurrences. This correlation is visualised in the linear regression model in Figure C1 (Appendix). Based on preliminary results, it appears that the most frequent syntactic functions, i.e. those functions which speakers are exposed to most often, have the lowest non-standard frequencies. This includes, for instance, subjects before finite verbs and subject complements. At first view, the observation clashes with the usual relation between conservatism and frequency, which assumes that traditional (non-standard) phenomena survive in frequent grammatical relations (cf. Krug 2003: 18).

Instead, the observed tendency is more in line with the usual path of morphological innovations, which are assumed to establish themselves first in the less salient syntactic roles and spread from there to the more prominent roles (cf. Ihalainen 1991: 105). Since the investigated phenomena do not represent dialectal innovations but have all been in the language for a considerable amount of time (as shown in 3.1), we are faced with a puzzle: traditional phenomena which do not exhibit the typical distribution of traditional phenomena but the typical distribution of linguistic innovations. However, the results are compatible with the effects of uncertainty described above. A logical explanation on cognitive grounds suggests that speakers feel more secure in their linguistic choices regarding syntactic constructions they use of hear more frequently. In less frequent constructions, on the contrary, uncertainty is more likely to creep in, causing inconsistency in the linguistic behaviour, and thus naturally leading to the production of non-standard occurrences. Elsewhere, Emonds (1986: 114) has described the connection between hypercorrection and missing internalisation, saying that "[w]hen speakers of a language internalize a construction in that language, they do not, after the period of acquisition itself, persistently overgeneralize the construction into paradigms where it is not acceptable." Hymes (1997: 13) has elaborated on the link between occurrence and appropriateness in an extended notion of speaker competence: "members of a speech community are aware of commonness, rarity, previous occurrence or novelty, of many features of speech, and [...] this knowledge enters into their definitions and evaluations of ways of speaking." Accordingly, frequently observed uses can be expected to be judged more acceptable, and acceptable uses can be expected to be used more frequently. An explanation for the quantitative tendency observed in FRED will probably move along these lines.

Collocational effects

Collocational effects are often ignored in dialectological studies and are, therefore, rarely approached systematically. They get mentioned indirectly. For instance, *them days* is the typical example for non-standard demonstratives found in the literature. In this study, I have tried to stress the importance of the collocational environment as a potential determinant of pronoun variation. The applicability of collocational analyses to dialect data was demonstrated in chapters 13 (variation between genitive structures) and 15 (collocational preferences of different demonstrative determiners). It should also be mentioned that collocationally motivated non-standard uses represent exceptions to the linear regression described under frequency effects above. Take demonstrative *them* as an

example. Demonstrative determiners are highly frequent in the corpus. Based on the general tendency observed in other functions and constructions, we would expect to get a low non-standard frequency in this function. However, the opposite is the case due to the strong attraction of demonstrative *they* and, especially, *them* to the time noun *days* (see Table 15.2).

Processing effort

Several phenomena studied in the empirical part include so-called pleonastic pronouns, i.e. overt realisations of pronouns judged unnecessary, unacceptable or illogical by some grammarians. Besides their special pragmatic uses (see 17.3), such additional pronouns can also be argued to serve the reduction of processing effort or cognitive load. It is due to the linearity of language and the velocity of spontaneous oral conversation that the overt expression of a pronoun instead of zero can be used to enhance the understandability of an utterance. In so-called dangling participles, for instance, the use of an overt pronoun provides a straightforward means of disambiguation (see 10.5).

In the literature, a significant reduction of processing effort has, for example, been attributed to resumptive pronouns (cf. Huang 2000: 254). From the standpoint of natural morphology, it can be argued that resumptives and other pleonastics achieve an increase of semiotic transparency, the transparency being "optimal if nothing interferes between meaning and phonetic form, i.e. if no obstacle increases the receiver's processing task." (Dressler 1985: 323). While it is generally agreed that zeros refer to highly accessible mental entities in the discourse (cf. Ariel 1990: 106), it stands to reason that the use of an overt pronoun instead of zero reduces the effort needed to identify, for example, a verbal subject (especially in a non-pro-drop language like English).

Pragmatic aspects

Among the most influential determinants in Figure 17.1 are the different pragmatic and discourse-pragmatic aspects detected in this study, including the speaker's general need for expressivity, the speaker viewpoint, discourse prominence, and politeness and modesty. Pronoun pragmatics is a highly interesting and complex field; the next section brings together the multiple findings mentioned throughout this study.

17.3 Pronoun pragmatics

In this last section, attention will be drawn once more to pragmatic aspects. The pragmatic uses of non-standard pronouns are manifold. References and examples can be found in almost every chapter in this study. In summary, the present investigation supports other studies which acknowledge that a rule-based sentence grammar and a purely formalist approach do not suffice to describe the functional scope of personal pronouns and self-forms in English. Many of the phenomena discussed in the empirical part are best understood from a pragmatic perspective. At the same time, it should be kept in mind that pragmatic uses, similar to the non-pragmatic influences discussed before, are non-categorical. Pragmatic aspects can, but do not have to, be marked on pronouns. To a certain extent, this limits the explanatory potential of concepts like subject of consciousness (Zribi-Hertz 1989) or relative discourse prominence (Baker 1995) to those instances where pragmatic marking is detectable.

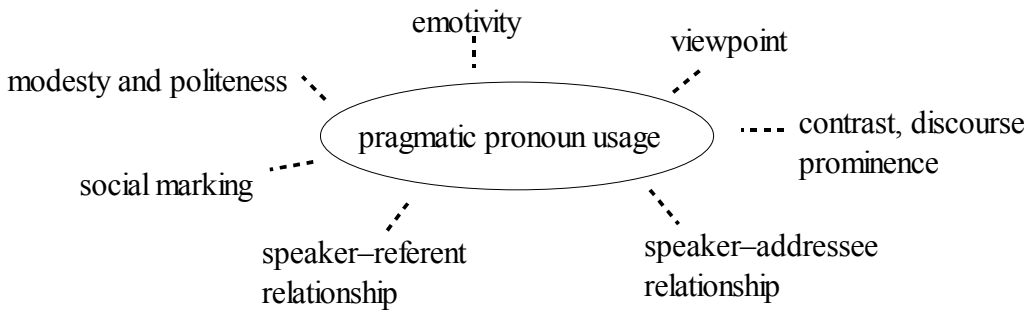


Figure 17.2: Pragmatic pronoun usage

The lowest common denominator of the different aspects depicted in Figure 17.2 is the speakers' need for expressivity:

> Speakers tend to develop different meanings for the competing forms that exist in the dialects, whereas the ideology of the standard forces us to [...] eliminate variation. The standardized variety therefore forgoes a significant potential for creating expressive meaning in favour of a unitary and purportedly more 'logical' solution. (Cheshire and Stein 1997b: 7)

From a structural point of view, it is the relative redundancy of case marking in a fixed word order language like English which facilitates a functional re-attribution of case features (one of the possible developments of functionally empty categories, cf. Labov 1966: 20). In addition, overt pronouns can be used instead of standard zero to convey meanings on top of the propositional content of a sentence, including the speaker's attitude towards the sentence proposition (see the Working principle of ethic datives and corresponding constructions in (464)). We observe a circular relationship: morphosyntactic variability allows for pragmatic pronoun usage, and pragmatic pronoun usage results in morphosyntactic variation.

A detailed analysis of pronoun pragmatics would exceed the scope of this study. Nevertheless, pragmatic aspects were shown to be vitally important for the interpretation of the phenomena under discussion, which is why at least a brief recapitulation is in order to depict the great range of expressive possibilities.

To begin with, pronoun choice can be used to express politeness and modesty. Speakers can use specific pronominal expressions to signal reservedness when referring to themselves, or to pay deference to the addressee. Typical examples are the use of independent self-forms as described in 8.5,[186] the use of plural *us* with singular reference (6.1.1), and the use of 2PL *you all* (11.2). Descriptions of polite self-forms usually mention 1SG *myself* (e.g., Tieken-Boon van Ostade 1994: 230)[187] and 2SG *yourself* (ibid., 'inverted modesty'; cf. Baker 1995: 91), but polite meanings have also been mentioned in connection with third-person examples (cf. Hernández 2002: 274). Filppula (1999: 34), for example, mentions that "it may well be that in [Irish English], too, speakers consider it more polite to refer to a third party, who is usually not present, by means of a reflexive pronoun instead of the corresponding personal pronoun." Prescriptive language guides, of course, cast a critical eye on these independent self-forms:

> Some persons with false modesty would rather speak of themselves as *myself* than as *I*, *me*. They deliver such incorrect sentences as 'The chief sent a Christmas card to my wife and *myself*.' [...] Don't be timid with personal pronouns; as a matter of fact, *I* and *me* are shorter, less emphatic than *myself*. (Stratton 1949: 155)

186 Note that I have argued against the view that self-forms in argument positions are intensifiers attached to covert head nouns.

187 "The non-reflexive -*self* pronominals are usually found with the first person singular, the author him/herself, and it seems only natural, given polite human behaviour, that they might want to make their own presence, say in a list of names, somewhat less obtrusive. The use of *myself* may therefore be regarded as a kind of modesty device. One result of the use of *myself* rather than *I* is that the stress falls on the second syllable, thus deflecting attention from the 'I'."

In coordination, moving a 1SG form to the final position ('NP *and me*' instead of '*me and* NP') is also associated with polite linguistic behaviour (cf. Kuno 1987: 233, 'modesty principle'; compare ch. 8.5). The use of *and I* in object function, however, is considered polite by some people, and stilted by others (cf. Angermeyer and Singler 2003; Honey 1995; Householder 1987; Redfern 1994; compare the notes on hypercorrection above).

Gendered pronouns are a good example for the pragmatic use of gender features. Many examples discussed in this study point towards a pragmatic system which revolves around the speaker viewpoint. Gendered forms can be used for the topicalisation of the focal referent, but mainly they express the speaker's personal involvement and emotional attachment to the referent (compare 7.2.2).

Other examples for emotive language involve pleonastic pronouns as described in chapter 12. Here, it is the use of an overt pronoun instead of standard zero, and the position of the pronoun rather than its form, which carries the additional meaning. A certain degree of emotivity was attributed to overt *you* imperatives (12.1), non-standard benefactives (12.3), ethic PPs (12.4), disjunctive pronouns (12.5), and also the *it* in *walk it* (6.2).

Emotivity, of course, forms part of various other aspects which have a notable influence on pronoun choice.[188] Specific pronouns can serve to specify the speaker–addressee relationship, for example by encoding politeness (independent self-forms), closeness (singular *us*), or solidarity (sentences like *How are we feeling today?*, cf. Harley and Ritter 2002a: 507). Specific pronouns can also serve to specify the speaker–referent relationship, especially the speaker's emotional attachment. In the present study, this aspect was illustrated by the use of gendered pronouns and switches between gendered and non-gendered forms in 7.2.4 (in a slightly different function, consider the 'animation effect' of pleonastic reflexives in 12.2.3).

One of the most important discourse-pragmatic functions of non-standard pronouns is their use as contrasting devices, i.e. the demarcation of specific referents compared to other entities in the discourse. As was once pointed out by Jespersen (1949: 170–171), many uses that are spontaneously interpreted as emphatic are not about emphasis but about contrast. Contrast can of course be obtained by a variety of means, including word stress

188 On a more general level, a heightened degree of emotionality in the conversation can also affect internal variation, possibly leading to an increase of non-standard frequencies. Compare Visser (1963: 56), who states that "[a]ccording to Wright, The English Dialect Grammar §402 the personal pronoun is often repeated in the northern dialects in recriminatory talk, as 'thou great lout, thou.'." In one of the FRED transcripts, LAN_012, the transcriber noted "Much more strongly dialectal in some parts than in others, especially when informant gets excited about topic, and when recollecting school memories."

and sentence structure (e.g., disjunctives as in *Me, I...*).[189] Non-standard pronoun case forms have been observed to appear in utterances that imply a contrast or comparison between different entities. This includes coordination, comparative prepositions such as *like*, *than* or *as*, and exclusive expressions like *but* (cf. Zribi-Hertz 1989; Stern 2004). Frequently, the use of non-standard case in these contexts not only serves to reinforce the contrast but to spotlight the one entity with the highest relative contextual importance or discourse prominence (cf. Baker 1995).

Finally, pragmatic pronouns (usually independent self-forms) can be used to describe an internal point of view, i.e. "the point of view associated with a participant in the clause or sentence" (cf. Kemmer 1995: 67, 'viewpoint reflexives'). Cantrall's concept of viewpoint (1973; 1974) and related concepts such as logophoricity (Kuno 1987), subject of consciousness (Zribi-Hertz 1989) and subjectivity (Kemmer 1995; Stein 1997)[190] have been used in the literature to explain the use of independent self-forms which identify the respective referent "as the cognizer of the situation" (Stern 2004: 277). In the analyses in Part II, the influence of viewpoint was considered in connection with different phenomena, especially the so-called snake sentences where the use of logophoric self-forms is commonly expected to indicate a shift in perspective (see 9.3). Unfortunately, the corpus results showed that spontaneous speech differs fundamentally from the examples usually found in the literature (often minimal pairs). In the empirical examples, differences in viewpoint and different degrees of empathy can hardly ever be deduced from the pronoun's form.

In summary, it has to be acknowledged that variation in spontaneous speech is not, or not entirely, predictable by theories based on more formal data or illustrative minimal pairs. However, different discourse-pragmatic approaches do provide explanations for occurrences that can not be accounted for by case assignment rules (neither structural nor default) or any of the other determinants mentioned above. The most important conclusion, therefore, is that any account of pronoun variation in English is incomplete without considering pragmatic aspects, and that "pragmatics may no longer be treated as an 'epiphenomenon at best'" (Huang 2000: 331).

189 In pro-drop languages, the overt use of pronouns instead of zero suffices to signal emphasis or contrast. Consider Catalan *Tots els músics pensen que són irresistibles* vs. *Tots els músics pensen que ells són irresistibles* 'All musicians think that *they* are irresistible' (cf. Huang 2000: 233).

190 Compare Stein (1997), on the subjectivity of O-form subjects in tensed clauses such as *Her's a bright spark*; Brinton (1995), on the representation of non-speaker subjectivity in literary discourse.

Chapter 18
Conclusion and outlook

With the present study, I hope to have brought some clarity to the much discussed and multi-layered topic of non-standard pronouns in English. By using a corpus-driven method, the data were approached in an un-preconceived fashion which not only allowed to identify well-known frequent phenomena but also some rarer non-standard uses. It was thus possible to describe the full functional range of personal pronoun behaviour as it presents itself in the spontaneous conversations selected for this study, and to identify a number of dominant distributional tendencies and patterns of variation which may reappear in other varieties of English and other datasets. The most important overarching tendencies observable in FRED can be summarised as follows.

In terms of geographical distribution, the investigated pronoun phenomena show general supraregionality. They represent generalised features of spoken English rather than regional characteristics (see 16.1). In some cases, the empirical results were in conflict with previous accounts, necessitating a reformulation of traditional definitions, for example for pronoun exchange (8.6) and gendered pronouns (7.3).

Non-standard pronoun usage tends to be non-categorical, complementary distribution being the exception rather than the rule. Any appropriate framework must account for this variability. Variation in pronoun choice affects almost all syntactic functions in all dialect areas. The main difference between the individual phenomena lies in differing degrees of variation (gradient phenomena). Semantically equivalent or near-equivalent variants are neither entirely determined by the syntactic context and dialect, nor do they occur in free variation. Predictions about individual phenomena can, therefore, only have the form of distributional tendencies.

A breakdown of the different phenomena by relative frequencies showed that non-standard uses, i.e. uses contradicting the prescriptive rules of written standard English, cover an extremely wide range from 0.01% to almost 100%. At the same time, however, clustering was observed at the lower and higher ends of this range: most phenomena are either rare (< 10% and < 20%) or highly frequent (> 90%). In the latter group, prescriptive claims which contradict the empirically substantiated behaviour become hard to maintain.

In addition, it became apparent that the perceived salience of a specific phenomenon does not need to tally with its frequency in actual performance data. In fact, some of the

most widely discussed non-standard uses in the literature range among the least frequent phenomena in the corpus (16.2).

Different degrees of variation were evidenced for the four morphosyntactic categories of person, number, gender and case. The findings were summarised in the Hierarchy of morphosyntactic categories, a prioratisation which reflects the importance of each category for the correct processing of pronominal expressions: person > number > gender > case. This empirically deduced hierarchy was shown to correlate with the major typological trends in the realisation of pronoun features (17.1).

Among the four morphosyntactic categories, case is by far the most variable, or least robust, category. The basic structural condition for this variability is the fact that English is a fixed word order language where structural case does not need to manifest itself morphologically (cf. Bobaljik 2008). The empricial results in this study indicate that pronoun case in English may be following different possible developments of functionally or semantically empty grammatical categories as described by Greenberg (1991): firstly, the disappearance or partial disappearance of case distinctions (default case, see 16.3); and secondly, the reinterpretation or functional re-attribution of specific case forms, for instance as emphatic/ contrastive expressions or for other pragmatic purposes (see 17.2 and 17.3). In the investigated interviews, the loosened connection between pronoun case and agreement could not be shown to affect the communication process in any negative way. It was, therefore, argued that the only utterances where case serves its original purpose of establishing syntactic dependencies and thematic relations are those utterances where absence of differential case marking would cause ambiguity (9.3, 17.2).

The analyses presented in this study provide strong empirical evidence against the structuralist 'one form – one meaning' principle. The variation reflected in different non-standard phenomena shows that a purely formalist approach cannot explain the actual behaviour of personal pronouns and self-forms in Modern English, especially in the spoken medium. Instead, a complex array of extra-syntactic determinants is at work which can only be captured by a combined approach (17.2). Pragmatic uses, in particular, play an important role in pronoun choice (17.3). Regarding collocational influences, it was shown how the investigation of dialect phenomena could in the future benefit from new methods such as collexeme analysis (15.5).

Overall, pronoun behaviour in spontaneous speech does not conform to expectations based on more formal data and minimal pair examples, the reason being that what appears

ambiguous out of context usually becomes clear within context. In face-to-face conversations like the interviews used in this study, part of the necessity for morphology "to regulate expressive traffic" is lost, as compared to registers where "the absence of nonlinguistic, gestural and situational context information necessitates the support of the conventionalized and socially controlled organization principles to ensure its functioning." (Stein 1994: 6). Immediate practical implications of these observations concern language teaching. The demand of Emonds (1986: 124) that "the real emphasis in grammar teaching for native English speakers should be re-directed to an explicit linguistic formulation and appreciation of the differences in *natural language* class and ethnic group dialects" seems boldly optimistic. It is, however, feasible for teachers to advert to existing varieties, and to convey a sense for linguistic diversity and respect for vernaculars and the people who use them.

Venturing a look into the future of English, the development of case assignment is bound to be of great interest, especially given the current status of English as a high-contact language (cf. Trudgill 2009, on language contact and simplification). Several other issues which had to remain unanswered in this study may also provide incentives for future research. They include, for instance:

- the future development of generic question tags (ch. 6.4);
- the future development of gendered pronouns (ch. 7);
- the apparent correlation between coordination, O-form subjects and singular concord (ch. 8.3.2);
- the areal distribution of causal and temporal adverbial participles with '*with* + pronoun + V-*ing*' (apparent predominance in MID and N data, ch. 10.5);
- the future coexistence of different overt *you* imperative patterns (ch. 12.1);
- the exact nature of lexical/ collocational motivation in English genitives (ch. 13.4), and the future distribution of the two genitive structures with pronominal modifier *it* (ch. 13.2; the historical development of *s*- vs. *of*-genitives with 3SGn modifiers could also benefit from revision);
- the influence of geographic area and collocational context on the use of possessive *we*, including other varieties of English (ch. 14.1.1);
- and the apparent avoidance of possessive pronouns ending in -*s* (ch. 14.2).

Particularly interesting from a more general perspective are those determinants of variation which affect pronouns in different syntactic environments and, possibly, different varieties of English. Among the various recurrent determinants identified in 17.2, some aspects especially worth exploring are the role of emphasis, including the extent to which correlation can be observed between non-standard usage and stress; the avoidance of ambiguity, a linguistic principle for which appropriate tests still need to be developed in order to measure the influence on linguistic choices; and, finally, the apparent connection between the overall frequency of specific syntactic functions and the frequency of non-standard occurrences in these functions.

Appendix and References

Appendix A – Historical information

According to Mühlhäusler/ Harré (1990: 276), the earliest system distinguished "only two pronouns, a first and a second person singular." Compare Greenberg (1997) for 1SG suppletion between the nominative (*eĝ(h)om*) and non-nominative case (*m-*). According to Schmidt (1994: 184), the morphological distinction can be interpreted as a relict of the opposition agent (later NOM in stressed position) vs. non-agent (originally in unstressed/ enclitic position). Compare Howe (1996), for an account of personal pronouns in the early Germanic languages; compare Faltz (1985: 210/ 211), for the encoding of reflexivity in Proto-Germanic.

Tables A2–A10 are based on the *Cambridge History of the English Language* (Blake 1992) and Baker (2007). The tables present a simplified account of a pronominal system which at one time exhibited a considerable variety of forms and spellings, including enclitic forms like *nic* 'not I' (cf. Brunner 1965[3]: 258–267). Note the historical distinction in the second person between singulars in *t-* (OE *þū*, *þē*) and plurals in *y-*/*u-*/*w-* (OE *ġē*, *ēow*).

Table A1: The Indo-European personal pronoun paradigm after Schmidt (1978)

IndoEuropean	1SG	2SG	1PL	2PL
NOM	eĝ-H	tu-H	mes, later we-y	yuH-s (later form)
ACC	(e)me/mē	tu	mōs	us-me
DAT	(e)m-oy	tew, tu-bhey tew tew	ṃs-bhey	wos, wos-bhi
LOK				
GEN	(e)me-nē		ṃs-er-om	us-m-ey
			ṃser-om	wos(-om)
ABL	(e)med	tu-s	ṃs-^{e}/os	wos?
INSTR		tw-^{ē}/ō	---	wos?
POSS	(e)m^{e}/oy-o-	tew-o-, two-	ṃser-o-,mosro-	wos(e)r-o- (later form)

Table A2: 1SG pronouns from OE to ModE

1SG	OE	MidE	ModE	comments
NOM	ic, iċ	I,ich	I	ACC and DAT merge in MidE
ACC	mē,mec	me	me	
DAT	mē			
GEN	mīn	min,mi	my,mine	

Table A3: 2SG pronouns from OE to ModE

2SG	OE	MidE		ModE	comments
	formal/informal	informal	formal	formal/informal	
NOM	þū	thou	you	you	formal/informal distinction in MidE; ACC and DAT merged in MidE; subjective and objective identical in ModE
ACC	þē, þec	thee			
DAT	þē				
GEN	þīn	thy,thine	your,yours	your,yours	

Table A4: 3SGm pronouns from OE to ModE

3SGm	OE	MidE	ModE	comments
NOM	hē	he	he	ACC and DAT merged in MidE
ACC	hine	him	him	
DAT	him			
GEN	his	his	his	

Table A5: 3SGf pronouns from OE to ModE

3SGf	OE	MidE	ModE	comments
NOM	hēo	heo,sche,ho,he	she	DAT and GEN identical since OE; objective and GEN identical since MidE
ACC	hīe	hire,her,heore	her	
DAT	hire			
GEN			hers	

Table A6: 3SGn pronouns from OE to ModE

3SGn	OE	MidE	ModE	comments
NOM	hit	hit, it	it	NOM and ACC identical since OE; objectives identical since MidE; unification of forms in ModE
ACC		hit,it,him		
DAT	him			
GEN	his	his,its	its	

Table A7: 1PL pronouns from OE to ModE

1PL	OE	MidE	ModE	comments
NOM	wē	we	we	ACC and DAT identical since OE
ACC	ūs	us	us	
DAT				
GEN	ūre	ure,our	our,ours	

Table A8: 2PL pronouns from OE to ModE

2PL	OE	MidE	ModE	comments
	formal/informal	formal		
NOM	ġē	ye,you	you	ACC and DAT identical since OE; subjective and objective identical in ModE
ACC	ēow	you,ya		
DAT				
GEN	ēower	your	your,yours	

Table A9: Dual pronouns in OE

DUAL	1	2	comments
NOM	wit	ġit	ACC and DAT identical in OE; duals vanish in MidE
ACC	unc	inc	
DAT			
GEN	uncer	incer	

Table A10: 3PL pronouns from OE to ModE

3PL	OE	MidE	ModE	comments
NOM	hīe	hie,he,hi,ho,hie,þei,þai	they	NOM and ACC identical in OE; ACC and DAT identical since MidE; unification of forms in ModE
ACC	hīe	hem,ham,heom, þeim,þem,þam	them	
DAT	him			
GEN	hira	here,heore,hore,þair,þar	their,theirs	

Appendix B – Coding information

Table B1: Dialect area codes

Dialect area	Code
SE	1
SW	2
MID	3
N	4

Table B2: County codes

SE		SW		MID		N	
Kent	1	Cornwall	5	Leicestershire	10	Durham	13
London	2	Devon	6	Nottinghamshire	11	Lancashire	14
Middlesex	3	Oxfordshire	7	Shropshire	12	Northumberland	15
Suffolk	4	Somerset	8			Westmorland	16
		Wiltshire	9			Yorkshire	17

Table B3: Syntactic function table with corpus examples

POS	function	FRED example
10	subject	although **her** knows I ain't a hundred percent
11	subject, coord. initial	well, **me and** mi old chap was up there
12	subject, coord. final	and my dad **and him** went to join the army
13	subject, coord. middle	mother **and I and** mi sister came to the farm
14	question, no verb	{What school did you go to?} **Me**?
15	subject/ subj. compl., no verb	{Did you know the books?} No, **not me**, no.
16	subject/ subj. compl., no verb, coord.	{Who would milk the cows?} Oh, **I and** the wife.
17	by-subject	Everything else was done by hand, **by us**.
18	direct address of hearer	and I said to my old man, **You**, What d' you hit me
19	interjection/ exclamation	Oh goodness **me**!/ Oh, that's gone, **dearie me**.
20	object	...so the farmer didn't want **I**, did he?
21	object, coord. initial	they prayed **me and** mi brother to make them
22	object, coord. final	they put Peggy **and I** together
23	object, coord. middle	she only asked her own sisters **and me and** my husband
24	object, no verb	{your mum pushed you up there in a bassinet?} Not **me**.
30	subject complement	and there was **I** by there trying to light it
31	subject compl., coord. initial	there's just **I and** Brendan left
32	subject compl., coord. final	that were father **and I** on the binders
33	subject compl., coord. middle	and then there's John and Michael **and miself and** a girl
40	prepositional complement	he gave it to **myself**
41	prepositional compl., coord. initial	they'd all pitched onto **I and** Albee
42	prepositional compl., coord. final	my uncles used to drive the cart for Joshua **and they**
43	by + pronoun 'alone'/ 'on one's own'	it was a bit of a bore if you was **all by yourself** all night
44	'pronoun + V-*ing*' (object)	all for a copper to **keep us going**
47	for–to construction	Don't suppose you would like **for I to** draw your ship, cap'n?

Table B3 (cont.)

POS	function	FRED example
48	ECM construction	I said, They **want I to** go to Charlton.
34	ʻpronoun + V-*ing*ʼ (adverbial)	**Me being** light, I used to jump right at the top.
35	ʻpronoun + V-*ing*ʼ, coord. (adverbial)	**me or** mi mum or whoever it is **being** that bit of a boss
36	ʻpronoun + V-*ing*ʼ (gerund-participle)	I said, excuse **me asking** your age but...
38	ʻpronoun + V-*ing*ʼ (instead of zero)	instead o' **me gettin'** my ninety pound, I got seventeen
51	intensifier	He ought to pull that out **himself**.
52	reflexive	he had the tail of a shirt for **drying him on** go and **get you back** into bed mi lass
53	non-standard benefactive	I was standing there **having me** a drink... up jumped the pike, **have him a meal** if he could get one
60	pleonastic in imperative	**don't you ask** me a question like that
61	*mind you*	I had a cup of tea, mind **you**
62	pronoun + qualifier (synthetic)	all **us kids** used to go up there and have it and **we three** used to saw, go and do the timber sawing so **you each** got a ration o' herrin' **we all** got jobs/ **we** got **all** paid
62a	qualifier + pronoun (analytic)	There was **ten of we**!/ ʻcause she had to do for **all o' we**
62b	ʻ*all* + pronoun'	pretty nigh **all them** had Friday night
62c	ʻ*all* + pronoun', coord.	the sailors **and all them** used to come out
63	qualifier adjective + pronoun	So ʻcourse, **soft me**, I took them.
64	(generic) question tag	good idea though, **innit** we used to make our own bread, **didn't it** your brother got away with that case last year, **is it**
70	disjunctive, post-positioned	I love flowers, **me**.
71	disjunctive, pre-positioned	**Me**, I go twice every Sunday.
72	disjunctive, post-positioned, coord.	they used to go fishing, **him and** a man named Wood
73	disjunctive, pre-positioned, coord.	one of our mates **and misel'**, we were dare-devils
74	resumptive	Like a rubber collar thing that you could wash **en**.
80	possessive determiner	and we have done all **us lives** the doctor used to make **him own bottles**
81	*of*-genitive	the owner **of her**/ clever to get out of t' road **of you**
90	demonstrative determiner	when we had all **them** fish/ the Frenchmen used to come over in **they days** with onions

Appendix C – Supplementary analysis data

Table C1: Function matrix: S- and O-forms in the FRED corpus

syntactic function*	**S-forms**					**O-forms**					**homonym**
	I	**he**	**she**	**we**	**they**	**me**	**him**	**her**	**us**	**them**	**you**
subject	x	x	x	x	x	x	x	x	x	x	x
subject, coord. initial	x	x	x	x	x	x	x	x	---	x	x
subject, coord. middle	x	---	---	---	---	---	---	---	---	---	---
subject, coord. final	x	---	---	---	x	x	x	x	---	x	---
subject complement	x	x	x	x	---	x	x	x	x	x	x
subject compl., coord. initial	x	---	---	---	---	x	x	x	x	x	---
subject compl., coord. middle	---	---	---	---	---	x	x	---	---	---	---
subject compl., coord. final	x	---	---	---	---	x	x	x	---	x	---
subject/subject compl., no verb	---	x	---	---	---	x	x	x	x	x	---
subject/subject compl., coord., no verb	x	x	---	---	---	x	x	x	---	---	x
question, no verb	---	---	---	---	---	x	---	---	---	---	---
(generic) question tag	x	x	x	x	x	---	x	---	x	x	x
by-subject	---	---	---	---	---	x	x	---	x	---	---
object	x	x	x	x	x	x	x	x	x	x	x
object, coord. initial	---	---	---	---	---	x	x	---	x	x	---
object, coord. middle	---	---	---	---	---	x	---	---	---	---	---
object, coord. final	x	---	---	---	---	x	---	x	x	x	---
object, no verb	---	---	---	---	---	x	---	---	---	---	---
prepositional complement	x	x	x	x	x	x	x	x	x	x	x
prepositional compl., coord. initial	x	---	---	---	---	x	x	x	x	---	---
prepositional compl., coord. final	x	---	---	---	x	x	x	---	---	x	---
'pronoun + V-*ing*' (object)	---	---	---	---	---	x	x	x	x	x	x
'pronoun + V-*ing*', coord. (object)	x	---	---	---	---	x	---	---	---	---	---
'pronoun + V-*ing*' (adverbial)	---	---	---	x	x	x	x	x	x	x	x
'pronoun + V-*ing*', coord. (adverbial)	---	---	---	---	---	x	---	---	---	---	---
'pronoun + V-*ing*' (gerund-part.)	---	x	---	x	---	x	x	x	x	x	x
'pronoun + V-*ing*' (instead of zero)	---	---	---	---	---	x	---	---	---	---	x
for–to construction	x	---	---	x	x	x	x	x	x	x	x
ECM construction	x	---	---	x	---	x	---	x	---	x	---
interjection/exclamation	---	---	---	---	---	x	---	---	---	---	---
pleonastic in imperative	---	---	---	---	---	---	---	---	---	---	x
mind you	---	---	---	---	---	---	---	---	---	---	x
pronoun + qualifier noun or numeral	---	---	---	x	x	---	---	---	x	x	x
'*all* + pronoun'	---	---	---	---	x	---	---	---	---	x	---
'*all* + pronoun', coord.	---	---	---	---	---	---	---	---	---	x	---

Table C1 (cont.)

direct address of the hearer	---	---	---	---	---	---	---	---	---	---	x
qualifier adjective + pronoun	---	---	---	---	---	x	---	---	---	---	---
disjunctive, post-positioned	x	x	---	---	x	x	x	---	x	x	---
disjunctive, pre-positioned	---	---	---	---	---	x	x	---	x	x	---
disjunctive, coord., post-positioned	x	x	x	---	---	x	x	x	x	x	---
disjunctive, coord., pre-positioned	x	x	---	---	---	x	x	x	---	x	x
resumptive	---	x	x	---	x	---	x	---	---	x	---
reflexive (incl. usually intransitive V)	---	---	---	---	---	x	x	x	x	x	x
non-standard benefactive	---	---	---	---	---	x	x	---	---	---	x
of-genitive	x	---	---	---	x	---	x	x	x	x	x
possessive	---	---	---	x	---	---	x	---	x	x	x
demonstrative	---	---	---	---	x	---	---	---	---	x	---
TOTAL NUMBER OF CONTEXTS	**21**	13	8	12	15	**35**	29	21	22	30	21

*only confirmed functions

Table C2: Function matrix: self-forms in the FRED corpus

syntactic function*	**self-forms**							
	my-	**yourself**	**him-**	**her-**	**it-**	**our-**	**yourselves**	**them-**
subject	---	---	---	---	---	---	---	---
subject, coord. initial	---	---	---	---	---	---	---	---
subject, coord. middle	---	---	---	---	---	---	---	---
subject, coord. final	x	---	x	---	---	x	---	---
subject complement	x	---	---	---	---	---	---	---
subject compl., coord. initial	x	---	---	---	---	---	---	---
subject compl., coord. middle	x	---	---	---	---	---	---	---
subject compl., coord. final	x	---	x	---	---	---	---	---
subject/subject compl., no verb	x	x	---	---	---	---	---	---
subject/subject compl., coord., no verb	x	---	---	---	---	---	---	---
question, no verb	---	---	---	---	---	---	---	---
(generic) question tag	---	---	---	---	---	---	---	---
by-subject	---	---	---	---	---	---	---	---
object	x	x	---	x	---	x	---	x
object, coord. initial	---	---	---	---	---	---	---	---
object, coord. middle	x	---	---	---	---	---	---	---
object, coord. final	---	---	---	---	---	---	---	---
object, no verb	x	---	---	---	---	---	---	---
prepositional complement	x	---	x	x	---	x	---	x
prepositional compl., coord. initial	---	---	---	---	---	---	---	---
prepositional compl., coord. final	---	---	---	---	---	---	---	---
'by + pronoun' meaning 'alone'	x	x	x	x	x	x	---	x
'pronoun + V-*ing*' (object)	---	---	---	---	---	---	---	---
‘pronoun + V-*ing*’, coord. (object)	x	---	---	---	---	---	---	---
disjunctive, post-positioned	x	---	---	---	---	---	---	---
disjunctive, pre-positioned	x	---	x	---	---	x	---	x
disjunctive, coord., post-positioned	x	---	---	---	---	---	---	---
disjunctive, coord., pre-positioned	x	---	---	---	---	---	---	---
intensifier	x	x	x	x	x	x	---	x
reflexive (incl. usually intransitive V)	x	x	x	x	x	x	x	x
benefactive with usually monotransit. V	x	x	---	---	---	---	---	---
TOTAL NUMBER OF CONTEXTS	**20**	6	7	5	3	7	1	6

* some non-confirmed functions excluded

Table C3: Top attracted collexemes of demonstrative determiner *they*

DEM = *they*, most strongly attracted collexemes					
COLL	freq. DEM in C	freq. COLL in C	observed freq. DEM_COLL in C	expected freq. DEM_COLL in C	CollStr
days	147	1347	85	49	**9.23**
cows	147	10	3	0	2.32
two	147	33	5	1	2.2
pots	147	4	2	0	2.12
times	147	23	4	1	2.06

Table C4: Top attracted collexemes of demonstrative determiner *them*

DEM = *them*, most strongly attracted collexemes					
COLL	freq. DEM in C	freq. COLL in C	observed freq. DEM_COLL in C	expected freq. DEM_COLL in C	CollStr
days	1309	1347	752	438	**107.89**
places	1309	30	16	10	1.84
boats	1309	28	15	9	1.78
jobs	1309	19	11	6	1.71
docks	1309	3	3	1	1.47

Table C5: Top repulsed collexemes of demonstrative determiner *these*

DEM = *these*, most strongly repulsed collexemes					
COLL	freq. DEM in C	freq. COLL in C	observed freq. DEM_COLL in C	expected freq. DEM_COLL in C	CollStr
days	1713	1347	6	573	infinite
times	1713	23	2	9.79	3.35
sorts	1713	44	11	18.72	1.93
three	1713	11	1	4.68	1.69
two	1713	33	10	14	0.98

Table C6: Top repulsed collexemes of demonstrative determiner *those*

DEM = *those*, most strongly repulsed collexemes					
COLL	freq. DEM in C	freq. COLL in C	observed freq. DEM_COLL in C	expected freq. DEM_COLL in C	CollStr
things	857	110	12	23	2.51
people	857	90	10	19.16	2.07
boats	857	28	1	5.96	1.99
lads	857	19	0	4.04	1.98
men	857	48	4	10.22	1.83

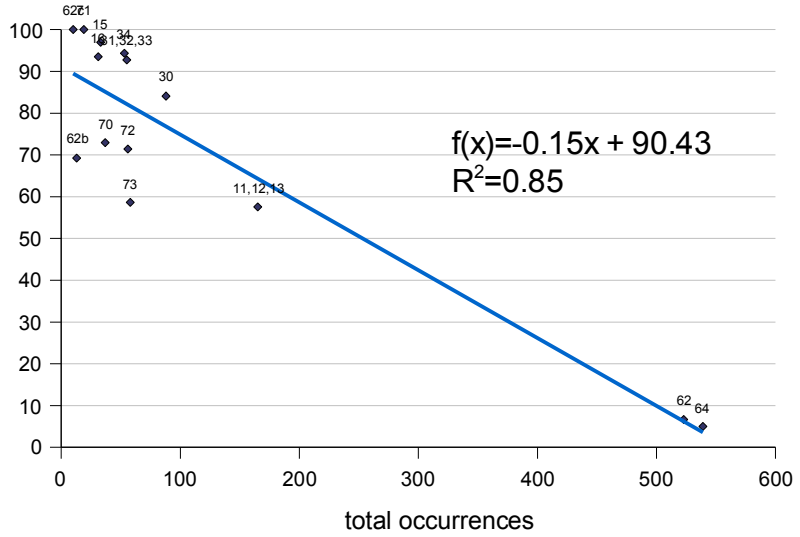
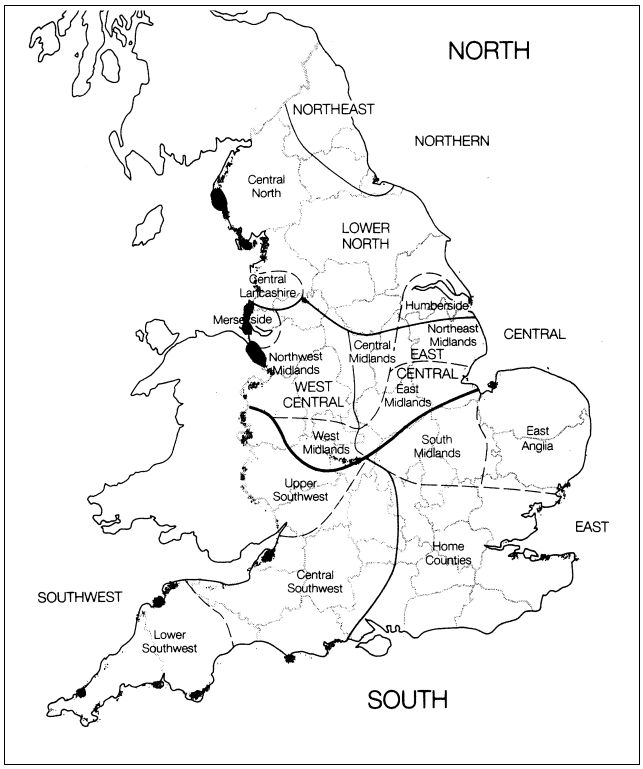
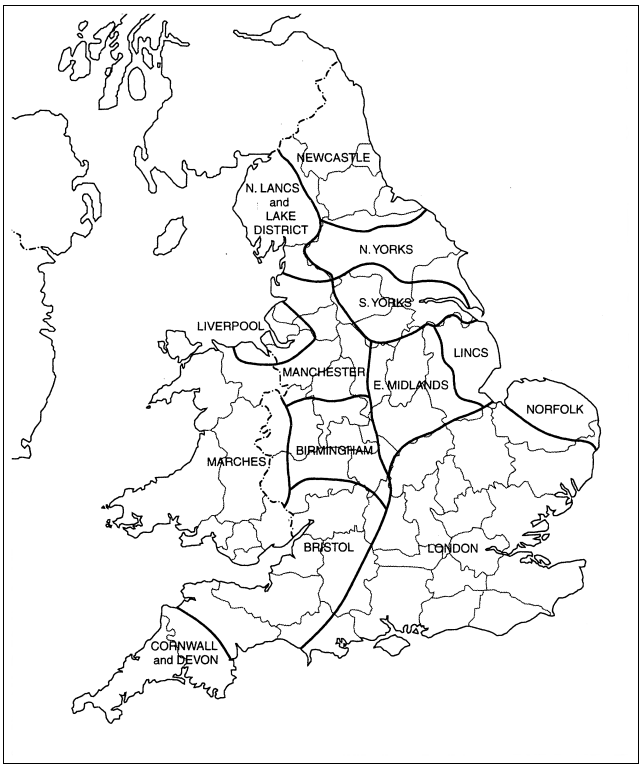


Figure C1: Linear regression model of syntactic functions and non-standard occurrences (preliminary observation discussed in 17.2; function codes correspond to Table B3)

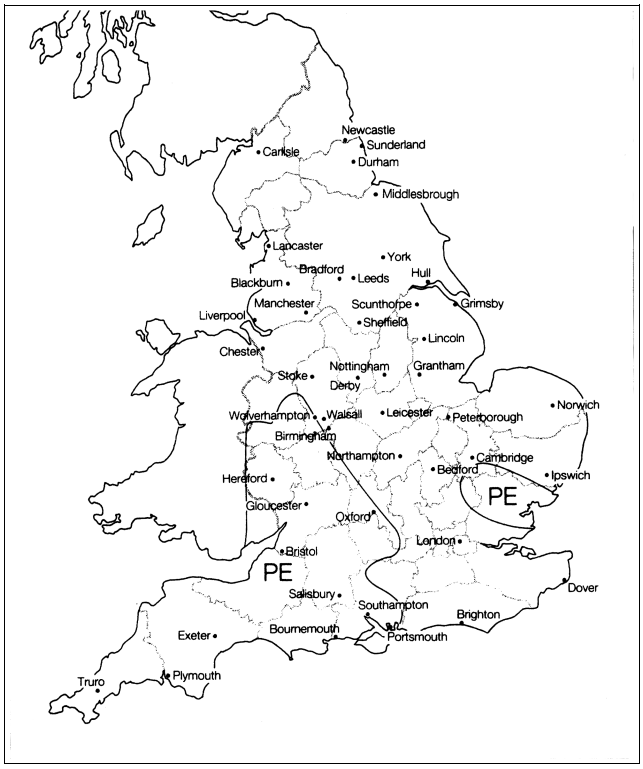
Appendix D – Maps



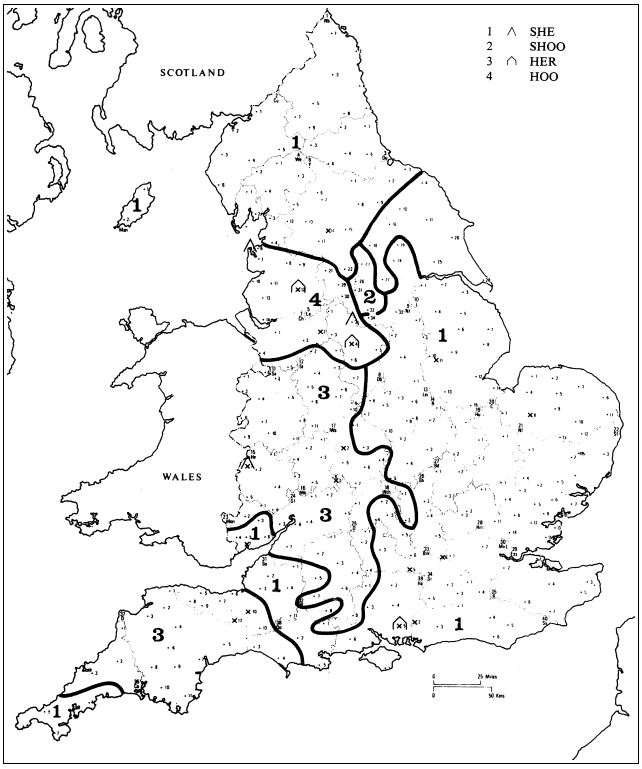
Map D1: England's modern dialect areas (Trudgill 1999[2] : 65)



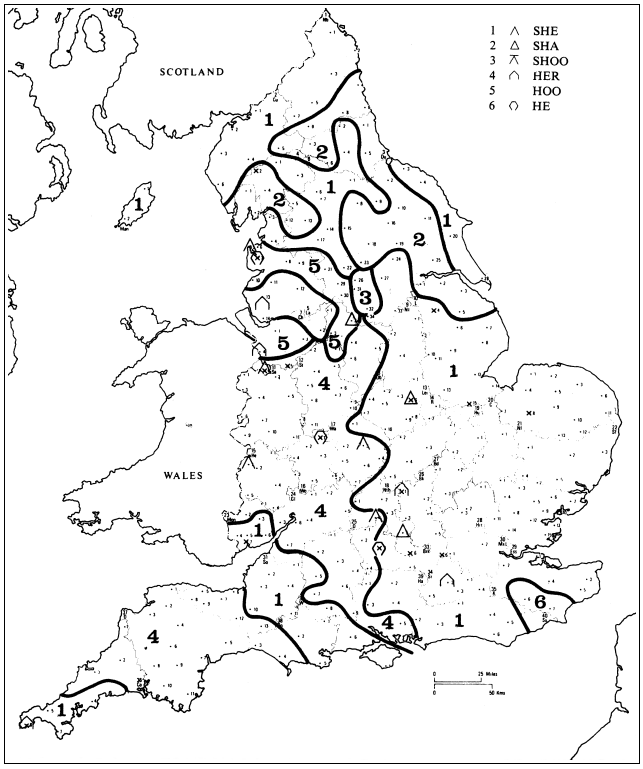
Map D2: Possible future dialect areas in England (Trudgill 1999[2]: 83)



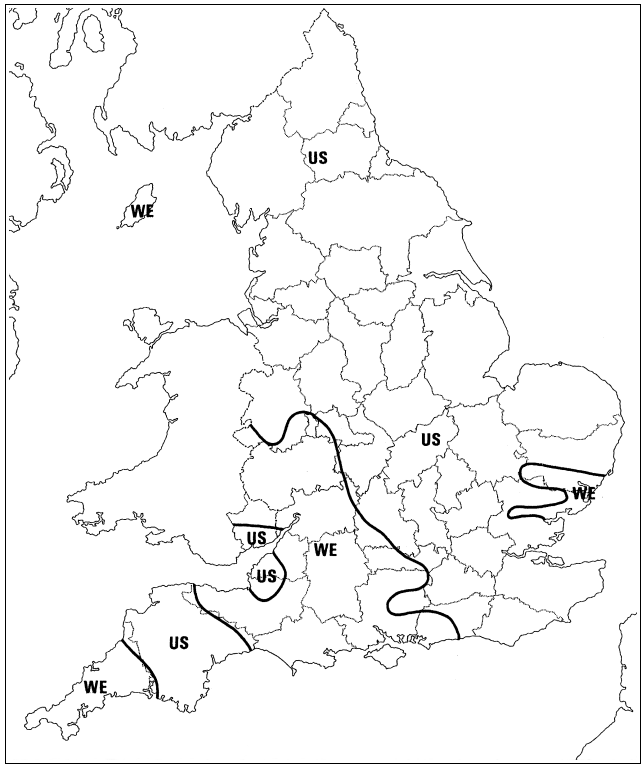
Map D3: Pronoun exchange areas according to Trudgill (1999[2]: 96)



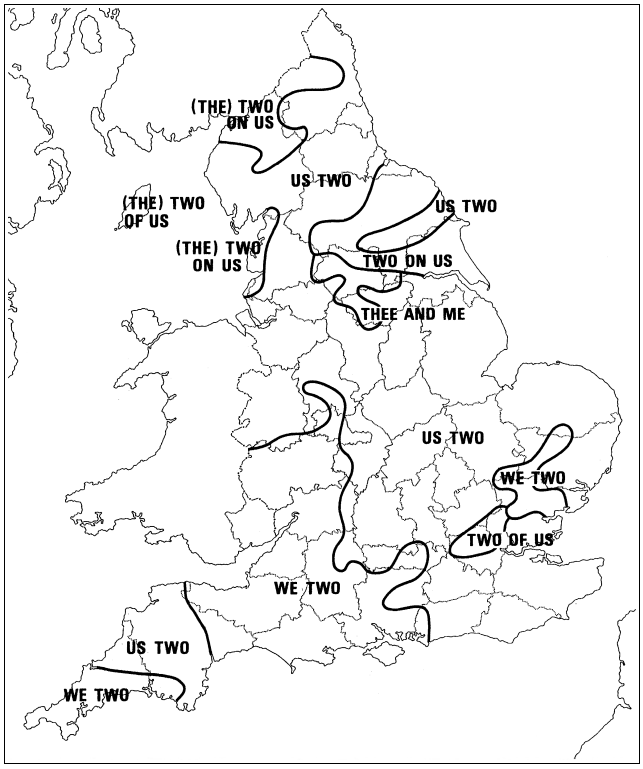
Map D4: 'SHE is' (LAE M68)



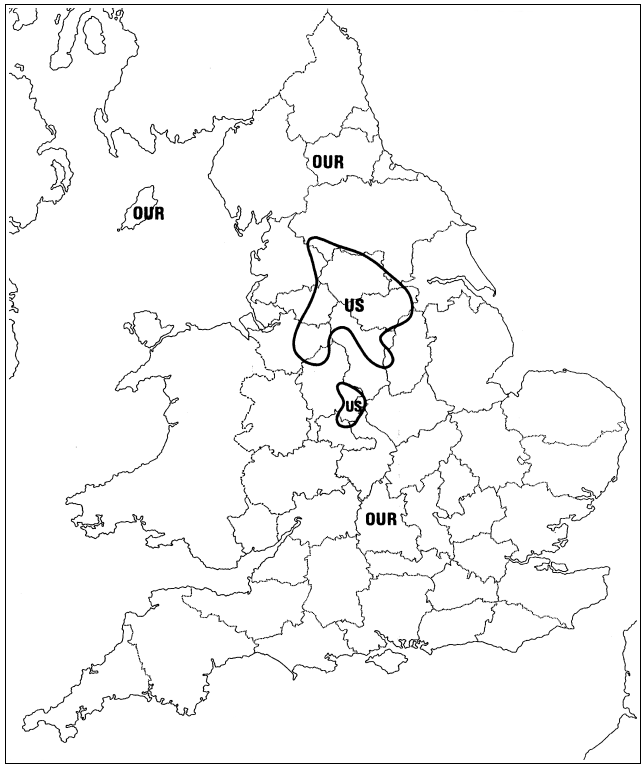
Map D5: '(is) SHE (married)' (LAE M69)



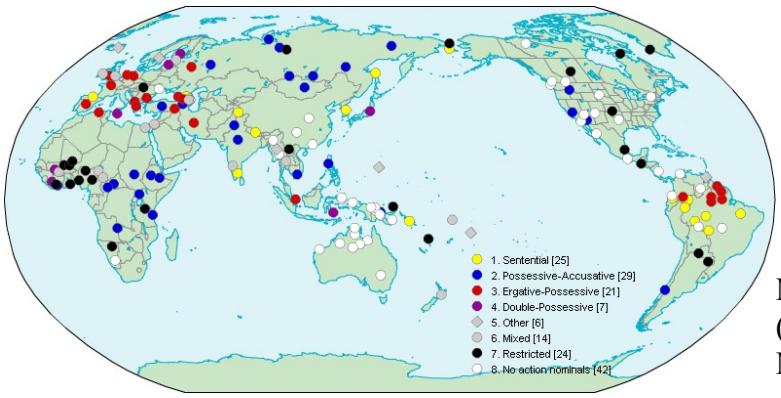
Map D6: ‘US’ (WM 178)



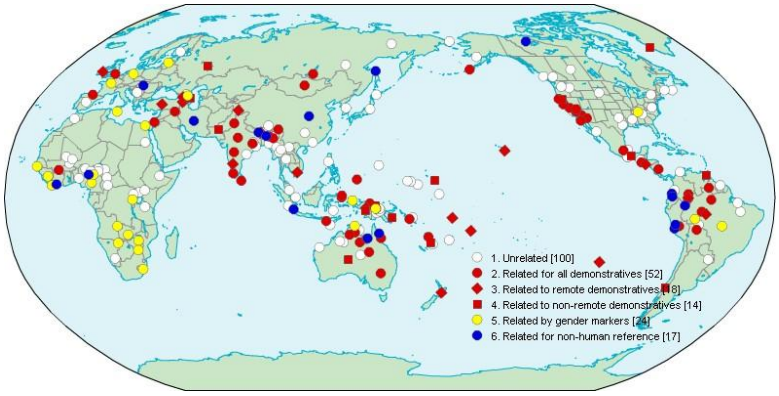
Map D7: ‘just WE TWO’ (WM 185)



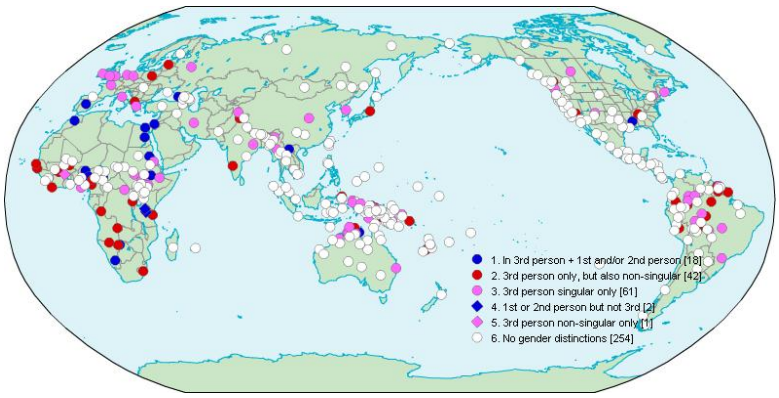
Map D8: ‘We have OUR own’ (WM 116)



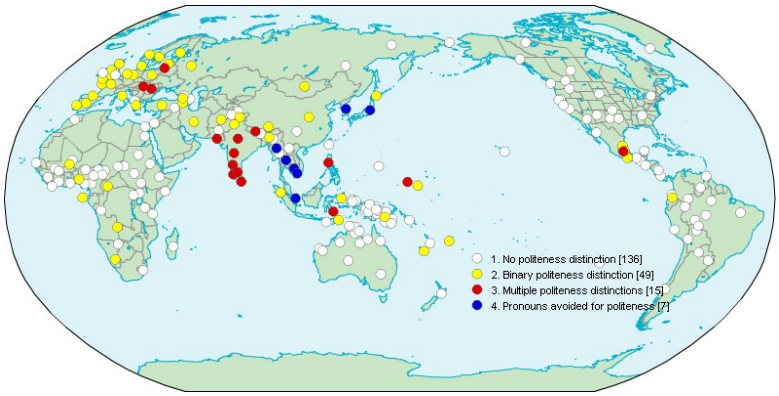
Map D9: Action nominal constructions (WALS Interactive reference tool: map by Maria Koptjevskaja-Tamm)



Map D10: Third person pronouns and demonstratives (WALS Interactive reference tool: map by D.N.S. Bhat)



Map D11: Gender distinctions in independent personal pronouns (WALS Interactive reference tool: map by Anna Siewierska)



Map D12: Politeness distinctions in pronouns (WALS Interactive reference tool: map by Johannes Helmbrecht)

References

Abney, Steven Paul. 1987. *The English Noun Phrase in Its Sentential Aspect*. Ph.D. thesis, Massachusetts Institute of Technology. MIT Working Papers in Linguistics.

Adger, David. 2006. "Combinatorial Variability." *Journal of Linguistics* 42(3): 503–530.

Alba-Salas, Josep. 2004. "Lexically Selected Expletives: Evidence from Basque and Romance." *SKY Journal of Linguistics* 17: 35–100.

Altenberg, Bengt. 1982. *The Genitive v. the Of-Construction: A Study of Syntactic Variation in 17th Century English*. Lund studies in English 62. Lund: CWK Gleerup.

Anderwald, Lieselotte. 2003. "Non-Standard English and Typological Principles: The Case of Negation." In: Rohdenburg and Mondorf (eds.), 507–529.

Anderwald, Lieselotte. 2004. "The Varieties of English Spoken in the Southeast of England: Morphology and Syntax." In: Kortmann and Schneider (eds.). 175–195.

Anderwald, Lieselotte, and Bernd Kortmann. 2002. "Typology and Dialectology: A Programmatic Sketch." In: Jan Berns and Jaap van Marle (eds.), *Present-Day Dialectology – Problems and Findings*, Berlin/ New York: Mouton de Gruyter, 159–171.

Anderwald, Lieselotte, and Susanne Wagner. 2007. "FRED – the Freiburg English Dialect Corpus – Applying Corpus-Linguistic Research Tools to the Analysis of Dialect Data." In: Joan C. Beal, Karen P. Corrigan and Hermann L. Moisl (eds.), *Creating and Digitizing Language Corpora, Vol.1: Synchronic Databases*, London: Palgrave Macmillan, 35–53.

Angermeyer, Philipp S., and John V. Singler. 2003. "The Case for Politeness: Pronoun Variation in Co-ordinate NPs in Object Position in English." *Language Variation and Change* 15: 171–209.

Anttila, Arto, and Vivienne Fong. 2004. "Variation, Ambiguity and Noun Classes in English." Special issue *Variation in Form versus Variation in Meaning. Lingua* 114 (9–10): 1253–1290.

Ariel, Mira. 1990. *Accessing Noun-Phrase Antecedents*. Croom Helm linguistics series. London: Routledge.

Audring, Jenny. 2006. "Pronominal Gender in Spoken Dutch." *Journal of Germanic Linguistics* 18: 85–116.

Auer, Peter. 2004. "Non-Standard Evidence in Syntactic Typology: Methodological Remarks on the Use of Dialect Data vs Spoken Language Data." In: Kortmann (ed.),

69–92.

Azevedo, Milton M. 1989. “Vernacular Features in Educated Speech in Brazilian Portuguese.” *Hispania* 72(4): 862–872.

Bache, Carl. 1985. “The Semantics of Grammatical Categories: A Dialectical Approach.” *Journal of Linguistics* 21(1): 51–77.

Baker, Carl L. 1995. “Contrast, Discourse Prominence, and Intensification, with Special Reference to Locally Free Reflexives in British English.” *Language* 71(1): 63–101.

Baker, Peter S. 2007. *Introduction to Old English*. Oxford: Blackwell.

Baker, Robert. 1770 (1968 reprint). *Reflections on the English Language*. Menston: Scolar Press.

Bakker, Willem Frederic. 1974. *Pronomen Abundans and Pronomen Coniunctum: A Contribution to the History of the Resumptive Pronoun within the Relative Clause in Greek*. Amsterdam: North-Holland Publishing.

Barber, Charles. 1976. *Early Modern English*. London: André Deutsch.

Barbiers, Sjef, Leonie Cornips and Susanne van der Kleij (eds.). 2002. *Syntactic Microvariation*. Electronic publication of the Meertens Institute, Amsterdam. http://www.meertens.knaw.nl/books/synmic/.

Barbiers, Sjef, Hans Bennis, Gunther de Vogelaer, Magda Devos and Margreet van der Ham. 2005. *SAND: Syntactische Atlas van de Nederlandse Dialecten, Vol. 1: Maps*. Amsterdam: Amsterdam University Press.

Barker, Chris, and David Dowty. 1993. “Non-Verbal Thematic Proto-Roles.” In: Amy Schafer (ed.) *Proceedings of the North-Eastern Linguistic Society 23, Vol. 1*, Amherst: GLSA, 49–62.

Barnes, William. 1886. *A Glossary of the Dorset Dialect*. London: Trübner.

Baron, Dennis. 1986. *Grammar and Gender*. New Haven: Yale University Press.

Bauer, Laurie, and Peter Trudgill (eds.). 1998. *Language Myths*. London: Penguin.

Beal, Joan. 1993. “The Grammar of Tyneside and Northumbrian English.” In: Milroy and Milroy (eds.), 187–213.

Beal, Joan. 2004. “English Dialects in the North of England: Morphology and Syntax.” In: Kortmann and Schneider (eds.), 114–141.

Bernstein, Cynthia. 2003. “Grammatical Features of Southern Speech: *Yall*, *Might Could*, and *Fixin To*.” In: Stephen J. Nagle and Sara L. Sanders (eds.), *English in the Southern United States*, Cambridge: Cambridge University Press, 106–118.

Bhat, Darbhe Narayana Shankara. 2004. *Pronouns*. Oxford Studies in Typology and Linguistic Theory. Oxford: Oxford University Press.

Bhat, Darbhe Narayana Shankara. 2005. “Third-Person Pronouns and Demonstratives.” In: Haspelmath et al. (eds.), 178–181.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Black, James R., and Virginia Motapanyane. 1996. *Microparametric syntax and dialect variation*. Amsterdam Studies in the Theory and History of Linguistic Science 139. Amsterdam/ Philadelphia: Benjamins.

Blake, Barry J. 2001[2]. *Case*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Blake, Norman F. 1992. *The Cambridge History of the English Language, Vol. II: 1066 – 1476*. Cambridge: Cambridge University Press.

Bliss, Alan Joseph. 1979. *Spoken English in Ireland 1600–1740: Twenty-Seven Representative Texts*. Dublin: Dolmen Press.

Boas, Franz. 1911. *Handbook of American Indian Languages*. Bureau of American Ethnology, Bulletin 40. Washington: Government Print Office.

Bobaljik, Jonathan D. 2008. “Where's Phi? Agreement as a Post-Syntactic Operation.” In: Daniel Harbour, David Adger and Susana Béjar (eds.), *Phi Theory: Phi-Features across Modules and Interfaces*, Oxford Studies in Theoretical Linguistics 16, Oxford: Oxford University Press, 295–328.

Bodine, Ann. 1998[2]. “Androcentrism in Prescriptive Grammar: Singular 'They', Sex-Indefinite 'He', and 'He or She'.” In: Deborah Cameron (ed.), *The Feminist Critique of Language: A Reader*, London: Routledge, 124–138.

Bolinger, Dwight L. 1977. *Meaning and form*. English Language Series 11. London: Longman.

Bornkessel, Ina, Matthias Schlesewsky and Angela D. Friederici. 2003a. “Contextual Information Modulates Initial Processes of Syntactic Integration: The Role of Inter- versus Intrasentential Predictions.” *Journal of Experimantal Psychology: Learning, Memory, and Cognition* 29(5): 871–882.

Bornkessel, Ina, Matthias Schlesewsky and Angela D. Friederici. 2003b. “Eliciting Thematic Reanalysis Effects: The Role of Syntax-Independent Information during Parsing.” *Language and Cognitive Processes* 18(3): 269–298.

Bosch, Peter. 1985. “Constraints, Coherence, Comprehension. Reflections on Anaphora.” In: Emel Sözer (ed.), *Text Connexity, Text Coherence*, Hamburg: Buske, 299–319.

Boyland, Joyce Tang. 2001. “Hypercorrect Pronoun Case in English? Cognitive Processes that Account for Pronoun Usage.” In: Joan Bybee and Paul Hopper, *Frequency and the*

Emergence of Linguistic Structure, Amsterdam/ Philadelphia: John Benjamins, 383–404.

Branco, António, and Palmira Marrafa. 1997. "Subject-Oriented and Non Subject-Oriented Long-Distance Anaphora: An Integrated Approach." In: Byung-Soo Park and Jong-Bok Kim (eds.), *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation (PACLIC 11)*, Seoul: Kyung Hee University, 21–30.

Brinton, Laurel J. 1995. "Non-Anaphoric Reflexives in Free Indirect Style: Expressing the Subjectivity of the Non-Speaker." In: Stein and Wright (eds.), 173–194.

Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage.* Studies in Interactional Sociolinguistics 4. Cambridge: Cambridge University Press.

Brunner, Karl. 1965[3]. *Altenglische Grammatik: nach der angelsächsischen Grammatik von Eduard Sievers*. Tübingen: Niemeyer.

Buchholtz, Mary. 2000. "The Politics of Transcription." *Journal of pragmatics* 32(10): 1439–1465.

Bühler, Karl. 1934/ 1999[3]. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Fischer/ Stuttgart: Lucius und Lucius.

Burchfield, Robert W. (ed.). 1994. *The Cambridge History of the English Language, Vol. V: English in Britain and Overseas: Origins and Developments*. Cambridge: Cambridge University Press.

Campbell, Alistair. 1962. *Old English Grammar*. Oxford: Clarendon Press.

Cantrall, William R. 1973. "Reflexive Pronouns and Viewpoint." *Linguistische Berichte* 28: 42–50.

Cantrall, William R. 1974. *Viewpoint, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton.

Cardinaletti, Anna, and Michael Starke. 1994. "The Typology of Structural Deficiency: On the Three Grammatical Classes." University of Venice *Working Papers in Linguistics* 4(2): 41–109.

Carminati, Maria Nella. 2005. "Processing Reflexes of the Feature Hierarchy (Person > Number > Gender) and Implications for Linguistic Theory." *Lingua* 115: 259–285.

Chambers, Jack K. 2004. "Dynamic Typology and Vernacular Universals." In: Kortmann (ed.), 127–145.

Chambers, Jack K., and Peter Trudgill. 1998[2]. *Dialectology*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Cheshire, Jenny, Viv Edwards and Pamela Whittle. 1993. "Non-Standard English and

Dialect Levelling." In: Milroy and Milroy (eds.), 53–96.

Cheshire, Jenny, and Viv Edwards. 1998. "Lessons from a Survey of British Dialect Grammar." In *Links & Letters* 5: 61–73.

Cheshire, Jenny, and Dieter Stein (eds.). 1997a. *Taming the Vernacular – From Dialect to Written Standard Language.* London: Longman.

Cheshire, Jenny, and Dieter Stein. 1997b. "The Syntax of Spoken Language." In: Cheshire and Stein (eds.), 1–12.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Studies in Generative Grammar, 9. Dordrecht: Foris.

Cleland, Alexandra A., and Martin J. Pickering. 2003. "The Use of Lexical and Syntactic Information in Language Production: Evidence from the Priming of Noun-Phrase Structure." *Journal of Memory and Language* 49: 214–230.

Corbett, Greville G. 1991. *Gender*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Corbett, Greville G. 2000. *Number*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Coupland, Nikolas. 1988. *Dialect in Use: Sociolinguistic Variation in Cardiff English*. Cardiff: University of Wales Press.

Croft, William. 1990. *Typology and Universals.* Cambridge: Cambridge University Press.

Crowley, Tony. 1989. *Standard English and the Politics of Language*. Urbana: University of Illinois Press.

Curzan, Anne L. 2003. *Gender Shifts in the History of English*. Cambridge: Cambridge University Press.

Dahl, Östen. 2001. "Principles of Areal Typology." In: Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.), 2001, *Language Typology and Language Universals: An International Handbook, Vol. 1*, Berlin/ New York: De Gruyter, 1456–1470.

Dalrymple, Mary, and Ronald M. Kaplan. 2000. "Feature Indeterminacy and Feature Resolution." *Language* 76(4): 759–798.

De Smet, Hendrik, and Liesbet Heyvaert. 2008. "Gerunds, Participles or Gerund-Participles? A Usage-Based Perspective on Some of Their Distributional, Formal and Semantic Properties." Paper presented at the first conference of the International Society for the Linguistics of English, Freiburg, 8–11 October (ISLE 2008).

De Smet, Hendrik, and Liesbet Heyvaert. 2009. "The Meaning of *-ing*. Semantics and Pragmatics". Paper presented at the Third International Conference on the Linguistics of

Contemporary English (ICLCE3). University of London, 14–17 July 2009. (preparing paper for submission)

De Vogelaer, Gunther, and Gert De Sutter. 2011. "The Geography of Gender Change: Pronominal and Adnominal Gender in Flemish Dialects of Dutch." *Language Sciences* 33(1): 192–205.

De Vries, Mark. 2005. "Coordination and Syntactic Hierarchy." *Studia Linguistica* 59: 83–105.

Denison, David. 1998. "Syntax." In: Suzanne Romaine (ed.), *The Cambridge History of the English Language, Vol. IV: 1776–1997*, Cambridge: Cambridge University Press, 92–329.

Dressler, Wolfgang U. 1985. "On the Predictiveness of Natural Morphology." *Journal of Linguistics* 21(2): 321–337.

Duncan, Pauline. 1972. "Forms of the Feminine Pronoun in Modern English Dialects." In: Martyn Wakelin (ed.), *Patterns in the Folk Speech of the British Isles*, London: Athlone, 182–200.

Dunkling, Leslie. 1990. *A Dictionary of Epithets and Terms of Address*. London: Routledge.

Ellis, Alexander John. 1869–1889. *On Early English Pronunciation: With Especial Reference to Shakespeare and Chaucer, Vol. 1–5*. London: Trübner.

Elworthy, Frederic Thomas. 1875–1886. *The Grammar of the Dialect of West Somerset (containing The Dialect of West Somerset, Grammar of West Somerset Dialect, West Somerset Word-Book, or Glossary)*. English Dialect Society, 17. London: Trübner.

Emonds, Joseph. 1986. "Grammatically Deviant Prestige Constructions." In: Michael Brame, Heles Contreras and Frederick J. Newmeyer (eds.), *A Festschrift for Sol Saporta*, Seattle: Noit Amrofer, 93–129.

Erdmann, Peter. 1978. "It's I, It's Me: A Case for Syntax." *Studia Anglica Posnaniensia* 10: 67–80.

Evans, Bergen, and Cornelia Evans. 1957. *A Dictionary of Contemporary American Usage*. New York: Random House.

Faltz, Leonard M. 1985. *Reflexivization: A Study in Universal Syntax*. Outstanding Dissertations in Linguistics series. New York/ London: Garland.

Fanselow, Gisbert. 2002. "Quirky 'Subjects' and Other Specifiers." In: Kaufmann and Stiebels (eds.), 227–250.

Farmer, Ann K., and Robert M. Harnish. 1987. "Communicative Reference with Pronouns." In: Jef Verschueren and Marcella Bertuccelli-Papi (eds.), *The Pragmatic*

Perspective: Selected Papers from the 1985 International Pragmatics Conference, Amsterdam/ Philadelphia: John Benjamins, 547–565.

Filppula, Markku. 1999. *The Grammar of Irish English – Language in Hibernian Style*. London: Routledge.

Filppula, Markku. 2004. "Irish English: Morphology and Syntax." In: Kortmann and Schneider (eds.), 73–101.

Fischer, Susann. 2002. *The Catalan Clitic System: A Diachronic Perspective on its Syntax and Phonology*. Interface Explorations 5. Berlin/ New York: Mouton de Gruyter.

Forchheimer, Paul. 1953. *The Category of Person in Language*. Berlin: Walter de Gruyter.

Frajzyngier, Zygmunt, and Traci S. Curl (eds.). 2000. *Reflexives: Forms and Functions*. Amsterdam/ Philadelphia: John Benjamins.

Geniušiene, Emma. 1987. *The Typology of Reflexives*. Berlin/ New York: Mouton de Gruyter.

Givón, Talmy. 2001. *Syntax: An Introduction, Vol. 1*. Amsterdam/ Philadelphia: John Benjamins.

Goldwater, Micah B., and Jeffrey T. Runner. 2006. "Coreferential Interpretations of Reflexives in Picture Noun Phrases: An Experimental Approach." In: Pascal Denis et al. (eds.), *Proceedings of the 2004 Texas Linguistics Society Conference*, Somerville, MA: Cascadilla Proceedings Project, 28–34.

Grano, Thomas. 2006. " 'Me and Her' Meets 'He and I': Case, Person, and Linear Ordering in English Coordinated Pronouns." Unpublished BA thesis, Stanford University. http://home.uchicago.edu/~tgrano/uht.pdf.

Green, Lisa J. 2002. *African American English – A Linguistic Introduction*. Cambridge: Cambridge University Press.

Greenbaum, Sidney. 1996. *The Oxford English Grammar*. Oxford: Oxford University Press.

Greenbaum, Sidney, and Gerald Nelson. 2002[2]. *An Introduction to English Grammar*. London: Longman.

Greenbaum, Sidney, and Randolph Quirk. 1990. *A Student's Grammar of the English Language*. Harlow: Longman.

Greenberg, Joseph H. 1963. "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements." In: Joseph H. Greenberg (ed.), *Universals of Language: Report of a Conference Held at Dobbs Ferry, New York, April 13–15*, Cambridge, Mass.: MIT Press, 73–113.

Greenberg, Joseph H. 1966. *Language Universals: With Special Reference to Feature*

Hierarchies. The Hague: Mouton.

Greenberg, Joseph H. 1991. “The Last Stages of Grammatical Elements: Contractive and Expansive Desemanticization.” In: Elizabeth C. Traugott and Bernd Heine, *Approaches to Grammaticalization*, Amsterdam: John Benjamins, 301–14.

Greenberg, Joseph H. 1997. “The Indo-European First and Second Person Pronouns in the Perspective of Eurasiatic, Especially Chukotkan.” *Anthropological Linguistics* 39(2): 187–195.

Gries, Stefan Th. 2007. Coll.analysis 3.2a. A program for R for Windows 2.x; R version 2.9.1. http://www.linguistics.ucsb.edu/faculty/stgries/teaching/groningen/.

Gries, Stefan Th., Beate Hampe and Doris Schönefeld. 2005. “Converging Evidence: Bringing Together Experimental and Corpus Data on the Association of Verbs and Constructions.” *Cognitive Linguistics* 16.4: 635–76.

Gries, Stefan Th., and Anatol Stefanowitsch. 2004. “Extending Collostructional Analysis: A Corpus-Based Perspective on 'Alternations'.” *International Journal of Corpus Linguistics* 9.1: 97–129.

Gries, Stefan Th., and Anatol Stefanowitsch. 2010. “Cluster Analysis and the Identification of Collexeme Classes.” In: John Newman and Sally Rice (eds.), *Empirical and Experimental Methods in Cognitive/ Functional Research*, Stanford, CA: Center for the Study of Language and Information, 73–90.

Gutzmann, Daniel. 2007. “Eine Implikatur konventioneller Art: Der Dativus Ethicus.” *Linguistische Berichte* 211: 277–308.

Halliwell, James O. 1887. *A Dictionary of Archaic and Provincial Words, Obsolete Phrases, Proverbs, and Ancient Customs, from the Fourteenth Century: In Two Volumes*. London: John Russell Smith.

Harley, Heidi, and Elizabeth Ritter. 2002a. “Person and Number in Pronouns: A Feature-Geometric Analysis.” *Language* 78(3): 482–526.

Harley, Heidi, and Elizabeth Ritter. 2002b. “Structuring the Bundle: A Universal Morphosyntactic Feature Geometry.” In: Horst J. Simon and Heike Wiese (eds.), *Pronouns: Grammar and Representation*, Amsterdam/ Philadelphia: John Benjamins, 23–39.

Harris, John. 1993. “The Grammar of Irish English.” In Milroy and Milroy (eds.), 139–186.

Harris, Martin. 1981. “It's I, It's Me: Further Reflections.” *Studia Anglica Posnaniensia* 13: 17–20

Harris, Martin. 1991. “Demonstrative Adjectives and Pronouns in a Devonshire Dialect.” In: Trudgill and Chambers (eds.), 20–28.

Haspelmath, Martin. 2008. “A Frequentist Explanation of Some Universals of Reflexive Marking.” *Linguistic Discovery* 6.1: 40–63. http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/2/xmlpage/1/issue/29.

Haspelmath, Martin. 2010. “Framework-Free Grammatical Theory.” In: Bernd Heine and Heiko Narrog (eds.), *The Oxford Handbook of Grammatical Analysis*. Oxford: Oxford University Press, 341-365.

Haspelmath, Martin, Matthew S. Dryer, David Gil and Bernard Comrie (eds.). 2005. *The World Atlas of Language Structures* (abbrev. WALS). Oxford: Oxford University Press.

Heacock, Paul (gen.ed.). 2008[2]. *Cambridge Dictionary of American English.* Cambridge: Cambridge University Press.

Helmbrecht, Johannes. 2005. “Politeness Distinctions in Pronouns.” In: Haspelmath et al. (eds.), 186–189.

Henderson, T. S. T. 1985. “Who Are *We*, Anyway? A Study of Personal Pronoun Systems.” *Linguistische Berichte* 98: 300–309.

Henry, Alison. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford Studies in Comparative Syntax. Oxford: Oxford University Press.

Hernández, Nuria. 2002. “A Context Hierarchy of Untriggered Self-Forms in English.” Ekkehard König and Volker Gast (eds.), *Reflexives and Intensifiers: The Use of Self-Forms in English*. Special issue 3/2002 of the *Zeitschrift für Anglistik und Amerikanistik (ZAA) – A Quarterly of Language, Literature and Culture*, 269–284.

Hernández, Nuria. 2006. *User’s Guide to FRED: Freiburg Corpus of English Dialects*. University of Freiburg online publications. http://www.freidok.uni-freiburg.de/volltexte/2489.

Heyvaert, Liesbet. 2008. “On the Constructional Semantics of Gerundive Nominalizations.” *Folia Linguistica* 42(1): 39–82.

Heyvaert, Liesbet, Hella Rogiers and Nadine Vermeylen. 2005. “Pronominal Determiners in Gerundive Nominalization: A 'Case' Study.” *English Studies* 86(1): 71–88.

Hilpert, Martin. 2006. “Distinctive Collexeme Analysis and Diachrony.” *Corpus Linguistics and Linguistic Theory* 2.2: 243–57.

Hinrichs, Lars, and Benedikt Szmrecsanyi. 2007. “Recent Changes in the Function and Frequency of Standard English Genitive Constructions: A Multivariate Analysis of Tagged Corpora.” *English Language and Linguistics* 11(3): 437–474.

Hogg, Richard M. (gen. ed.). 1992–99. *The Cambridge History of the English Language, 6 vols*. Cambridge: Cambridge University Press.

Holmes, Janet. 1995. *Men, Women and Politeness*. Real language series. London: Longman.

Honey, John. 1995. "A New Rule for the Queen and I?" *English Today* 11(4): 3–8.

Hope, Jonathan. 2000. "Rats, Bats, Sparrows and Dogs: Biology, Linguistics and the Nature of Standard English." In: Laura Wright (ed.), 49–56.

Householder, Fred W. 1987. "Some Facts about *Me* and *I*." *Language Research* 23(2): 163–184.

Howe, Stephen. 1996. *The Personal Pronouns in the Germanic Languages: A Study of Personal Pronoun Morphology and Change in the Germanic Languages from the First Records to the Present Day*. Studia Linguistica Germanica 43. Berlin/ New York: De Gruyter.

Huang, Yan. 2000. *Anaphora – A Cross-Linguistic Study*. Oxford Studies in Typology and Linguistic Theory. New York: Oxford University Press.

Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Huddleston, Rodney, and Geoffrey K. Pullum. 2005. *A Student's Introduction to English Grammar*. Cambridge: Cambridge University Press.

Hübler, Axel. 1998. *The Expressivity of Grammar: Grammatical Devices Expressing Emotion across Time*. Topics in Linguistics 25. Berlin/ New York: Mouton de Gruyter.

Hughes, Arthur, and Peter Trudgill. 1996[3]. *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles*. London: Arnold.

Hyman, Eric. 2006. "The All of You-All." *American Speech* 81(3): 325–331.

Hymes, Dell. 1997. "The Scope of Sociolinguistics." In: Nikolas Coupland and Adam Jaworski (eds.), *Sociolinguistics: A Reader and Coursebook*, Basingstoke: Macmillan, 12–22.

Ihalainen, Ossi. 1985. " 'He Took the Bottle and Put 'n in His Pocket': The Object Pronoun *It* in Present-Day Somerset." In: Wolfgang Viereck (ed.), *Focus on: England and Wales* (Varieties of English Around the World 4), Amsterdam/ Philadelphia: John Benjamins, 153–161.

Ihalainen, Ossi. 1991. "On Grammatical Diffusion in Somerset Folk Speech." In: Trudgill and Chambers (eds.), 104–119.

Ihalainen, Ossi. 1994. "The Dialects of England since 1776." In: Burchfield (ed.), 197–274.

Ingram, David. 1978. "Typology and Universals of Personal Pronouns." In: Joseph H. Greenberg (ed.), *Universals of Human Language*, *Vol. 3: Word Structure*, Stanford, CA: Stanford University Press, 213–247.

Jaffe, Alexandra. 2000. “Introduction: Non-Standard Orthography and Non-Standard Speech.” *Journal of Sociolinguistics* 4(4): 497–513.

Jespersen, Otto. 1928. *An International Language*. London: Allen & Unwin. Retrievable from http://interlanguages.net/AIL.html.

Jespersen, Otto. 1949. *A Modern English Grammar on Historical Principles, Vol. VII: Syntax*. Original 1909, completed and published by Niels Haislund. Copenhagen: Ejnar Munksgaard.

Ježek, Elisabetta, and Paolo Ramat. 2009. “On Parts-of-Speech Transcategorization.” *Folia Linguistica* 43(2): 391–416.

Johannessen, Janne Bondi. 1998. *Coordination*. Oxford Studies in comparative syntax. Oxford: Oxford University Press.

Johnson, Samuel. 1805[9]. *A Dictionary of the English Language*. London: Rivington.

Jonson, Ben. 1640/ 1972. *The English Grammar*. Facsimile. Menston: Scolar Press.

Joos, Martin. 1967. *The Five Clocks: A Linguistic Excursion into the Five Styles of English Usage*. New York: Harcourt, Brace and World.

Kaufmann, Ingrid, and Barbara Stiebels (eds.). 2002. *More than Words: A Festschrift for Dieter Wunderlich*. Studia Grammatica 53. Berlin: Akademie Verlag.

Keenan, Edward L., and Bernard Comrie. 1977. “Noun Phrase Accessibility and Universal Grammar.” *Linguistic Inquiry* 8: 63–99.

Kemmer, Suzanne. 1993. *The Middle Voice*. Typological Studies in Language 23. Amsterdam/ Philadelphia: John Benjamins.

Kemmer, Suzanne. 1994. “Middle Voice, Transitivity, and the Elaboration of Events.” In: Barbara Fox and Paul J. Hopper (eds.), *Voice: Form and Function*, Amsterdam/ Philadelphia: John Benjamins, 179–230.

Kemmer, Suzanne. 1995. “Emphatic and Reflexive -*Self*: Expectations, Viewpoint, and Subjectivity.” In: Stein and Wright (eds.), 55–82.

Kerswill, Paul, and Ann Williams. 1999. “Dialect Recognition and Speech Community Focusing in New and Old Towns in England: The Effects of Dialect Levelling, Demography and Social Networks.” In: Clive Upton and Katie Wales (eds.), *Dialectal Variation in English: Proceedings of the Harold Orton Centenary Conference 1998*, Leeds Studies in English 30, University of Leeds: School of English, 205–237.

Kimball, Geoffrey D. 1991. *Koasati Grammar*. Studies in the Anthropology of North American Indians. Lincoln: University of Nebraska Press.

Kiparsky, Paul. 2002. “Disjoint Reference and the Typology of Pronouns.” In: Kaufmann and Stiebels (eds.), 179–226.

Kiparsky, Paul. 2008. “Universals Constrain Change; Change Results in Typological Generalizations.” In: Jeff Good (ed.), *Language Universals and Language Change*, Oxford: Oxford University Press, 23–53.

Kiss, Tibor. 2001. “Anaphora and Exemptness: A Comparative Treatment of Anaphoric Binding in German and English.” In: Dan Flickinger and Andreas Kathol (eds.), *Proceedings of the 7th International HPSG Conference, UC Berkeley, 22–23 July, 2000*. Stanford, CA: Center for the Study of Language and Information. http://csli-publications.stanford.edu/.

Kjellmer, Göran. 1971. *Context and Meaning*. Gothenburg Studies in English 22. Stockholm: Almqvist & Wiksell.

Kjellmer, Göran. 1986. “ ‘Us Anglos Are a Cut Above the Field’: On Objective Pronouns in Nominative Contexts.” *English Studies* 67(5): 445–449.

Kjellmer, Göran. 2001. “ ‘It Comes Time’: A Look at Existential *It*.” *English Studies* 82(4): 328–335.

Klein, Lawrence. 1994. “‘Politeness’ as linguistic ideology in late seventeenth- and eighteenth-century England.” In: Stein and Tieken-Boon van Ostade (eds.), 31–50.

Koch, Peter, and Wulf Oesterreicher. 1985. “Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgebrauch.” *Romanistisches Jahrbuch* 36: 15–43.

Koch, Peter, and Wulf Oesterreicher. 1994. “Schriftlichkeit und Sprache.” In: Hartmut Günther and Otto Ludwig (eds.), *Handbuch Schrift und Schriftlichkeit I*, Berlin/ New York: Mouton de Gruyter, 587–604.

König, Ekkehard. 1991. *The Meaning of Focus Particles: A Comparative Perspective*. London: Routledge.

König, Ekkehard, and Peter Siemund. 1997. “On the Development of Reflexive Pronouns in English: A Case Study in Grammaticalization.” In: Uwe Böker and Hans Sauer (eds.), *Anglistentag 1996 Dresden – Proceedings of the Conference of the German Association of University Teachers of English*, *Vol. XVIII*, Trier: Wissenschaftlicher Verlag, 95–108.

König, Ekkehard, and Peter Siemund. 2000. “Intensifiers and Reflexives: A Typological Perspective.” In: Frajzyngier and Curl (eds.), 41–74.

Koktova, Eva. 1999. *Word-Order Based Grammar*. Trends in Linguistics – Studies and Monographs 121. Berlin/ New York: Mouton de Gruyter.

Koptjevskaja-Tamm, Maria. 2005. “Action Nominal Constructions.” In: Haspelmath et al. (eds.), 254–257.

Kortmann, Bernd. 1999. “Typology and Dialectology.” In: Bernard Caron (ed.),

Proceedings of the 16th International Congress of Linguists, Paris 1997, Elsevier Science: CD-ROM article 0060.

Kortmann, Bernd. 2001. "In the Year 2525… Reflections on the Future Shape of English." *Anglistik: International Journal of English Studies* 12: 97–114.

Kortmann, Bernd. 2002. "New Prospects for the Study of English Dialect Syntax: Impetus from Syntactic Theory and Language Typology." In: Barbiers et al. (eds.), 185–213.

Kortmann, Bernd (ed.). 2004b. *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*. Berlin/ New York: Mouton de Gruyter.

Kortmann, Bernd, and Edgar Schneider, in collaboration with Kate Burridge, Rajend Mesthrie and Clive Upton (eds.). 2004. *A Handbook of Varieties of English. Vol. 2: Morphology and Syntax*. Berlin/ New York: Mouton de Gruyter.

Kortmann, Bernd, and Benedikt Szmrecsanyi. 2004. "Global Synopsis: Morphological and Syntactic Variation in English." In: Kortmann and Schneider (eds.), 1142–1202.

Kristensson, Gillis. 1997. "The Dialects of Middle English." In: Raymond Hickey and Stanisław Puppel (eds.), *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on His 60th Birthday*, Berlin/ New York: De Gruyter, 655–664.

Krug, Manfred. 1998. "String Frequency: A Cognitive Motivating Factor in Coalescence, Language Processing, and Linguistic Change." *Journal of English Linguistics* 26(4): 286–320.

Krug, Manfred. 2003. "Frequency as a Determinant in Grammatical Variation and Change." In: Rohdenburg and Mondorf (eds.), 7–67.

Kruisinga, Etsko. 1905. *A Grammar of the Dialect of West Somerset: Descriptive and Historical*. Bonner Beiträge zur Anglistik 18. Bonn: Hanstein.

Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse, and Empathy*. Chicago: Chicago University Press.

Labov, William. 1966. "The Linguistic Variable as a Structural Unit." *Washington Linguistic Review* 3: 4–22.

Labov, William. 1970. "The Logic of Non-Standard English." In: Frederick Williams (ed.), *Language and Poverty: Perspectives on a Theme*, Chicago: Markham, 153–187.

Labov, William. 1971. "The Notion of System in Creole Studies." In: Dell Hymes (ed.), *Pidginization and Creolization of Languages – Proceedings of a Conference Held at the University of the West Indies at Mona, Jamaica*, Cambridge: Cambridge University Press, 447–472.

Labov, William. 1972. "Some principles of linguistic methodology." *Language in Society* I: 97–120.

Labov, William. 1989. “The Child as Linguistic Historian.” *Language Variation and Change* 1: 85–94.

Labov, William. 1996. “When Intuitions Fail.” In: L. McNair, K. Singer, L. Dolbrin and M. Aucon (eds.), *Papers from the Parasession on Theory and Data in Linguistics*, Chicago Linguistic Society 32, 77–106.

Labov, William. 1997. “Some Further Steps in Narrative Analysis.” *Journal of Narrative and Life History* 7: 395–415.

Langacker, Ronald W. 1990. *Concept, Image and Symbol: The Cognitive Basis of Grammar*. Berlin/ New York: Mouton de Gruyter.

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar, Vol.II: Descriptive Application*. Stanford, CA: Stanford University Press.

Lass, Roger. 1987. *The Shape of English: Structure and History*. London: Dent.

Lass, Roger. 1990. “How to do Things with Junk: Exaptation and Language Evolution.” *Journal of Linguistics* 26(1): 79–102.

Lass, Roger (ed.). 1999a. *The Cambridge History of the English Language, Vol. III: 1476–1776*, Cambridge: Cambridge University Press.

Lass, Roger. 1999b. “Phonology and Morphology.” In: Roger Lass (ed.), 56–186.

Leech, Geoffrey N., and Jan Svartvik. 1994[2]. *A Communicative Grammar of English*. London: Longman.

Lees, Robert B. 1960. *The Grammar of English Nominalizations*. Bloomington: Indiana University Press.

Leonard, Sterling Andrus. 1929 (1962, reissued). *The Doctrine of Correctness in English Usage, 1700–1800*. Madison: University of Wisconsin Press.

Levinson, Stephen C. 1987. “Pragmatics and the Grammar of Anaphora: A Partial Pragmatic Reduction of Binding and Control Phenomena.” *Journal of Linguistics* 23: 379–434.

Levinson, Stephen C. 1997. *Pragmatics*. Cambridge: Cambridge University Press.

Lowsley, Barzillai. 1888. *A Glossary of Berkshire Words and Phrases*. English Dialect Society, 56. London: Trübner.

Lowth, Robert. 1762 (1967 facsimile reprint). *A Short Introduction to English Grammar*. Menston: Scolar Press.

McLennan, Sean. 2004. “Guy, Guys, and Gender Neutrality.” *Proceedings from the Annual Meeting of the Chicago Linguistic Society* 40(1): 211–219.

Marshall, William. 1787/ 1788/ 1789/ 1790/ 1796. *Provincialisms of East Norfolk/ East Yorkshire/ the Vale of Glocester/ the Midland Counties/ West Devonshire*. Reprinted

Glossaries, 1873. English Dialect Society, 1. London: Trübner.

Mathiot, Madeleine, and Marjorie Roberts. 1979. "Sex roles as revealed through referential gender in American English." In M. Mathiot (ed.), *Ethnolinguistics: Boas, Sapir and Whorf Revisited*, The Hague: Mouton, 1–47.

Mayhew, Anthony L. 1908. " 'Dear': 'O Dear No!' " *Notes and Queries* (s10–X, 257): 434–435.

Miethaner, Ulrich. 2000. "Orthographic Transcriptions of Non-Standard Varieties: The Case of Earlier African-American English." *Journal of Sociolinguistics* 4(4): 534–560.

Miller, Jim. 1993. "The Grammar of Scottish English." In: Milroy and Milroy (eds.), 99–138.

Miller, Jim. 2004. "Scottish English: Morphology and Syntax." In: Kortmann and Schneider (eds.), 47–72.

Milroy, James, and Lesley Milroy (eds.). 1993. *Real English: The Grammar of English Dialects in the British Isles*. London: Longman.

Mitchell, Bruce. 1985. *Old English Syntax, Vol. I: Concord, the Parts of Speech, and the Sentence*. Oxford: Clarendon.

Montgomery, Michael B. 2004. "Appalachian English: Morphology and Syntax." In: Kortmann and Schneider (eds.), 245–280.

Morrison, Estelle Rees. 1926. " 'You all' and 'we all'." *American Speech* 2(3): 133.

Morse-Gagne, Elise. 2003. *Viking Pronouns in England: Charting the Course of THEY, THEIR, and THEM*. Ph.D. thesis, University of Pennsylvania. Available from ProQuest, Paper AAI3087442.

Mühlhäusler, Peter. 2004. "Norfolk Island-Pitcairn English (Pitkern Norfolk): Morphology and Syntax." In: Kortmann and Schneider (eds.), 789–804.

Mühlhäusler, Peter, and Rom Harré. 1990. *Pronouns and People: The Linguistic Construction of Social and Personal Identity*. Oxford: Blackwell.

Murray, James A.H., Henry Bradley, W. A. Craigie and C. T. Onions (eds.). 1961/ 1989[2]/ online edition. *The Oxford English Dictionary* (abbrev. OED). Oxford: Clarendon Press.

Murray, Lindley. 1816[3]. *An English Grammar*. London: Longman.

Nevalainen, Terttu. 2000. "Gender Differences in the Evolution of Standard English: Evidence from the Corpus of Early English Correspondance." *Journal of English Linguistics* 28(1): 38–59.

Nevalainen, Terttu. 2006. *An Introduction to Early Modern English*. Edinburgh: Edinburgh University Press.

Nevalainen, Terttu, and Helena Raumolin-Brunberg. 1994. "*Its* Strength and the Beauty of

It: The Standardization of the Third Person Neuter Possessive in Early Modern English." In: Stein and Tieken-Boon van Ostade (eds.), 171–216.

Newmeyer, Frederick J. 2004. "Typological Evidence and Universal Grammar." *Studies in Language* 28(3): 527–548.

Nichols, Johanna, and Balthasar Bickel. 2005. "Possessive Classification." In: Haspelmath et al. (eds.), 242–245.

Noyer, Robert Rolf. 1992. *Features, Positions and Affixes in Autonomous Morphological Structure*. Ph.D. thesis, Massachusetts Institute of Technology. http://hdl.handle.net/1721.1/12895.

Oakes, Michael P. 1998. *Statistics for Corpus Linguistics*. Edinburgh Textbooks in Empirical Linguistics. Edinburgh: Edinburgh University Press.

Olson, David R. 1994. *The World on Paper: The Conceptual and Cognitive Implications of Writing and Reading*. Cambridge: Cambridge University Press.

Orton, Harold. 1962. *Survey of English Dialects: Introduction*. Leeds: E. J. Arnold and Son.

Orton, Harold, and Wilfrid J. Halliday (Vol. 1)/ Michael V. Barry (Vol. 2)/ Philip M. Tilling (Vol. 3)/ Martyn F. Wakelin (Vol. 4). 1962–1971. *Survey of English Dialects: The Basic Material, 4 vols* (abbrev. SED). Leeds: E. J. Arnold and Son.

Orton, Harold, Stewart Sanderson and John Widdowson. 1978. *The Linguistic Atlas of England* (abbrev. LAE). London: Croom Helm.

Osselton, Noel. 1988. "Thematic Genitives." In: Graham Nixon and John Honey (eds.), *An Historic Tongue: Studies in English Linguistics in Memory of Barbara Strang*, London: Routledge, 138–144.

Parker, Frank, Kathryn Riley and Charles Meyer. 1988. "Case Assignment and the Ordering of Constituents in Coordinate Constructions." *American Speech* 63(3): 214–233.

Parker, Frank, Kathryn Riley and Charles Meyer. 1990. "Untriggered Reflexive Pronouns in English." *American Speech* 65(1): 50–69.

Patrick, Peter L. 2004. "Jamaican Creole: Morphology and Syntax." In: Kortmann and Schneider (eds.), 407–438.

Pawley, Andrew, and Frances H. Syder. 1983. "Natural Selection in Syntax: Notes on Adaptive Variation and Change in Vernacular and Literary Grammar." *Journal of Pragmatics* 7: 551–79.

Penke, Martina, and Anette Rosenbach. 2004. "What Counts as Evidence in Linguistics? An Introduction." *Studies in Language* 28(3): 480–526.

Petyt, Ken. 1985. *Dialect and Accent in Industrial West Yorkshire*. Amsterdam/ Philadelphia: John Benjamins.

Phythian, Brian Arthur. 1988. *Correct English*. Teach Yourself books. Sevenoaks: Hodder and Stoughton.

Pickett, Joseph P. (ex.ed.). 2000[4]. *The American Heritage Dictionary of the English Language*. Boston/ New York: Houghton Mifflin.

Pietsch, Lukas. 2005. *Variable Grammars: Verbal Agreement in Northern Dialects of English*. Linguistische Arbeiten 496. Tübingen: Niemeyer.

Pinker, Steven. 1994. *The Language Instinct: The New Science of Language and Mind*. New York: Morrow.

Priestley, Joseph. 1762. *A Course of Lectures on the Theory of Language and Universal Grammar*. Reprinted in: Joseph Priestley, 1833, *English Grammar: Lectures on the Theory of Language and Universal Grammar, and an Oratory and Criticism*, London: Rowland Hunter.

Poutsma, Hendrik. 1929[2]. *A Grammar of Late Modern English*. Groningen: Noordhoff.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1972. *A Grammar of Contemporary English*. London: Longman.

Quirk, Randolph, and Sidney Greenbaum. 1984. *A University Grammar of English, Based on A Grammar of Contemporary English*. London: Longman.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Reaser, Jeffrey, and Benjamin Torbert. 2004. "Bahamian English: Morphology and Syntax." In: Kortmann and Schneider (eds.), 391–406.

Redfern, Richard K. 1994. "Is *Between You and I* Good English?" In: Greta D. Little and Michael Montgomery (eds.), *Centennial Usage Studies*, American Dialect Society 78, Tuscaloosa/ London: Alabama Press, 187–193.

Reuland, Eric J. 1983. "Governing *-ing*." *Linguistic Inquiry* 14: 101–136.

Richardson, Malcolm. 1980. "Henry V, the English Chancery and Chancery English." *Speculum* 55: 726–750.

Richards, Marc. 2008. "Defective Agree, Case Alternations, and the Prominence of Person." In: M. Richards, and A. L. Malchukov (eds.). *Linguistische Arbeitsberichte* 86, Leipzig: University of Leipzig Press, 137–161.

Riley, Kathryn, and Frank Parker. 1998. *English Grammar: Prescriptive, Descriptive, Generative, Performance*. Boston: Allyn and Bacon.

Rissanen, Matti. 1999. "Syntax." In: Roger Lass (ed.), 187–331.

Rocha, Marco. 2000. “A Corpus-Based Study of Anaphora in English and Portuguese.” In: Simon Botley and A. M. McEnery (eds.), *Corpus-Based and Computational Approaches to Discourse Anaphora*, Amsterdam/ Philadelphia: John Benjamins, 81–94.

Rohdenburg, Günter, and Britta Mondorf (eds.). 2003. *Determinants of Grammatical Variation in English*. Topics in English Linguistics 43. Berlin/ New York: Mouton de Gruyter.

Rosenbach, Anette. 2002. *Genitive Variation in English: Conceptual Factors in Synchronic and Diachronic Studies*. Berlin/ New York: Mouton de Gruyter.

Rosenbach, Anette. 2003. “Aspects of Iconicity and Economy in the Choice between the S-Genitive and the Of-Genitive in English.” In: Rohdenburg and Mondorf (eds.), 379–411.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Bloomington: Indiana University Linguistics Club. (Reprinted 1986 as *Infinite Syntax*, Norwood, N.J.: Ablex.)

Rot, Sándor. 1994. “English Dialectology and Problems of Historical Stylistics of Decoding.” In: Wolfgang Viereck (ed.), *Verhandlungen des Internationalen Dialektologenkongresses Bamberg 1990, Bd. 3: Regionalsprachliche Variation, Umgangs- und Standardsprachen* (Zeitschrift für Dialektologie und Linguistik, Beiheft 76), Stuttgart: Franz Steiner, 475–487.

Safir, Kenneth J. 2004. *The Syntax of Anaphora*. Oxford: Oxford University Press.

Saha, Prosanta K. 1987. “Strategies of Reflexivization in American English.” *American Speech* 62(3): 211–234.

Salzmann, Martin. 2006. *Resumptive Prolepsis: A Study in Indirect A'-Dependencies*. Utrecht: LOT Dissertation Series 136.

Sampson, Geoffrey, David Gil and Peter Trudgill (eds.). 2009. *Language complexity as an evolving variable*. Oxford: Oxford University Press.

Sankoff, David (ed.). 1978. *Linguistic Variation: Models and Methods*. New York/ London: Academic Press.

Sankoff, David, and Suzanne Laberge. 1978. “The Linguistic Market and the Statistical Explanation of Variability.” In: Sankoff (ed.), 239–250.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and World.

Sasse, Hans-Jürgen. 1993. “Syntactic Categories and Subcategories.” In: Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, Theo Vennemann (eds.) *Syntax: An International Handbook of Contemporary Research*, Handbooks of Linguistics and Communication 9, Berlin: De Gruyter, 646–686.

Schachter, Paul. 1985. “Pronouns and Other Pro-Forms.” In: Timothy Shopen (ed.), *Language Typology and Syntactic Description, I: Clause Structure*, Cambridge: Cambridge University Press, 24–35.

Schmidt, Gernot. 1978. *Stammbildung und Flexion der Indogermanischen Personalpronomina*. Wiesbaden: Otto Harrassowitz.

Schmidt, Karl Horst. 1994. “Zum Personalpronomen und der Kategorie ‘Person’ im Kartvelischen und Indogermanischen.” *Historische Sprachforschung* 107: 179–193.

Schütze, Carson T. 2001. “On the Nature of Default Case.” *Syntax* 4(3): 205–238.

Sebba, Mark. 2004. “British Creole: Morphology and Syntax.” In: Kortmann and Schneider (eds.), 196–208.

Shetter, William Z. 2000. “My Dog Died on Me – Pronouns of ‘Caring’.” In: Shetter (ed.), *Language Miniatures: A Selection of Essays*, Indiana University Custom Publishing. http://mypage.iu.edu/~shetter/miniatures/ethical.htm.

Shields, Kenneth. 1994. “The Role of Deictic Particles in the IE Personal Pronoun System.” *WORD* 45(3): 307–315.

Siemund, Peter. 2003. “Varieties of English from a Cross-Linguistic Perspective: Intensifiers and Reflexives.” In: Rohdenburg and Mondorf (eds.), 479–506.

Siemund, Peter. 2008. *Pronominal Gender in English – A Study of English Varieties from a Cross-Linguistic Perspective*. Routledge Studies in Germanic Linguistics 10. London: Routledge.

Siewierska, Anna. 2004. *Person*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Siewierska, Anna. 2005. “Gender Distinctions in Independent Personal Pronouns.” In: Haspelmath et al. (eds.), 182–185.

Sigurðsson, Halldór Ármann. 1992. “The Case of Quirky Subjects.” *Working Papers in Scandinavian Syntax* 49: 1–26.

Sinclair, John McHardy. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Singler, John Victor. 2004. “Liberian Settler English: Morphology and Syntax.” In: Kortmann and Schneider (eds.), 879–897.

Siren, Kathleen A., and Kim A. Wilcox. 1990. “The Utility of Phonetic versus Orthographic Transcription Methods.” *Child Language Teaching and Therapy* 6 (2): 127–146.

Smith, John Charles. 2006. “How to Do Things without Junk: The Refunctionalization of a Pronominal Subsystem between Latin and Romance.” In: Jean-Pierre Y. Montreuil and

Chiyo Nishida (ed.), *New Perspectives on Romance Linguistics: Selected Papers from the 35th Linguistic Symposium on Romance Languages (LSRL), February 2005, Austin, Texas*, *Vol. II: Phonetics, Phonology and Dialectology*, Amsterdam/ Philadelphia: John Benjamins, 183–205.

Stefanowitsch, Anatol. 2003. "Constructional Semantics as a Limit to Grammatical Alternation: The Two Genitives of English." In: Rohdenburg and Mondorf (eds.), 413–443.

Stefanowitsch, Anatol, and Stefan Th. Gries. 2003. "Collostructions: Investigating the Interaction of Words and Constructions." *International Journal of Corpus Linguistics* 8.2: 209–243.

Stein, Dieter. 1994. "Sorting out the Variants: Standardization and Social Factors in the English Language 1600–1800." In: Stein and Tieken-Boon van Ostade (eds.), 1–17.

Stein, Dieter. 1997. "Syntax and Varieties." In: Cheshire and Stein (eds.), 35–50.

Stein, Dieter, and Ingrid Tieken-Boon van Ostade (eds.). 1994. *Towards a Standard English 1600–1800*. Berlin/ New York: Mouton de Gruyter.

Stein, Dieter, and Susan Wright (eds.). 1995. *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge: Cambridge University Press.

Stern, Nancy. 2004. "The Semantic Unity of Reflexive, Emphatic, and Other -*Self* Pronouns." *American Speech* 79(3): 270–280.

Stratton, Clarence. 1949. *Guide to Correct English*. New York: Whittlesey House, McGraw-Hill Book Company.

Sundby, Bertil, Anne K. Bjørge and Kari E. Haugland. 1991. *A Dictionary of English Normative Grammar, 1700–1800*. Studies in the History of the Language Sciences 63. Amsterdam/ Philadelphia: John Benjamins.

Swan, Michael. 1992. *Practical English Usage*. Oxford: Oxford University Press.

Sweet, Henry. 1900. *A New English Grammar, Logical and Historical. Part I: Introduction, Phonology and Accidence*. Oxford: Clarendon.

Szemerényi, Oswald J. L. 1996. *Introduction to Indo-European Linguistics*. Oxford: Clarendon Press.

Szmrecsanyi, Benedikt. 2006. *Morphosyntactic Persistence in Spoken English: A Corpus Study at the Intersection of Variationist Sociolinguistics, Psycholinguistics, and Discourse Analysis*. Trends in Linguistics – Studies and Monographs 177. Berlin/ New York: Mouton de Gruyter.

Szmrecsanyi, Benedikt. forthcoming. *Morphosyntactic Variation in British English Dialects: Exploring Dialect Grammars in the Aggregate Perspective*. Cambridge:

Cambridge University Press.

Szmrecsanyi, Benedikt, and Nuria Hernández. 2007. *Manual of Information to Accompany the Freiburg Corpus of English Dialects Sampler FRED-S*. University of Freiburg online publications. http://www.freidok.uni-freiburg.de/volltexte/2859.

Szmrecsanyi, Benedikt, and Lars Hinrichs. 2008. "Probabilistic Determinants of Genitive Variation in Spoken and Written English: A Multivariate Comparison across Time, Space, and Genres." In: Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta and Minna Korhonen (eds.), *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*, Amsterdam/ Philadelphia: John Benjamins, 291–309.

Szmrecsanyi, Benedikt, and Bernd Kortmann. 2009. "Between Simplification and Complexification: Non-Standard Varieties of English around the World." In: Sampson et al. (eds.), 64–79.

Tannen, Deborah. 2007[2]. *Talking Voices – Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Taylor, John R. 1996. *Possessives in English: An Exploration in Cognitive Grammar*. Oxford: Oxford University Press.

The British National Corpus, Version 2. 2001. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. http://www.natcorp.ox.ac.uk/.

Thompson, Sandra A. 1973. "On Subjectless Gerunds in English." *Foundations of Language* 9: 3, 374–383.

Tieken-Boon van Ostade, Ingrid. 1994. "Standard and Non-Standard Pronominal Usage in English, with Special Reference to the Eighteenth Century." In: Stein and Tieken-Boon van Ostade (eds.), 217–242.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Studies in Corpus Linguistics 6. Amsterdam/ Philadelphia: John Benjamins.

Traugott, Elizabeth Closs. 1972. *A History of English Syntax*. New York: Holt, Rinehart, and Winston.

Traugott, Elizabeth Closs. 1992. "Syntax." In: Richard Hogg (ed.), *The Cambridge History of the English Language I: Old English*, Cambridge: Cambridge University Press, 168–289.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, Peter. 1999[2]. *The Dialects of England*. Oxford: Blackwell.

Trudgill, Peter. 2003. *A Glossary of Sociolinguistics*. Edinburgh University Press.

Trudgill, Peter. 2004a. "The Dialect of East Anglia: Morphology and Syntax." In:

Kortmann and Schneider (eds.), 142–153.

Trudgill, Peter. 2004b[2]. *Dialects*. Language Workbooks series. London: Routledge.

Trudgill, Peter. 2009. "Sociolinguistic Typology and Complexification." In: Sampson et al. (eds.), 98–109.

Trudgill, Peter, and Jack K. Chambers (eds.). 1991. *Dialects of English: Studies in Grammatical Variation*. London: Longman.

Upton, Clive, David Parry and J. D. A. Widdowson. 1994. *Survey of English Dialects: The Dictionary and Grammar* (abbrev. SED). London: Routledge.

Upton, Clive, Stewart Sanderson and John Widdowson. 1987. *Word Maps: A Dialect Atlas of England* (abbrev. WM). London: Croom Helm.

Vallins, George Henry. 1952. *Good English – How to Write It*. London: André Deutsch.

Van Gelderen, Elly. 2000a. *A History of English Reflexive Pronouns: Person,* Self, *and Interpretability*. Amsterdam/ Philadelphia: John Benjamins.

Van Gelderen, Elly. 2000b. "Bound pronouns and non-local anaphors: The case of earlier English." In: Frajzyngier and Curl (eds.), 187–226.

Van Marle, Jaap. 1997. "Dialect versus Standard Language: Nature versus Culture." In: Cheshire and Stein (eds.), 13–34.

Viereck, Wolfgang. 1975. *Lexikalische und grammatische Ergebnisse des Lowman-Survey von Mittel- und Südengland, 2 vols.* (abbrev. LSV). München: Fink.

Viereck, Wolfgang. 1991. *The Computer Developed Linguistic Atlas of England* (abbrev. CLAE). In collaboration with Heinrich Ramisch; computational production Harald Händler et al. Tübingen: Niemeyer.

Visser, Frederikus Th. 1963–73. *An Historical Syntax of the English Language, 4 vols*. Leiden: Brill.

Wagner, Susanne. 2002. "We don' say she, do us? Pronoun Exchange – a feature of English dialects?" Manuscript retrievable from http://www.tu-chemnitz.de/phil/english/ling/download/Wagner2002PronounExchange.pdf.

Wagner, Susanne. 2004a. *Gender in English Pronouns: Myth and Reality*. University of Freiburg online publications. http://www.freidok.uni-freiburg.de//volltexte/1412.

Wagner, Susanne. 2004b. "English Dialects in the Southwest: Morphology and Syntax." In: Kortmann and Schneider (eds.), 154–174.

Wagner, Susanne. 2004c. " 'Gendered' Pronouns in English Dialects – A Typological Perspective." In: Kortmann (ed.), 479–496.

Wakelin, Martyn F. 1983. "The Stability of English Dialect Boundaries." *English World-Wide* IV(1): 1–15.

Wakelin, Martyn F. 1984. “Rural Dialects in England.” In: Peter Trudgill (ed.), *Language in the British Isles*, Cambridge: Cambridge University Press, 70–93.

Wakelin, Martyn F. 1986. *The Southwest of England*. (Varieties of English Around the World 5.) Amsterdam/ Philadelphia: John Benjamins.

Wales, Katie. 1994. “Royalese: The Rise and Fall of ‘The Queen’s English’.” *English Today* 10(3): 3–10.

Wales, Katie. 1996. *Personal Pronouns in Present-Day English*. Cambridge: Cambridge University Press.

Wales, Katie. 2002. “ ‘Everyday Verbal Mythology’: Pronouns and Personification in Non-Literary English.” In: Sybil Scholz, Monika Klages, Evelyn Hantson and Ute Römer (eds.), *Language, Context and Cognition: Papers in Honour of Wolf-Dietrich Bald's 60th Birthday*, München: Langenscheidt-Longman, 327–335.

Werner, Abraham. 1973. “The ethic dative in German.” In: Ferenc Kiefer and Nicolas Ruwet (eds.), *Generative Grammar in Europe* (Foundations of Language, Supplementary Series 13), Dordrecht: Reidel.

Weiß, Helmut. 2004. “A Question of Relevance: Some Remarks on Standard Languages.” *Studies in Language* 28(3): 648–674.

Wierzbicka, Anna. 1996. *Semantics – Primes and Universals*. Oxford: Oxford University Press.

Winford, Donald, and Bettina Migge. 2004. “Surinamese Creoles: Morphology and Syntax.” In: Kortmann and Schneider (eds.), 482–516.

Wolfram, Walt. 2004. “Urban African American Vernacular English: Morphology and Syntax.” In: Kortmann and Schneider (eds.), 319–340.

Woolford, Ellen. 2006. “Lexical Case, Inherent Case, and Argument Structure.” *Linguistic Inquiry* 37(1): 111–130.

Wright, Joseph (ed.). 1898–1905. *The English Dialect Dictionary: Being the Complete Vocabulary of All Dialect Words Still in Use, or Known to Have Been in Use During the Last Two Hundred Years, 6 volumes*. London: Frowde.

Wright, Joseph. 1905. *The English Dialect Grammar: Comprising the Dialects of England, of the Shetland and Orkney Islands, and of Those Parts of Scotland, Ireland and Wales Where English Is Habitually Spoken*. Oxford: Frowde.

Wright, Laura (ed.). 2000. *The Development of Standard English, 1300–1800: Theories, Descriptions, Conflicts*. Cambridge: Cambridge University Press.

Wright, Susan. 1997. “ ‘Ah’m Going for to Give Youse a Story Today’: Remarks on Second Person Plural Pronouns in Englishes.” In: Cheshire and Stein (eds.), 170–184.

Yamamoto, Mutsumi. 1999. *Animacy and Reference: A Cognitive Approach to Corpus Linguistics*. Studies in Language, Companion Series, 46. Amsterdam/ Philadelphia: John Benjamins.

Young, David J. 1984. *Introducing English Grammar*. London: Hutchinson.

Zipf, George Kingsley. 1949. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Hafner.

Zribi-Hertz, Anne. 1989. "Anaphor Binding and Narrative Point of View: English Reflexive Pronouns in Sentence and Discourse." *Language* 65(4): 695–727.

Zwart, Jan-Wouter. 2005. "A Comparative Approach to Syntactic Change in the History of English." *English Language and Linguistics* 9(1): 157–179.

Deutsche Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der funktionalen Vielfalt ausgewählter Pronomenformen im gesprochenen Britischen Englisch. Im Mittelpunkt stehen nicht-standardmäßige Verwendungen von Subjekt-, Objekt- und Reflexivformen im spontansprachlichen Dialog (*I*, *me*, *myself*, etc.). In der Standardvarietät, genauer in präskriptiven Grammatiken des geschriebenen Standardenglisch, ist die Verwendung dieser Wortformen isomorph, das heißt es besteht eine Eins-zu-eins-Zuordnung zwischen Form und Funktion: Subjektformen erfüllen Subjektfunktionen, Objektformen erfüllen Objektfunktionen, und Reflexivformen finden Verwendung als Reflexivpronomen (*she saw herself in the mirror*) oder emphatische Pronomen (*she saw it herself*). Abweichungen von dieser Norm haben sowohl in der einschlägigen Literatur als auch in der allgemeinen Öffentlichkeit (z.B. Internetforen) bereits vielfach Beachtung gefunden. Vor allem einige neuere Arbeiten aus den Bereichen der Varietätenlinguistik und Diskursanalyse beschäftigen sich mit möglichen Erklärungen unterschiedlicher Phänomene, die im Gegensatz zur ursprünglich in der generativen Linguistik verteidigten komplementären Verteilung von Pronomen und Anaphern, sowie Subjekt- und Objektformen, stehen. Dabei dominieren, trotz aktueller Verfügbarkeit entsprechender Daten, weiterhin intuitiv-theoretische Erklärungsmodelle, welche die Gültigkeit präskriptiver, ebenfalls auf intuitiver Logik basierender Normen einschränken, bisher jedoch keine allgemein akzeptierte Alternative bieten.

Aus dieser Situation heraus ergeben sich für die vorliegende Arbeit folgende Zielsetzungen. Erstens: Das Erfassen des Themas in seiner Gesamtheit, d.h. die phänomenübergreifende, überregionale Beschreibung des pronominalen Paradigmas in seiner tatsächlichen Verwendung. Zweitens: Das Einordnen der empirischen Ergebnisse in den breiteren Zusammenhang, inklusive historischer und typologischer Zusammenhänge, sowie die Erörterung der beobachteten Phänomene unter Einbeziehung bereits vorhandener Studien und Erklärungsansätze. Dabei stehen zwei Dinge im Vordergrund: Einerseits die bislang ausstehende empirisch begründete Analyse einzelner Phänomene; andererseits das Herausstellen allgemeiner distributioneller Tendenzen, welche sich über unterschiedliche Phänomene und regionale Varietäten erstrecken.

Zum Erreichen der genannten Ziele wurde ein korpusgesteuerter, deskriptiver Ansatz gewählt. Dieser unterscheidet sich vom korpusbasierten Ansatz darin, dass die vorhandenen Daten nicht der Überprüfung oder Veranschaulichung bereits vorhandener

Theorien dienen, sondern selbst den Untersuchungsgegenstand vorgeben. Konkret bedeutet dies, dass durch Extraktion und Klassifizierung der entsprechenden Wortformen alle nicht-standardmäßigen Vorkommen im Korpus identifiziert und anschließend analysiert wurden. Diese Vorgehensweise ermöglicht es, die allgemeine Verwendung der untersuchten Pronomen und das vorhandene Ausmaß an Variation in seiner Gesamtheit zu erfassen.

Die Arbeit untergliedert sich in drei Hauptabschnitte: Einleitung (I), Empirische Analysen (II) und Zusammenschau (III).

Teil I führt den Leser in die Vorgehensweise und Zielsetzung der Arbeit ein, die angewandte Terminologie, sowie grundlegende Leitlinien zur Kategorisierung empirischer Vorkommen. Es wird kurz auf die Bedeutung spontansprachlicher Daten eingegangen, sowie auf die für die nachfolgenden Analysen relevanten historischen Hintergründe.

Teil II beinhaltet die zentralen empirischen Analysen. Diese basieren auf den England-Daten des am Englischen Seminar der Universität Freiburg erstellten FRED Korpus (Freiburg Corpus of English Dialects), eine umfangreiche Sammlung von Interviews, welche ursprünglich die historisch orientierte Befragung von einheimischen Zeitzeugen zum Ziel hatten (‘oral history’). Die so entstandenen Unterhaltungen eignen sich sehr gut für dialektologische Studien. Sie beinhalten spontansprachliche Daten traditioneller Dialekte, deren Sprecher sich stark mit ihrer nativen Umgebung identifizieren und im Laufe ihres Lebens allgemein weniger als andere Sprecher mit sprachbeinflussenden Faktoren in Berührung kamen (minimale Schulbildung, geringe Mobilität, etc.). Die große Datenmenge von 1,5 Millionen Wörtern (nur der England-Teil) ermöglicht eingehende Untersuchungen hoch- wie auch niedrigfrequenter Phänomene. Die orthographischen Transkripte eigenen sich für gängige Konkordanz-Programme, und die vorhandenen Tonaufnahmen ermöglichen eine zusätzliche Überprüng der Texte (siehe beiliegende CD-ROM). Untersuchungsgegenstand sind alle im Korpus nachgewiesenen Phänomene, die auf Auswechselbarkeit unterschiedlicher Formen hinweisen. So befassen sich einzelne Kapitel zum Beispiel mit sogenannten Gendered Pronouns, dem vieldiskutierten Pronoun Exchange, der Kasusvariation in Partizipial- und Gerundialkonstruktionen, und der Verteilung der zwei Genitive im Englischen (*s*- und *of*-genitive). Weitere Themen sind Pronomen, die Teil von Nominalphrasen sind (sogenannte ‘qualified pronouns’ wie *us girls*), und pleonastische Pronomen, welche in vielen Grammatiken als unnötig oder gar unlogisch bezeichnet werden (z.B. offenes *you* in Imperativen oder pleonastische Reflexiva). Die letzten zwei empirischen Kapitel befassen sich mit der Verwendung von

Personalpronomen in Funktionsbereichen anderer pronominaler Kategorien anstelle von Possessiv- und Demonstrativpronomen. Jedes Kapitel schließt mit einer kurzen themenspezifischen Zusammenfassung.

Teil III schließlich bietet eine zusammenfassende Auswertung der mannigfaltigen Analyseergebnisse, einschließlich der sich hieraus ergebenden Implikationen für die Untersuchung morphosyntaktischer Variation im Allgemeinen, und nicht-standardmäßiger Verwendung von Pronomen im Besonderen. Rückblickend richtet sich das Augenmerk noch einmal auf die folgenden Fragen: Welche üblicherweise mit traditionellen Britischen Dialekten in Verbindung gebrachten nicht-standardmäßigen Verwendungen wurden in den Korpusdaten gefunden? Wie sind diese regional verteilt? In wiefern stimmen die empirischen Ergebnisse mit der bisherigen Literatur überein? Welche sprachlichen und außersprachlichen Faktoren beeinflussen effektiv die Wahl bestimmter Pronomenformen im spontanen müdlichen Dialog, und ist diese Wirkung systematisch?

Die Antwort auf diese Fragen stellt sich, in gekürzter Form, wie folgt dar. Beinahe alle untersuchten Nicht-Standard-Phänomene sind in den verschiedenen Dialektregionen vertreten und sind daher als allgemeine Merkmale von gesprochenem Englisch zu sehen. Dies steht teilweise im Widerspruch zu früheren Studien und hebt die Bedeutung überregionaler Untersuchungen hervor. Eine Ausnahme ist hierbei die nicht-standardmäßige Verwendung von Subjektformen, welche charakteristisch für den Südwesten des Landes ist. Selbst hier spricht die Verteilung verschiedener Formen jedoch deutlich gegen den andernorts proklamierten Isomorphismus. Die untersuchten Phänomene werden charakteristischerweise nicht kategorisch verwendet; manche Textabschnitte beinhalten sogar semantisch identische formale Varianten. Hinzu kommt ein hohes Ausmaß an nicht vorhersehbarer sprecherinterner Variation. Unterschiede zwischen verschiedenen Phänomenen, sowie zwischen verschiedenen Dialektregionen, sind daher immer gradueller Natur, d.h. sie bestehen in unterschiedlichen Variationsgraden. In konkreten Zahlen decken die berechneten Nicht-Standard-Frequenzen einen unerwartet weiten Bereich ab, der von 0,1% bis über 90% reicht (im Mittel 32,5%). Eine erkennbare Ballung findet sich allerdings einmal im niedrigen Frequenzbereich (< 20%; z.B. Subjektformen in Objektposition und umgekehrt) und andererseits im sehr hohen Frequenzbereich (z.B. beinahe-kategorische Objektformen in Sätzen ohne Verb).

Die Verteilung unterschiedlicher Pronomenformen lässt sich aus mehreren Perspektiven beleuchten. Hierzu gehört einmal die vielfach höhere funktionale Flexibilität von Objektformen, sowie deren Bevorzugung in allen Konstruktionen mit unklaren

Kongruenzregeln. In zahlreichen syntaktischen Positionen haben sich Objektformen dort, wo präskriptive Regeln weiterhin Subjektformen fordern, zum unmarkierten 'default' entwickelt (üblicherweise 70% bis 80 %). Es ist zu erwarten, dass hochfrequente nicht-standardmäßige Verwendungen aufgrund der allgemeinen Akzeptanz mit der Zeit Eingang in korpusbasierte und andere Grammatiken finden werden. Gleichzeitig kann man beobachten, wie der Versuch präskriptiver Grammatiken, sich gegen die Verbreitung hochfrequenter Objektformen zu wehren, zu Hyperkorrekturen führt (z.B. Verwendung von *and I* in koordinierten Nominalphrasen mit Objektfunktion).

Eines der wichtigsten Ergebnisse der vorliegenden Arbeit betrifft die Priorisierung verschiedener morphosyntaktischer Kategorien, welche deren Bedeutung für die korrekte Verarbeitung pronominaler Ausdrücke widerspiegelt (im Englischen sowie im typologischen Vergleich): Person > Numerus > Genus > Kasus. Mit seiner festen Satzstellung bietet das Englische vor allem Spielraum für Variation im Bereich des Kasus, da die syntaktische Funktion eines Wortes aus seiner Stellung abgeleitet werden kann.

Am Ende der Arbeit wird noch einmal die Bedeutung einer gesamtheitlichen Vorgehensweise hervorgehoben. Es ergibt sich ein komplexes Bild. Neben grundlegenden syntaktischen Aspekten (struktureller Kasus und 'default' Kasus), ist eine Vielzahl einflussreicher Faktoren zu beobachten: Diatopische Variation, Sprechervariablen, Emphatischer Sprachgebrauch, Vermeidung von Ambiguität, Hyperkorrektur und Vermeidung bestimmter Formen aufgrund sprachlicher Unsicherheit, Frequenzeffekte, Einfluss bestimmter Kollokationen, und die Verminderung des kognitiven Aufwands während der Sprachverarbeitung. Eine besondere Rolle kommt der pragmatischen Verwendung bestimmter Pronomen zu, welche im letzten Diskussionskapitel unter 'Pronoun pragmatics' zusammengefasst ist. Hier spielt vor allem das grundlegende Bedürfnis nach Expressivität eine Rolle, sowie das Ausdrücken von Höflichkeit, die Sprecherperspektive und die Beziehung zwischen Sprecher und bezeichneter Entität. Die Ergebnisse zeigen, dass außersprachliche Faktoren ein fester Bestandteil sprachlicher Variation sind, und dass sie besonders bei der Verwendung der englischen Personalpronomen nicht länger ignoriert werden können.

Die Arbeit schließt mit einer Zusammenfassung der wichtigsten Ergebnisse und Schlussfolgerungen aus der vorliegenden Untersuchung und einem Ausblick auf mögliche weiterführende Themen und Forschungsansätze.