MODALITY ON TREK: DIACHRONIC CHANGES IN WRITTEN SOUTH AFRICAN ENGLISH ACROSS TEXT AND CONTEXT

G.P. Wasserman

20390750

Thesis submitted for the degree Doctor of Philosophy
in English at the Vaal Triangle Campus of the North-West University

Promoter: Prof. A.J. van Rooy

SUMMARY AND KEY TERMS

SUMMARY

This study describes the diachronic development of modality in South African English (henceforth SAfE) from the early 19th century up to its contemporary state (1820s to 1990s) in the registers of letters, news, fiction/narrative and non-fiction, on the basis of the theoretical framework of sociohistorical linguistics and the empirical approach of corpus linguistics. Both quantitative and qualitative analyses are conducted for modal and quasi-modal verbs, by means of the newly compiled historical corpus of SAfE and ICE-SA (with the addition of Afrikaans corpora for comparison). The study explores general frequency changes, register-internal changes and macro- and microsemantic changes, with the focus of the main semantic analysis more strongly on the obligation and necessity cluster[1]. A set of parameters is compiled for analysing the strength of obligation in the modals *must* and *should*, and the quasi-modal *HAVE to*, and is applied in the microsemantic analyses. The findings are compared with the trends for modality in other native Englishes, such as American, British and Australian English (cf. e.g. Mair & Leech, 2006; Collins, 2009a; Leech, 2011), in an attempt to present a complete and comprehensive description of SAfE modality, as opposed to the traditional approach of focusing on peculiar features.

It is reported that the trends of modality in SAfE correspond to those of other native varieties in some cases, but do not correspond in others. The modals of SAfE for example have declined more and the quasi-modals have increased less over the 20th century than in other native varieties of English. One particular case in which

[1] An earlier and more compact version of some of the aspects of the methodology and findings of this study (mostly regarding synchronic frequencies and the semantics of the obligation and necessity cluster) has been published in an article by Wasserman and Van Rooy (2014) (the author and the promoter). The current study however presents newer and much more extensive results, by making use of e.g. a larger, newer version of the historical corpus of SAfE and a refined list of parameters for analysing degrees of obligative strength (see Chapter 3), as well as much more comprehensive investigations (regarding all the modals and quasi-modals), interpretations and comparisons.

SAfE is reported to be unique among other varieties, is the quantitative and qualitative trends for *must*, which has some implications for the manifestation of the democratisation process. *Must* in SAfE has not declined significantly over the 20th century (as it has in other native varieties) and has become less face threatening, since uses with a median (weaker) degree of force are just as frequent as those with a higher degree of force by the 1990s (unlike in other native varieties, where *must* has become restricted to high-degree obligative contexts). Based on sociohistorical, as well as linguistic evidence (on both quantitative and qualitative levels), language contact with Afrikaans is posited as the main influence for the increased use of *must* in contexts that are not face threatening. Extrapolating from the semantic findings, some new insights are offered regarding the phase in which SAfE finds itself within Schneider's (2003) model of the evolution of New Englishes, and some support is offered for Bekker's (2012:143) argument that "SAfE is ...the youngest of the colonial varieties of English", especially in the Southern Hemisphere. Ultimately, this thesis offers a piece in the larger puzzle that is SAfE, both in terms of linguistic (textual) and sociohistorical (contextual) aspects.

KEY TERMS

South African English, SAfE, WSAfE, modality, modals, quasi-modals, English grammar, English morphosyntax, English verb phrase, corpus linguistics, historical linguistics, diachronic linguistics, sociolinguistics, register studies, grammaticalisation, new dialect formation, World Englishes, Southern Hemisphere Englishes, Post Colonial Englishes, language contact, English semantics, Afrikaans modals.

ACKNOWLEDGEMENTS

I owe the realisation of this thesis to the support of many persons, first of which is my promoter, Professor Bertus van Rooy, to whom I owe my introduction to the field of linguistics and indeed my interest in its pursuit. On top of the academic nurture I received from you since my first year at university, I thank you, firstly, for your invaluable mentorship and intellectual guidance in discovering, learning, and mastering skills and knowledge, secondly, for your help and support in all practical matters concerning this research project[2], including all logistical and infrastructural aspects, as well as countless hours spent reading the different versions of the text and providing extensive, constructive and patient editorial guidance, and thirdly, for providing me with opportunities and encouragement that have enabled me to develop and keep developing in the academic environment. The existence of this thesis is a testimony to your instruction.

My deepest thanks to my parents, who taught me language and a love for it – you have been my primary teachers in life. I thank my dad, Kobie, for instilling in me a love for knowledge, argument and discovery, as well as a deep-founded respect for personal and collective history. Thank you for the hours of nightly philosophising and for the motivation to study language at university. Your love and interest have given me the courage to pursue excellence. I thank my mom, Christa, for reading me stories since I had ears to hear, for still teaching me new words or old expressions every day, for giving me an awareness of language and the enjoyment it can bring (even if this was done by teaching me how to do crossword puzzles or how to swear in Setswana), for playing countless games with me, and for instilling in me a hunger for knowledge about the world. Your love and support have made me who I am.

I dedicate this thesis to my husband, Wimpie: thank you for accompanying me throughout this journey. You have stretched and grown with me on this long, sometimes uphill, path, with more love, friendship and support than I could wish for.

[2] I would like to acknowledge the financial support I received from the North-West University (NWU) and the National Research Foundation (NRF) in the form of grants, which have helped to fund the research presented in this thesis.

May this be a monument to your patience in the wake of my fluctuating episodes of despair and procrastination, and optimistic (sometimes frantic) typing through many nights, as well as my atonement for the time you spent at the giving end of hundreds of meals and cups of tea, and thousands of words of encouragement.

To the Inventor of language and my Creator, I offer You this work in full appreciation that it is only the fruit of what You have begun and of what You have finished. *Aan God al die eer!*

TABLE OF CONTENTS

Summary and key terms ..... **iii**
Acknowledgements ..... **v**
List of figures ..... **xiii**
List of tables ..... **xv**

1 Introduction **1**
1.1 RATIONALE AND SCOPE 1
1.1.1 SAfE grammar 2
1.1.2 The modal system in context 6
1.1.3 SAfE and theoretical issues 8
1.2 SOCIOLINGUISTIC CONSIDERATIONS 12
1.2.1 Labels: SAfE vs. WSAfE 12
1.2.2 Social variables 13
1.3 THE INTERPLAY OF EMPIRICAL AND THEORETICAL CONSIDERATIONS 14
1.4 NOTATION STYLE 15
1.5 RESEARCH QUESTIONS, OBJECTIVE AND GOALS 16
1.6 STRUCTURE OF THESIS 18
1.7 CONCLUSION 18

2 Literature review **21**
2.1 INTRODUCTION 21
2.2 THEORETICAL FRAMEWORKS AND THE EMPIRICAL APPROACH 22
2.2.1 Main theoretical framework: sociohistorical linguistics 23
2.2.1.1 Historical Linguisitcs 25
2.2.1.1.1 Grammaticalisation theory 37
2.2.1.2 Sociolinguistics 47

2.2.1.2.1 Frameworks for English varieties 51
2.2.1.2.2 Politeness theory 60
2.2.2 Main empirical approach: corpus linguistics 65
2.2.2.1 Register studies 67
2.3 SOCIAL HISTORY AND LANGUAGE CHANGE 75
2.3.1 General reasons for language change 75
2.3.1.1 Change in English: migration, contact and identity 77
2.3.2 The World, Postcolonial and Southern Hemisphere Englishes 81
2.3.2.1 A *tabula rasa* context 82
2.3.2.2 The Founder Effect 83
2.3.2.3 Colonial lag 85
2.3.2.4 Contact in sociolinguistic paradigms 86
2.3.3 A sociohistorical overview of SAfE 93
2.3.3.1 Migration 93
2.3.3.2 The extended formation of SAfE and its contemporary state 99
2.3.3.3 Toward a framework for SAfE 101
2.3.3.4 Contact with Afrikaans 112
2.3.3.4.1 Historical overview 112
2.3.3.4.2 Identity and its linguistic implications 117
2.4 MODALITY AND THE MODAL SYSTEM 123
2.4.1 The English modal system: description and development 124
2.4.1.1 Description: approaches, terminology, grammatical characteristics and semantics 124
2.4.1.2 The development of the modal class 147
2.4.1.2.1 The semantic development of *must* 152
2.4.1.3 The development of the quasi-modal class 157
2.4.2 Modality and English varieties 162
2.4.2.1 General trends in native Englishes 162
2.4.2.2 Trends related to polysemy and monosemy 169
2.4.2.3 Synchronic variation in native Englishes 175
2.4.3 Modality in SAfE 179
2.4.3.1 English and Afrikaans modal cognates 181
2.5 CONCLUSION 190

3 Methodology 193
3.1 INTRODUCTION 193
3.2 RESEARCH DESIGN 194
3.2.1 Study population 194
3.2.2 Corpora 196
3.2.2.1 The historical corpus of SAfE 197
3.2.2.2 A corpus of contemporary SAfE 208
3.2.2.3 Corpora of contemporary Afrikaans 209
3.3 GENERAL ANALYTIC AND QUANTITATIVE METHODS 210
3.3.1 Tools 210
3.3.2 Methods 210
3.4 INTERPRETATIVE, QUALITATIVE METHODS 211
3.4.1 Macro- and microsemantic analyses 211
3.4.1.1 Microsemantic parameters 214
3.5 CONCLUSION 222

4 Results 223
4.1 INTRODUCTION 223
4.2 GENERAL RESULTS 224
4.2.1 Diachronic results 224
4.2.1.1 Introduction 224
4.2.1.2 The permission, possibility and ability cluster 243
4.2.1.2.1 *Can* and *could* 244
4.2.1.2.2 *May* and *might* 250
4.2.1.2.3 *BE able to* 257
4.2.1.3 The prediction and volition cluster 259
4.2.1.3.1 *Will* and *would* 260
4.2.1.3.2 *Shall* 270
4.2.1.3.3 *BE going to* and *WANT to* 278
4.2.1.4 The obligation and necessity cluster 285
4.2.1.4.1 *Must, should* and *HAVE to* 287
4.2.1.4.2 *NEED to*, *need* and *ought* 298

4.2.1.4.3 *(HAVE) got to, BE supposed to*, *(had) better* and *BE to* 306
4.2.1.5 Summary: diachronic results 316
4.2.2 Synchronic results 320
4.2.2.1 Introduction 320
4.2.2.2 Modals and quasi-modals in speech and writing 324
4.2.2.3 Results for contemporary Afrikaans 335
4.2.2.4 Summary: synchronic findings 339
4.2.3 Conclusion: general results 341
4.3 SEMANTIC RESULTS OF THE MAJOR CONTENDERS IN THE OBLIGATION AND NECESSITY CLUSTER 342
4.3.1 Introduction 343
4.3.2 *Must* 344
4.3.2.1 Macrosemantics of *must* 344
4.3.2.1.1 Dynamic *must* 348
4.3.2.1.2 Epistemic *must* 350
4.3.2.2 Microsemantics of deontic *must* 351
4.3.2.2.1 *Must*: high and median degrees of obligation 352
4.3.2.2.2 Formulaic *must* 373
4.3.2.2.3 Deontic *must* in spoken SAfE and deontic *moet/moes* in spoken Afrikaans 376
4.3.3 *Should* 389
4.3.3.1 Macrosemantics of *should* 390
4.3.3.1.1 Epistemic *should* 394
4.3.3.1.2 Quasi-subjunctive *should* 395
4.3.3.1.3 Preterite *should* 400
4.3.3.2 Microsemantics of deontic *should* 403
4.3.3.2.1 *Should*: high and median degrees of obligation 404
4.3.3.2.2 Formulaic *should* 427
4.3.4 *HAVE to* 430
4.3.4.1 Macrosemantics of *HAVE to* 430
4.3.4.1.1 Dynamic *HAVE to* 433
4.3.4.1.2 Epistemic *HAVE to* 436
4.3.4.2 Microsemantics of deontic *HAVE to* 436

4.3.2.2.1 *HAVE to*: high and median degrees of obligation.....437
4.3.2.2.2 Formulaic *HAVE to*.....458
4.3.5 Comparative summary: the semantic ecology of obligation and necessity.....460
4.3.6 Conclusion: semantic results.....470
4.4 CONCLUSION.....472

5 Conclusion.....475
5.1 INTRODUCTION.....475
5.2 SUMMARY OF FINDINGS ACCORDING TO RESEARCH QUESTIONS, OBJECTIVE AND GOALS.....475
5.3 THEORETICAL IMPLICATIONS FOR THE DEVELOPMENT OF SAFE.....492
5.3.1 Grammaticalisation in SAfE.....492
5.3.2 The development of SAfE: propagation of features, stages of development and language contact.....495
5.4 RECOMMENDATIONS FOR FURTHER RESEARCH.....502
5.5 CONCLUSION.....503

Appendix 1
Word counts for the Historical Corpus of SAfE.....505

Appendix 2
Diachronic and synchronic raw and normalised frequencies of modals and quasi-modals in SAfE.....506

Appendix 3
Log likelihood scores of modal and quasi-modal change in SAfE between periods to assess statistical significance.....513

Appendix 4

Synchronic comparative frequencies of *must, should* and *HAVE to* in SAfE and other native varieties 515

Appendix 5

Raw and normalised diachronic frequencies for the macro- and microsemantic analyses of *must*, *should* and *HAVE to* in SAfE 516

Appendix 6

Raw and normalised synchronic frequencies for *must* in SAfE and *moet/moes* in Afrikaans according to subject 526

Bibliography 527

LIST OF FIGURES

1 Map of Southern Africa ca. 1898 97

2 The Three-Stage Koinéization Model of the Formation of SAfE 109

3 A contemporary map of South Africa 200

4 Some patterns of linguistic frequency change 225

5 Overall normalised frequencies of modals and quasi-modals across time 226

6 Frequency trends of modals and quasi-modals across time 227

7.1 Frequency trends of combined modals and quasi-modals in the four separate registers across time 232

7.2 Frequency trends of modals in the four separate registers across time 234

7.3 Frequency trends of quasi-modals in the four separate registers across time 235

8 Frequency trends of the permission/possibility/ability cluster across time ... 243

9 Frequency trends of the prediction/volition cluster across time 260

10 Frequency trends of the obligation/necessity cluster across time 286

11 Frequency changes of *must, should* and *HAVE to* in SAfE over time 295

12 Frequency changes of *must, should* and *HAVE to* in BrE and AmE over time 296

13 Modal and quasi-modal frequencies in contemporary SAfE for written and spoken registers 321

14 Speech/writing ratios for quasi-modals across native varieties of English ... 322

15 Combined spoken and written frequencies of *must, should* and *HAVE to* in ICE-corpora across native English varieties 330

16 The spoken and written frequencies of *must, should* and *HAVE to* in SAfE, BrE and AmE 332

17 The spoken and written frequencies of semantic equivalent modals and quasi-modals in SAfE and Afrikaans in the obligation and necessity cluster 338

18 Macrosemantic frequency changes of *must* in SAfE over time 344

19 Macrosemantic frequency changes of SAfE *must* per register over the 20th century 346

20 Microsemantic frequency changes in SAfE deontic *must* over time 352

21 Macrosemantic frequency changes of *should* in SAfE over time 390

22 Macrosemantic frequency changes of SAfE *should* per register over the 20th century 393

23 Frequencies of the five types of quasi-subjuctive uses of *should* in SAfE over time 397

24 Microsemantic frequency changes in SAfE deontic *should* over time 403

25 Macrosemantic frequency changes of *HAVE to* in SAfE over time 430

26 Macrosemantic frequency changes of SAfE *HAVE to* per register over the 20th century 432

27 Microsemantic frequency changes in SAfE deontic *HAVE to* over time 437

28 Comparative frequency changes of epistemic modality in *must*, *should* and *HAVE to* in SAfE over time 461

29 Comparative frequency changes of deontic modality in *must*, *should* and *HAVE to* in SAfE over time 463

30 Comparative frequency changes of dynamic modality in *must* and *HAVE to* in SAfE over time 465

31 Comparative microsemantic frequency changes of deontic *must*, *should* and *HAVE to* in SAfE over time 466

32 Proportional tendencies for high- and median-degree *must*, *should* and *HAVE to* to occur with objective and subjective sources in SAfE in the 1990s 468

LIST OF TABLES

1 Afrikaans and English modal cognates/semantic equivalents 189

2 Total word counts for registers in the historical corpus of SAfE per period . 198

3 Modal categories based on semantic clusters 212

4 Parameters for high and median degrees of obligation..... 215

5.1 Overall normalised frequencies of modals over time 241

5.2 Overall normalised frequencies of quasi-modals over time 241

6.1 Normalised synchronic frequencies of modals in written and spoken SAfE. 324

6.2 Normalised synchronic frequencies of quasi-modals written and spoken in SAfE..... 324

7 Afrikaans modal and quasi-modal frequencies per 100k words with English semantic equivalents 337

8 Percentages of high-degree deontic uses with subjective, objective or indeterminate sources of obligation in *must* 354

9 Percentages of median-degree deontic uses with subjective, objective or indeterminate sources of obligation in *must* 356

10 Percentages of formulaic deontic uses with subjective, objective or indeterminate sources in *must* 374

11 Macrosemantic frequencies per 100 words for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register 377

12 Microsemantic deontic frequencies per 100 words for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register 378

13 Percentages of high-degree deontic uses with subjective, objective or indeterminate sources of obligation in *should* 405

14 Percentages of median-degree deontic uses with subjective, objective or indeterminate sources of obligation in *should* 406

15 Percentages of formulaic deontic uses with subjective, objective or indeterminate sources of obligation in *should*, with raw numbers in brackets 427

16 Percentages of high-degree deontic uses with subjective, objective or indeterminate sources in *HAVE to* 438

17 Percentages of median-degree deontic uses with subjective, objective or indeterminate sources in *HAVE to* 441

18 Percentages of formulaic deontic uses with subjective, objective or indeterminate sources in *HAVE to* 458

19 Word counts for the Historical Corpus of SAfE 505

20.1 Raw diachronic frequencies of modals in SAfE 506

20.2 Raw diachronic frequencies of quasi-modals in SAfE 507

21.1 Normalised diachronic frequencies of modals in SAfE per 100 000 words .. 508

21.2 Normalised diachronic frequencies of quasi-modals in SAfE per 100 000 words 509

22 Combined diachronic raw and normalised (per 100 000 words) frequencies of auxiliaries in SAfE 510

23.1 Raw synchronic frequencies of modals in SAfE 511

23.2 Raw synchronic frequencies of quasi-modals in SAfE 511

24.1 Normalised synchronic frequencies of modals in SAfE per 100 000 words . 511

24.2 Normalised synchronic frequencies of quasi-modals in SAfE per 100 000 words 511

25 Combined synchronic raw and normalised (per 100 000 words) frequencies of auxiliaries in SAfE 512

26.1 Log likelihood scores of modals over time 513

26.2 Log likelihood scores of quasi-modals over time 514

27 Combined spoken and written frequencies of *must, should* and *HAVE to* in ICE-corpora across native English varieties per million words 515

28 The spoken and written frequencies of *must, should* and *HAVE to* per million words in ICE-SA compared to F-LOB and Frown, and the BNC and LCSAE 515

29.1 Raw macrosemantic frequencies of SAfE *must, should* and *HAVE to* over time 516

29.2 Normalised macrosemantic frequencies of SAfE *must, should* and *HAVE to* per 100 words (%) over time 517

30.1 Raw macrosemantic frequencies of SAfE *must* per register over the 20th century 518

30.2 Normalised macrosemantic frequencies of SAfE *must* per register per 100 words (%) over the 20th century 518

31.1 Raw macrosemantic frequencies of SAfE *should* per register over the 20th century 519

31.2 Normalised macrosemantic frequencies of SAfE *should* per register per 100 words (%) over the 20th century 519

32.1 Raw macrosemantic frequencies of SAfE *HAVE to* per register over the 20th century 520

32.2 Normalised macrosemantic frequencies of SAfE *HAVE to* per register per 100 words (%) over the 20th century 520

33.1 Raw microsemantic frequencies of deontic SAfE *must, should* and *HAVE to* over time 521

33.2 Normalised microsemantic frequencies of deontic SAfE *must, should* and *HAVE to* per 100 words (%) over time 522

34.1 Raw frequencies of subjective, objective and indeterminate sources with high-degree, median-degree and formulaic *must, should* and *HAVE to* in SAfE over time 523

34.2 Normalised frequencies of subjective, objective and indeterminate sources with high-degree, median-degree and formulaic *must, should* and *HAVE to* per 100 words (%) in SAfE over time 524

35.1 Raw frequencies of the constructions of quasi-subjunctive *should* in SAfE over time 525

35.2 Normalised frequencies of the constructions of quasi-subjunctive *should* per 100 words (%) in SAfE over time 525

36 Raw numbers and percentages of person of subject with SAfE *must* and Afrikaans *moet/moes* in the spoken register 526

37 Raw and normalised macrosemantic frequencies for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register 526

38 Raw and normalised microsemantic deontic frequencies for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register 526

CHAPTER I

INTRODUCTION

"Time is change; we measure its passing by how much things alter." (Nadine Gordimer, 1966)

1.1 RATIONALE AND SCOPE

In the almost two centuries that the English language has shared its fate with a diverse mix of South African societies and their languages, it has faced some of the most complex social contexts to become the widespread lingua franca amongst ten other official languages. This thesis is interested in how we might 'measure' this relatively short span of time in the history of a language in terms of both how much its social context has changed and how much it has altered in its structure and meaning as a result of that context. The title of this study is inspired by the Great Trek, a South African event that was essentially a reaction to the oppression of the Afrikaner nation (Dutch descendants) by the British government at the Cape Colony in the early 19th century. Indeed the many treks of the English language into and within the country, with their winding paths of exploration and exploitation, and oppression and liberation, as well as its range of contact situations with neighbours and co-trekkers along the way, have been 'great' in their own right: they have shaped a unique variety among the other Englishes of the world.

The grammar of South African English (henceforth SAfE), here regarded as the variety spoken by descendants of the 19th century British settlers also called White South African English, has hitherto received little attention in research, as also noted by e.g. Lass (2002). Although some studies have been done on the grammar of Black South African English (e.g. De Klerk, 2006; Van Rooy, 2006; 2008; 2011 and Botha,

2013) and Indian South African English (e.g. Mesthrie, 1992), earlier researchers were generally unconvinced that SAfE grammar merits attention. Lanham and MacDonald (1979) indeed felt that there is little to be said about SAfE grammar, as it is primarily a distinctive accent. Titlestad (1996:168) likewise claimed, "if one excludes accent, then one is hard put to identify SAfE apart from a body of colloquial terms." Consequently, the phonological component of SAfE has attracted more interest and has indeed been thoroughly described by Bekker (2009; 2012). This situation is however not a unique one, as much the same tendency exists in the study of other Southern Hemisphere varieties such as New Zealand English, as noted by Hundt (1998). The lexical features of the variety have also been comprehensively documented by Branford and Branford (1991) in their *Dictionary of South African English*, as well as by Silva (1996) in her *Dictionary of South African English on Historical Principles*.

On top of the general neglect of grammatical study, a particular tradition in SAfE studies has restricted the scope of its research potential even further. By traditionally documenting only those peculiar features of SAfE, previous researchers inevitably focused on what makes this variety different from the English of its colonial parent variety and other varieties across the globe, although Cheshire (1991:1) deems this true for the description of many varieties of English. In contrast to such emphasis on divergence (or lack thereof), both Schneider (2007) and Trudgill's (2004) recent theoretical models of the evolution of New Englishes point toward the many corresponding facets shared by Post Colonial Englishes (PCEs), and specifically by Southern Hemisphere Englishes (SHEs) – mainly SAfE, Australian English and New Zealand English. Hence, studies that will facilitate this perspective focused on drawing similarities are wanting.

1.1.1 SAfE grammar

The first prominent listing of the peculiarities of English in South Africa regarding grammar, as well as other linguistic domains, is by Beeton and Dorner (1975) in their monument of mid to late 20th-century prescriptivism: *A Dictionary of English Usage in Southern Africa*. It aims to provide, firstly, "a glossary of local vocabulary and

idiom, together with judgments on the desirability and efficacy of each of the words listed", secondly, "a record of mistakes and problems common not only to English-speaking South Africans, but to all people using English", thirdly, "a record of mistakes and problems characteristically South African, and guidance in overcoming them", in the fourth place, "a record of problems encountered by those South Africans for whom English is not the home language", and, lastly, "a list of departures from Standard English pronunciation", together with, as expected, "guidance in the matter of pronunciation" (1975:iv).

In their review of perceived 'acceptable' and 'unacceptable' idiomatic uses, Beeton and Dorner (1975:xiii) list, for instance, the peculiar uses of the *busy* + *-ing* construction, as well as *now* and *must* among South Africans, and suggest that they should be grouped among the 'unacceptable' idioms:

> "...one gains an inkling of how fine and subjective are the lines of distinction between acceptable and unacceptable idiom. But that such a line of distinction should be attempted seems to us beyond question: otherwise we should, for example, merely surrender to the slovenly (but highly infectious) South African habit of using 'just now' to indicate almost any time in the foreseeable future. ... We lay great, and incorrect, emphasis on 'being busy' ('he is busy talking to so and so'), and see an invitation far too frequently in terms of compulsion: 'Must I call on him?' (for 'Would he like me to call on him')." (Beeton & Dorner, 1975:xiii.)

Furthermore, Beeton and Dorner's entries for *must* suggest that it is used to express other meanings than that of a high or strong degree of obligation, since the their 'guidance' warns against such other uses. The same applies to the quasi-modal *HAVE to*. The following entries appear (Beeton & Dorner, 1975:81;122 [original emphasis; my numbering]):

(a) **must**

must = have to, be forced to, & should not be used in a simple inquiry, eg '*Must* I sign here? = 'Am I forced to sign here?' & not, '*Shall* I sign here?' 'Do you *want/wish* me to sign here?' constructions wh do not suggest compulsion.

(b) **must/may**

'You *must* go home now' is a command; 'You *may* go home now' is permission.

(c) **have to/must**

should be avoided when replying to an enquiry for an interview, eg 'You *will have to/must* see the manager' x; 'You *will have to/must* wait' x; '*Will/would* you see the manager?' + '*Will/would* you mind waiting?' +

In the last quoted entry (c), Beeton and Dorner uses the 'x' at the end of the first and second examples to indicate errors in each case, whereas the '+'-sign indicates correct usage in the last two examples. The prescriptive tradition for English in South Africa, with such emphasis on error, is beautifully exemplified through Beeton and Dorner, and was only really transcended in the 1990s[1].

In the first real attempt at a grammatical description of the variety, Mesthrie and West (1995) outlined some characteristics of the input form of SAfE or Proto South African English. The input form or baseline for English in the country was the English of the British settlers of the early 19th century, which was represented by a corpus of letters mainly written to the Governor of the Cape colony (1995:111). Here mostly peculiar stylistic and striking grammatical features are noted, some of which are related to the verb phrase, for example, the 'adjective with infinitive' construction, omission of the *-s* marker on verbs with third person subject and other variations in verb morphology, as well as the unstressed *do* construction (1995:115;121;124;128).

Mesthrie (1996) later described some features of missionary English at the hand of unpublished letters written by Cape missionaries during the period 1800 to 1830 (1996:142), but mostly focused on the writing styles of, and exceptional instances of English usage by, individual authors. Even though this is not so relevant to this thesis in itself, seeing that modality is not mentioned, it is worth noting that the input-form, and a more specific input-source (missionary English), have been broadly analysed. Concerning Lanham and MacDonald's (1979) assumption that there is not much to be said about the grammar of SAfE (in their discussion of the phonological features of the variety), Mesthrie and West (1995:106) observed that this may be true for the variety at present, but not for its historical profile.

Moving into the twentieth century, and keeping with the traditional emphasis on peculiarities, Bowerman (2004b) compiled a list of grammatical features of contemporary SAfE, from mostly spoken data, according to him. Drawing on his list,

[1] This kind of prescriptivism links with the complaint tradition that characterises Schneider's (2003) Phase Three ('nativization') in his model for the evolution of New Englishes (see § 2.2.1.2.1 and § 5.3.2).

he for instance claimed that SAfE does not show much grammatical deviation from other standard Englishes (mainly standard BrE) apart from a few features, although he conceded that the morphosyntax of SAfE had not been well researched previously (2004b:472). In addition, Bowerman does not account for any of his documented features on the basis of corpus evidence. An example of one of the peculiar features he lists is the '*busy* + progressive' construction, which Lass & Wright (1986), as well as Beeton & Dorner (1975), had also previously mentioned.

In the same year that Bowerman listed his apparently anecdotal features, Jeffery and Van Rooy (2004) investigated the extended uses and functions of the temporal adverb *now* to a subjunct and emphasiser, based on data from ICE-SA[2]. Jeffery and Van Rooy suggested that, due to contact with Afrikaans (the language spoken by the descendants of the Dutch colonists), a construction originating from that language has been nativised into SAfE (2004:278) (compare also the allusion to another borrowed use of *now* mentioned in the above quote by Beeton & Dorner [1975:xiii]). On the same note, Schneider (2007:184) more recently listed a number of complementation patterns and distinctive preposition uses – all of which appear to be possible loan translations from Afrikaans.

However, in the particular case of the above-mentioned '*busy* + progressive' construction, such direct influence from Afrikaans has been challenged by both Lass and Wright (1986) and Mesthrie (2002). These authors argue that the use of this construction in SAfE cannot really be attributed to Afrikaans influence, as "the structure was an endogenous development reliant on already existing options within internal English structure, rather than a feature of language contact" (Mesthrie 2002:345). However, Lass and Wright (1986:217-218) argue that Afrikaans might be responsible for the lifting of a relevant semantic restriction in English, for which Mesthrie (2002:346-347) did not draw the same conclusion regarding qualitative semantics. Mesthrie however notes that the language might have had an effect on quantitative aspects (2002:358) and indeed regards Afrikaans as an influential factor in the general development of SAfE (cf. Lanham, 1982).

These studies provide welcome preambles to the variety's full grammatical description, but are either lacking in scope (only the input variety in Mesthrie and West [1995])) or empirical evidence (as in Bowerman [2004b]), or are divided

[2] The South African component of the International Corpus of English (see Jeffery, 2003).

regarding the extent of the influence of social issues such as language contact (compare Lass & Wright [1986], Mesthrie [2002] and Jeffery & Van Rooy [2004]). In addition, these studies still leave large parts of the historical profile of SAfE unexplored. It is evident that much still needs to be done in order to describe the grammar of the variety to its full potential. This thesis aims to provide a description that will increase that potential.

1.1.2 The modal system in context

As the first step of a more comprehensive description of SAfE grammar, a practical starting point is the grammatical element of modality, moreover the modal auxiliary group, including the core modal verbs, also called the modals (e.g. *can, must* and *will*), and the quasi-modal verbs, also called the quasi-modals (e.g. *BE able to, HAVE* to and *BE going to*). The development and current state of this system has attracted much academic interest in other major native English varieties, for example in the comparative, synchronic study of American, Australian and British English (henceforth AmE, AusE and BrE) by Collins (2009a). The historical development of English modality since its genesis through the process of grammaticalisation (e.g. Facchinetti *et al.*, 2003) until its contemporary state (e.g. Biber *et al.*, 1999) has also been extensively documented, whereas e.g. Leech (2003; 2011), Mair & Leech (2006), Leech & Smith (2006; 2009), Leech *et al.* (2009) and Mair (2014) have done some insightful research on the recent changes in modal auxiliary use, especially involving the interplay between modals and quasi-modals. These studies, as well as the English modal system will receive further attention in Chapter 2 and will serve as a comparative basis for my findings.

Essentially the only scholar to have listed features specifically relating to modality in SAfE (apart from Beeton & Dorner's prescriptive entries related to e.g. *must*) is Bowerman (2004b:477) – they are:

(i) The illocutionary force[3] of *must*: this strong obligative modal has much less social impact in WSAE[4] than in other varieties of English and often substitutes for polite should/shall, e.g. *you must turn left at the robots...*

(ii) Won't as a directive "softener": this use of *won't* softens a request, e.g. *won't you do me a favour?*

As mentioned in the previous section, Bowerman's focus on linguistic anomaly and his reliance on anecdotal evidence shows that a corpus-based account is clearly needed to obtain a more reliable picture of English modal use in the South African variety.

Drawing on these above-mentioned studies, the recent, exploratory inquiry into the modal patterns of SAfE by Rossouw[5] and Van Rooy (2012) indeed prompted interest into the deeper linguistic and extra-linguistic facets of modality in this variety, not only relating to its diachronic evolution, but also to its synchronic state, which will be the main foci of this thesis. In Rossouw and Van Rooy's pilot study, the general diachronic frequency patterns of SAfE modality, although appearing to be similar to its ancestor, British English (BrE), were in fact revealed not follow its patterns of change to the same extent, despite being somewhat conservative, since register-internal and modal-specific developments disclosed dissimilarities on a qualitative level[6]. Rossouw and Van Rooy's study, being mainly diachronic and exploratory, left gaps in synchronic description, sociolinguistic exposition and more in-depth semantic and pragmatic analyses. The subsequent lacuna in grammatical description and social interpretation ultimately impedes the potential of SAfE to be compared with other English varieties; thus, a more extensive study of SAfE modality, more so than that of Rossouw and Van Rooy (2012) is needed.

---

[3] Illocutionary force refers to "the intentions of speakers while speaking" (Crystal, 2008:446). The concept of an illocutionary act stems from the theory of speech acts and refers to "an act which is performed by the speaker by virtue of the utterance having been made" (2008:236). Crystal (2008:236) adds that examples of illocutionary acts "include promising, commanding, requesting, baptizing, arresting, etc." The act of commanding is relevant to the case of *must* as it is referred to by Bowerman (2004b) in terms of its strong obligative meaning. Illocutionary acts are contrasted with locutionary acts, which are acts of 'saying', and perlocutionary acts, which involve the "the effect [the act of 'saying'] has on the hearer" (Crystal. 2008:236).

[4] Also 'WSAfE' (White South African English).

[5] The author's maiden name.

[6] The results of Rossouw & Van Rooy's study (2012) will receive more attention in § 4.2.1.1.

I intend to provide not only an extended analysis of Rossouw and Van Rooy's findings, by means of a self-assembled and enlarged historical written corpus[7], but also to broaden the scope by adopting a more extensive temporal perspective and undertaking more rigorous investigations into contextual and semantic matters. The said temporal perspective of this thesis will involve a synchronic description of modality in SAfE, but will mainly maintain a strong retrospective and hence diachronic point of view – thus, trying to explain the present via the past. This perspective is aimed at ultimately creating a platform for drawing parallels with not only English varieties within South Africa itself, but also other varieties of English across the world, and offer comparability for both synchronic and diachronic approaches.

1.1.3 SAfE and theoretical issues

Since SAfE as a transplanted variety of English in the southern hemisphere is a product of colonialism and now part of an international community of English users, it is connected with three main sociolinguistic frameworks or paradigms of study, viz. the World Englishes (e.g. Kachru, 1992), Postcolonial Englishes (Schneider, 2003) and Southern Hemisphere Englishes (Trudgill, 2004), as mentioned earlier, and which will all receive full attention in § 2.2.1.2.1. The complexity of the linguistic situation in South Africa has led to some conjecture about how well, if at all, English in South Africa fits into the modals proposed by these paradigms (cf. Bekker, 2012; Van Rooy, 2010): as part of a conglomerate of Englishes in the world, or in terms of various factors or mechanisms (mainly cultural identity and dialect input) that institute the development of a new variety of English in a general postcolonial setting, as among e.g. American and Indian English, or specifically among the other Englishes of the southern hemisphere, e.g. Australian and New-Zealand English. In Kachru's three-circle model of World Englishes (1992), South Africa is not mentioned amongst the three categories of listings according to country, but Schneider (2003:243) remarks

[7] Seeing that no spoken data is available or obtainable for the 19th and early 20th centuries, the diachronic, as well as synchronic findings will only be based on written data in order to aid comparative methods, which are described in Chapter 3.

that South Africa has often "been sidestepped in such listings, or been classified somewhat forcedly" (see § 2.3.2).

Even when just considering the native or first language (L1) variety studied in this thesis, the complexity of the extensive linguistic contact situation with Afrikaans (a related West-Germanic language to English) (cf. Trudgill, 2004:5) is a complicating factor for the description of this variety (see § 2.3.3.4), which is less present in other varieties of L1 English, especially in the Southern Hemisphere[8]. The political context of the sanctioned policies of racial segregation under the apartheid regime (with Afrikaans as implementing language) from the mid to late 20th century is certainly unique in the world, and a factor that undoubtedly makes it difficult to classify SAfE along clear lines among other varieties (see § 2.3.3.4.1). Ongoing work by e.g. Schneider (2007), Van Rooy (2010), Van Rooy & Terblanche (2010) and Bekker (2012; 2013) seeks to deal with these issues and find solutions to the puzzle that is English in South Africa. The potential of the South African context as a means of understanding the role of language contact in the formation of New Englishes remains to be unlocked, as the gap in international discussions of not knowing enough about SAfE largely remains present.

One of the reasons behind this gap in research is the lack of historical corpora – the kind of resource that has led to the procurement of much linguistic evidence to support the claims of the proponents of the various frameworks for English varieties and hence validate their models. With the newly assembled historical corpus of SAfE and its ongoing compilation, this thesis aims to add such a resource to the arena of English studies. With this tool, I aim to make some contribution to the ongoing discussion of SAfE in terms of linguistic models and moreover the role of contact in such models.

It is interesting to note that the nature of English in South Africa has attracted much curiosity for almost as long as it has been used here. An 1870s article published

[8] For the Englishes of the Southern Hemisphere, Trudgill (2004:5) mentions that language contact with other European languages is present, but this does not so much involve contact with Germanic languages as with Romance languages (he notes for example that Falkland Islands English has had contact with Spanish). In the Northern Hemisphere, however, many varieties have had language contact with European languages, e.g. Canadian English with Canadian French (to some extent) and Irish English with Irish Gaelic, but it is only really in American English that contact with other Germanic languages has occurred (such languages include e.g. Dutch, German and Yiddish) (Trudgill, 2004:2;6).

in the *Cape monthly magazine* with the title *Cape English*[9], which is also included in the historical corpus, already raises many theoretical and linguistic issues and asks questions that are still debated over today, as the following extract from this rather lengthy article illustrates. Issues that are already raised here include increasing bilingualism (in both the Afrikaans and English communities), possible language shift to English among the Afrikaners, language contact (with Afrikaans), language spread, language change, identity, native-speaker authority, prestige, the selection of new features (lexical/phrasal/idiomatic), the role of children in the formation of a new variety, language standard, dominance, preservation, differences in spoken and written registers, social diversity in South Africa and the role of education.

> "It is no doubt a matter of congratulation among many who have at heart the advancement of civilization and refinement in the land, that the knowledge and use of the English language has so much increased of late years. And there is the full expectation that the development of our educational system will aid in the spread of this as "the vulgar tongue" of future generations of Africanders (using this term in no invidious sense). One would rejoice at such a result, not only from pardonable partiality for the language of the country to which we owe allegiance – which so many, who have never seen it, whose descent springs not from it, delight to call "home" – but also from an appreciation of the power and actual beauty and richness of the English language itself.
>
> One cannot, however, hold much converse with those who have lived here for any length of time (particularly in the country districts), without finding many occasions for questioning whether the source of the language is "the pure well of English undefiled" that many fancy it to be. There are several words in common use which would be sought in vain in any English dictionary; and there are many English words and phrases that are not used in the sense generally attributed to them by grammarians and lexicographers.
>
> Not to speak of the almost invariable use of "will" and "would" for "shall" and "should", we often hear such expressions as "by the house", meaning "at home" – "throwing with a stone" – "under" (Dutch "onder") used instead of "among". This un-English phraseology forms the every-day speech of those who are gradually attaining to the foremost places among us. They are, it may be, not yet found in our written tongue; but one cannot open a newspaper without seeing, especially in the correspondence columns, many an idiom which smacks of something else than the Anglo-Saxon tongue. ...
>
> Whatever value we may set on the other institutions and traditions of the mother country, and however we may differ in our estimate of them, few Englishmen – certainly no educated Englishman – will repudiate the obligation of colonists to guard the common treasure of the language, and hand it down to posterity unimpaired. The performance of this duty will be affected very much by the actual circumstances of the several colonies, and we should ignore the facts of human nature if we doubted that circumstances "over which they have no control", such as the hard physical features of a country, new social conditions, &c., will introduce changes into

[9] From: Lewin Robinson, A.M., *ed.* Selected Articles from the *Cape Monthly Magazine* (New Series 1870-76). Cape Town: J.C. Jutta, 1857-81.

> a language. We shall see this at once, if we think only of the rich exuberance of metaphor – due in a measure to the many-sidedness of the national life and character – with which the English language is overlaid, and ask ourselves what chance there is of this quality surviving in countries where pursuits, associations, and even the very skies are changed. ...
>
> There is no doubt that altered conditions of life will exercise a modifying influence on the language as well as on the character of a people; and where, as in this Colony, in addition to new surroundings, the language meets with a competitor claiming equal rights, some kind of compromise is nearly sure to follow. The two languages will live for a long time, we imagine, side by side, but occupying different spheres; business, commerce, and fashion espousing the one, and agriculture and all connected with it monopolizing the other. Of anything like fusion we feel no apprehension – but that the dominant language will levy contributions on the other is certain; and that in doing so it will lose something of its integrity is, to say the least, probable. It is bound to assert its vitality by growing, and its growth will be marked and regulated by the local influences brought to bear upon it. Annexing, adapting, modifying, and here and there, we fear, conceding something it will advance and develop until it take on the new geographical type, as surely as the transplanted shrub. New features will present themselves; but as long as the organic structure and substantial framework of the language remain, there need be no degeneracy. What we have to guard against, and may prevent, is corruption, not change. To check the one and regulate the other we must look to a diffused education bearing fruit in social refinement and critical taste. ...
>
> For many of the above [words/phrases] – for all, in fact, which have a local impress and a special fitness – we would bespeak a kindly hearing, and we think such additions will be no disfigurement to our language. That they will be admitted in any formidable numbers, we have little fear. The area of selection is too limited, the ordinary Dutchman's philosophy being of that "concentrated essence" kind, that a very limited vocabulary satisfies it. And though English children take readily enough – almost too readily – to the Dutch, adults find too many stumbling-blocks in the gutturals, the strange diphthongs, and coarse features it too frequently presents in unexpected quarters. We have known few English people acquire such a mastery over the language as to be thoroughly at home in a Dutch reunion. ...
>
> With respect to the rival language, "natural selection" will be busy for some time to come; and as intercourse grows, and the life of the Colony becomes more complicated, and speech itself develops new wants, new elements will be introduced; and so long as our literary priesthood guard the inner shrine of the temple from pollution, we trust our mother tongue will survive unimpaired, enriched though it may be, and invigorated by useful and picturesque additions. ..."

In this rather insightful piece, not untainted by colonial pride, it is truly fascinating that reference to a unique kind of modal use in the English variety of the country is already made here. The reference to *will* and *would* replacing *shall* and *should* will be investigated in the results section. The question I wish to raise at this point is that, if indeed these above-mentioned modals already overlap in their usage by the 1870s, can there be any reason to suppose that this kind of overlap was not already present or at least developing for other varieties as well? Chapters to come will explore this (see § 4.2.1.3.2 and § 4.3.3.1.3), along with the other theoretical issues raised not only in this historical article, but moreover by the linguistic scholars introduced above.

1.2 SOCIOLINGUISTIC CONSIDERATIONS

This section will discuss some important considerations related to sociolinguistic issues, such as labels and variables. Firstly, a clarification regarding the chosen label for the variety studied in this thesis will be given, followed by, secondly, a short exposition on the social variables that will be favoured in this study.

1.2.1 Labels: SAfE vs. WSAfE

The labels 'South African English' (SAfE) and 'White South African English' (WSAfE) have both come to be applied to the variety used by native speakers of English in South Africa (usually descendants of the 19th century British Settlers), but for the purpose of this thesis, the more concise term 'South African English' will be favoured to denote this speaker community. This subsection will discuss issues of terminology regarding this variety, where they are relevant to this thesis.

SAfE in its native-speaker context ideally encompasses Jeffery's (2003:343) prototypical "White English-speaking South African" (WESSA) in its entirety, but Jeffery argues that it would be misleading if this category does not include the fluently bilingual or multilingual (not necessarily English mother tongue) "educated in English to matriculation[10] level or beyond". However, I concede that the 'white' user of contemporary SAfE becomes increasingly difficult to define or to trace in terms of ancestry, and I am therefore also careful to claim absolutely that I only study the purely prototypical WESSA in this thesis. For earlier data this is much easier, as more well-known historical personalities are sometimes the authors of the texts I use, and many times biographical information about authors assists in ensuring the more accurate recognition of a native speaker.

Furthermore, Bekker (2009; 2012) for example distinguishes between three categories of SAfE in terms of traditionally 'white' speakers, which are based on a trichotomy applicable to all Southern Hemisphere Englishes (cf. Mitchell & Delbridge, 1965; Lass, 2002). These are: Cultivated SAfE (the prestigious variety

[10] The final year at secondary school level in South Africa, or Grade 12.

largely assuming British norms), General SAfE (the standard local variety) and Broad SAfE (a range of stereotyped varieties, typically associated with lower levels of education)[11]. The first two categories, Cultivated (CulSAfE) and General SAfE (GenSAfE), both fit into my understanding of SAfE or WSAfE in this study. The main difference between GenSAfE and the variety spoken by the WESSA is that GenSAfE is not confined to 'white' users of English – hence the term 'SAfE' is favoured above 'WSAfE' in this study. GenSAfE is therefore not necessarily the ideal label for the kind of SAfE I study here (since I endeavour to focus on native speakers), but it is not wholly independent of this label either[12]. The issue of labels for language varieties based on race has indeed always been a sensitive one in South Africa, and therefore brief notes on this matter is in order.

Even though this thesis will make reference to sociolinguistic issues in South Africa, it is not intent on perpetuating racial categories or equating language and race. However, it is a historical fact of SAfE that, at least until 1994, very clear ethnolects could be distinguished within the variety, as speech communities were forced to exist in relative isolation before the abolition of the apartheid system (cf. Coetzee-Van Rooy & Van Rooy, 2005:1). Since my data set spans the period until the 1990s, I will not engage in debates on the appropriate labels for varieties. The general native speaker variety of SAfE, its history and its present state, will receive further attention in Chapter 2.

1.2.2 Social variables

The sociolinguistic approach of the analyses relies on matters surrounding language variation, language contact, migration, identity and politeness. The subject of language variation includes issues of dialect and region. In the historical roots of SAfE it is apparent that some internal regional dialects enjoyed greater prestige than others, but this study will essentially not take up long discussions about the differentiation in sub-dialects according to their prestige, although this historical fact is important to consider for broad contextual reasons.

---

[11] Lanham (1967; 1978) and Lanham and Macdonald (1979), for example, prefer to refer to these categories as Conservative, Respectable and Extreme SAfE (cf. Lass, 2002:109-112).

[12] General SAfE (GenSAfE) is the basis for ICE-SA, which I include in my data for analysis.

This issue was however taken into consideration while compiling the corpora used in the analyses, as texts from various regions were covered in the data to ensure an even proportion of regional and hence internal dialect representation. Yet, it is beyond the scope and complexity of this study to engage in such micro-level textual investigations, contrasting between texts written in certain areas of the country. In the same way, other social variables such as age, gender and socio-economic background will not be scrutinised in the contextual analyses, but the corpora are compiled as being representative of a balanced mixture of these variables to ensure comparable data sets. Chapter 3 will provide more information on how the corpus used in this thesis was compiled.

The main social variables that will form the basis of discussion along with the linguistic variable (the modal system), are therefore geographical origin and ethnicity, especially for the data representing the 1870s to the 1990s, to ensure that the results will be representative of the specific variety of SAfE, as stipulated above.

1.3 THE INTERPLAY OF EMPIRICAL AND THEORETICAL CONSIDERATIONS

With a primarily corpus-based approach to my research I aim to ensure the empirical credibility of the analyses and findings, where real texts written (ideally) by descendants of the British Settlers will form the basis of investigation. Furthermore, a sociolinguistic perspective based on the historical environment and development of SAfE and its speech community will be integrated into every investigation and discussion to shed light on the context wherein these texts were written, and around which the English community in South Africa functioned.

The texts used in the corpus are fundamentally tied to the South African context, its political climate at a certain stage, the daily lives of individuals and the important events that helped shape the community every day across nearly two centuries. In the same way, an investigation which focuses purely on social context surrounding the variety will not be useful on its own and is tied to the corpus texts, because the influence of register developments are a crucial facet to consider when

dealing with written data, or any linguistic data for that matter (cf. Biber *et al.*, 1999). Aside from the diachronic investigation, a synchronic description of SAfE will also help to define the present situation of the variety, how it compares with other Southern Hemisphere varieties and, potentially, what the future might hold.

Although certain theoretical frameworks will influence the perspective adopted in my analyses (which will be explained in Chapter 2), this study will remain focussed on frequency patterns, and the explanation of such patterns from a sociolinguistic and semantic viewpoint. It is therefore not my intention to make a theoretical contribution to the grammatical and semantic domain of modality as such, but I do intend to increase the potential of SAfE to be described in terms of sociolinguistic theories, frameworks and models, such as Schneider's dynamic model of the evolution of New Englishes (2003) (see § 2.2.1.2.1 and § 2.3.3.3), as also mentioned in § 1.1.3 above.

1.4 NOTATION STYLE

For the purpose of notating modals and quasi-modals in this thesis, I will use small capitals together with italics to denote lexical items or lemmas in all their varying forms[13], for example in the quasi-modals *HAVE to*, where *HAVE* includes the forms *had, has, have* and *having*, as well as in *WANT to*, where *WANT* includes the forms *want*, *wanted* and *wanting*, and in *BE going to*, where *BE* includes the forms *am, 'm, are, 're, is, 's, was* and *were*, in accordance with the convention of e.g. Leech (2003), Mair and Leech (2006), Leech *et al.* (2009) and Coates (1983). However, in examples from corpora, the convention is not necessary, because the interpretation of a particular form like *has to* is clear from the context (cf. Leech *et al.*, 2009). Furthermore, in accordance with e.g. Leech (2003; 2011), Mair and Leech (2006) and Leech *et al.* (2009), the modals are not used in this way, but are merely italicised (unlike Coates, 1983); for example *can* includes the forms *can, can't* and *cannot*, and *will* includes the forms *will, won't* and *'ll*. This choice was made for comparative purposes with that of Leech and associates. This notation style is however not maintained when

[13] For the spoken ICE-SA data, as discussed in Chapter 3, the forms *wanna, gonna* and *gotta* were also included in the analysis of *WANT to, BE going to* and *(HAVE) got to* respectively (cf. Krug, 2000).

quoting work by others, e.g. Collins (2009a), who uses a different system including only italics for both modals and quasi-modals without small capitals for the quasi-modals.

1.5 RESEARCH QUESTIONS, OBJECTIVE AND GOALS

The six specific research questions that will be addressed in this study are based on the general paucity of research in the field of SAfE grammar and semantics, with modality as the point of departure towards its more comprehensive description. These questions are:

1. Has the use of modal and quasi-modal verbs changed in the (relatively short) history of SAfE (around 190 years)?
2. How have the different registers (social and business letters, news, fiction and other narrative texts, and non-fiction) evolved in terms of modal and quasi-modal use, i.e. what noticeable changes took place over time in the relationship between register and modality?
3. Are there reasons for such changes in (1) embedded in the relevant semantics?
4. What influence could social elements such as language contact have on the changes?
5. Are the changes similar or dissimilar to those that occurred in other native varieties of English?
6. Why might SAfE developed in this way in relation to other native Englishes?

In view of the above, the main objective is to conduct an empirical study of SAfE modality based on corpus evidence. Quantitative analyses (frequency and distributional patterns) and qualitative interpretations (semantic, pragmatic, and sociolinguistic explanations) will be carried out on the modal verbs *must, shall, should, can, could, may, might, will, would, ought* and *need* and the quasi-modal verbs *BE able to, HAVE to, BE going to, NEED to, BE supposed to, WANT to, BE to, (had) better* and *(HAVE) got to*. Comparisons of the results with that of other native

Englishes will also be undertaken. The ultimate objective is to provide a well-conceived puzzle piece in the broader description of SAfE grammar.

It is however the pursuit of eight particular goals which will ensure that this objective is reached. These are:

1. The compilation of a historical corpus of SAfE representing the written registers of social and business letters, news reportage, fiction and other narrative texts, as well as non-fiction from the early 19th to the late 20th century.
2. The identification of each occurrence of modal and quasi-modal verbs within the historical corpus as well as within the synchronic corpora of SAfE and Afrikaans.
3. The analysis of distributional patterns and patterns of change in general and in each register, and the presentation of the results in tabular and graphic form, as well as the illustration of the findings via examples from the corpus.
4. The in-depth analysis of each occurrence of *must, should* and *HAVE to* in terms of macro- and micro semantics, namely deontic (high- or median-degree obligation together with the subjective and objective source of the obligation), epistemic and dynamic meanings.
5. The comparison of the discerned semantic patterns across registers.
6. The comparison of the findings with modal distribution patterns in Afrikaans corpora, in order to establish whether or not contact with this language influenced and/or influences SAfE in the case of modality.
7. A general and modal-specific comparison of the findings with those of other native English varieties, both in terms of frequency and semantics.
8. A discussion of the sociohistorical and linguistic implications of the findings and comparisons.

In essence, therefore, I shall trace and investigate the trek of modality within written SAfE, leading up to its contemporary state, across text and context.

1.6 STRUCTURE OF THESIS

After this introductory chapter, Chapter 2 provides a critical review and synthesis of all the relevant theoretical and contextual issues raised and expanded on in the relevant literature. The main theoretical and empirical frameworks I adopt in this thesis will be surveyed, evaluated and integrated in terms of theories of language, English studies and the difficulty of placing SAfE within models of English varieties. Next, the social history of English (and language in general) will be explored, before discussing the South African setting, as well as the most important factors of language change in the social settings of the world and SAfE. Lastly, a discussion of the modal system will shed light on its development across the centuries, and more recently during and after its transplantation into the new territories of the English-speaking world and especially South Africa.

Chapter 3 will document the research methodology of this thesis, with particular emphasis on the general research design, the study population, the corpora, analysis tools and interpretative methods, including parameters for microsemantic analysis based on high and median levels of obligation (see § 2.4.1.1 and § 3.4.1.1). The results of these methodological procedures will be presented in Chapter 4, with the help of e.g. graphs showing trend lines of diachronic development and frequencies of variation, and illustrative examples from the corpora. Furthermore the interpretations, implications, and hypotheses drawn from the results will be expounded and elaborated in terms of the concepts discussed in Chapter 2. This chapter will furthermore present comparisons with other varieties of English, as also discussed in the literature review. The final chapter, Chapter 5, will summarise the findings, present conclusions and explore possibilities for future research.

1.7 CONCLUSION

The capacity and potential of SAfE to be compared with other varieties in terms of grammar from synchronic and diachronic perspectives are impeded by both the general lack of research on the subject and the type of research undertaken by

previous studies. This thesis aims to bridge this gap by providing a corpus-based account of the quantitative changes in the use of modals and quasi-modals, with the goal of explaining synchronic patters via diachronic analysis. Interpretations with a sociolinguistic approach will be employed in order to shed light on the tendencies noted through the analyses. Ultimately, comparisons with patterns found in other native varieties of English will contribute to establish the position of South African English in terms of modality, in comparison with other Englishes around the world.

CHAPTER 2

LITERATURE REVIEW

"...to remove the historical dimension from the study of language would have consequences as serious as removing the historical dimension from the study of man." (Watkins, 1989:798)

2.1 INTRODUCTION

This chapter will provide a critical literature review and synthesis of contextual and theoretical material relevant to the study of modality in SAfE on synchronic and diachronic levels. The chapter covers three main sets of ideas, starting with the general frameworks that create the foundation of the various approaches adopted in this thesis, then moving to an exploration of social history and context, including the factor of language contact, and lastly undertaking an examination of the grammatical features central to this study. Each of these three sets will in turn be divided into three levels of specificity: from the most general notions of language change and their applications in the English language, moving to the most prominent paradigms in studying varieties of English around the world and in particular the southern hemisphere, and culminating with an in-depth discussion of South African English itself.

To begin with, the main theoretical and empirical frameworks that underpin the research to follow in this study will be surveyed, evaluated and integrated in terms of theories of language, English studies and the difficulty of placing SAfE within models of English varieties. Next, an exposition of contextual factors in language change, as embodied in social history across the three levels of specificity, will be presented. This will illuminate the settings in which the above-mentioned theories and models function and are influenced by society in the past and present. The South

African setting is an especially vital element in contextualising the research of this thesis – as indicated in the title – in terms of history and the journey of the English language in the country.

By zooming in on one particular factor of language change, language contact, issues and debates concerning the importance and extent of the influence of this factor in the formation of English, new Englishes and SAfE (where it serves as the link between notions of migration and identity) will be raised and evaluated. Lastly, and on the most specific level, a discussion of the grammatical category investigated in this study, the modal system, will shed light on its development from Old to Contemporary English, its transplantation into the new territories of the English-speaking world and eventually its journey within the confines of South Africa, as is documented in the available literature. A concluding synthesis of the three sets of ideas and three internal levels of discussion will highlight the intricate relationships among these elements, which should enable the formation of a contextual and theoretical platform from which to launch the empirical study.

2.2 THEORETICAL FRAMEWORKS AND THE EMPIRICAL APPROACH

This section explores, evaluates and integrates the principal frameworks and approaches that are vital to the theoretical and empirical components of the present study. The main theoretical framework is the composite field of sociohistorical linguistics, as recently postulated by e.g. Peter Trudgill (2010), which is traditionally divided into the two independent disciplines of historical linguistics and sociolinguistics. Even though they are considered in merged form to create the theoretical foundation of this study, many constituents of the two fields are expounded separately to aid understanding. This is especially done to separate the first two sets of ideas, as mentioned in the Introduction to this chapter, which adhere to English development in general (historical linguistics) and varieties of English around the world (sociolinguistics), before SAfE is later given attention in 2.2.3.

The main empirical approach supporting the above-mentioned theoretical framework is corpus linguistics, as posited by e.g. Biber *et al.* (1998, 1999). Within this field, the sub-discipline of register studies, its approach to real data, as well as its quantitative and qualitative techniques will form the basis of the methodology for the present research. The literature survey explaining this empirical approach is deliberately presented within the same section of Chapter 2 reviewing literature related to theories, since this thesis not only makes use of a historical corpus of a specific variety of English in the southern hemisphere comprised of multiple registers, but also employs theoretical notions to explain empirical results. It is therefore evident that many facets of theory and method interact within the scope of this study and hence a systematic discussion of each element is necessary, before proceeding to further deliberation on contextual and linguistic matters.

2.2.1 Main theoretical framework: sociohistorical linguistics

Sociohistorical linguistics, as mentioned above, is the product of the theoretical fusion between two traditionally separate domains of linguistic study. Despite the more typical disciplinal dissociation of socio- and historical linguistics, many of their central elements amalgamate to form this hybrid approach. These include elements within their main propositions, such as studying a language within its context (historical and social), and elements embodied within their individual sub-theories, like a society's acceptance of new features into their varieties of English across time. In simple terms, language change (the focus of historical linguistics) is bound to a linguistic society (the focus of sociolinguistics) and, unequivocally, a society is bound to its history. Essentially, then, sociohistorical linguistics centres on the social history of a linguistic community, in order to explain the linguistic features of that community (Nevalainen, 1996; Trudgill, 2010). The social and historical elements embodied in the two subtheories are therefore inseparable within the scope of this thesis, as sections to come will explore.

This composited discipline was first advocated by Suzanne Romaine (1982), who emphasised the relevance of sociolinguistics to historical linguistics in using the past to explain the present, by investigating language variation within a particular

speech community over time (Romaine, 1982)[14]. This is exactly the approach of this thesis. Since Romaine (1982), the application of this sociohistorical approach to language has been further explored by e.g. Tieken-Boon van Ostade (1987), Romaine (1988), Kielkiewicz-Janowiak (1992), Milroy (1992), Nevalainen (1996; 2009), Trudgill (2001; 2010), to some extent Croft (2002), and Nevalainen & Raumolin-Brunberg (2003), to name a few. The focus of sociohistorical linguistic studies in these cases has not been static, but different angles and emphases arose within the field. The approach taken by Romaine (1988) still leaned more towards general linguistic study than the actual sociohistorical element, whereas Milroy (1992), in preferring the term historical sociolinguistics, focused more on the sociolinguistic element of study. However, as mentioned above, both the social and historical elements carry equal weight in the present linguistic research.

Perhaps the sociohistorical contribution to linguistics most relevant to this study is Peter Trudgill's (2010) application of this viewpoint to describe English varieties. He adopted the sociohistorical approach in order to "find solutions to puzzles about why certain language varieties are like they are" (2010:181) and came to the conclusion that these 'solutions' were consistently interconnected with contact between language varieties. This issue of language contact as a pivotal factor in language change is indeed at the core of this study and will be presented later in this chapter.

Owing to the fact that this discipline has its roots in both historical linguistics and sociolinguistics, each of these fields will receive individual attention in the sub-sections to come. While undertaking these separate discussions, the following sections not only seek to explore the literature concerned with these theories, but moreover seek to raise questions and strive towards answers regarding language change in its essence: the reasons for, mechanisms and processes in, and of course context of, this change.

[14] Earlier than this, Labov (1963) already implied the link between social factors and language change within a linguistic community, but it is Romaine (1982) that first made the link explicit. Labov indeed also continued to deal with these two fields in conjunction (e.g. Labov, 1966; 1994; 2001; 2010).

2.2.1.1 Historical linguistics

The primary interest in the field of historical linguistics, which is also called diachronic linguistics, is language change. Historical linguistic investigations typically ask questions about the history of language families or related languages and the history of particular languages regarding phonological, grammatical, lexical, semantic and orthographic features (Baugh and Cable, 1978; Lehmann, 1973; Lass, 1987; Hock, 1991). However, historical linguists may also, as their investigations go further and further into history, find themselves asking deeper theoretical questions about language history in general, i.e. the origin of human language (Lehmann, 1973:242). This thesis does not address this theoretical issue as such, but incorporates insights from e.g. Tomasello (2000) in the next section in order to understand how languages gradually acquire structure through the process of grammaticalisation.

As for the more typical questions asked in historical linguistic investigations, which are mentioned above, it is important to note the orientation of the approach itself. Though it seems apparent from the name *diachronic linguistics* that this approach only deals with the development of a language across time, it does not exclude synchronic analysis in principle, seeing that the state of a language at a particular point in time is useful to provide a comparative base for linguistic change. Therefore, both perspectives are required for a complete description of a language (Lass, 1987:24).

The notion of diachrony and synchrony being compatible areas of investigation has however not always been as clear-cut in the discipline itself. The development of historical linguistic ideas will therefore receive attention in this section, by firstly considering the controversies involving diachronic and synchronic description (which include issues of standard and variation), secondly exploring texts as a means of investigation and thirdly emphasising the importance of considering cultural and social factors.

The first consideration when conceptualising the historical linguistic discipline is time orientation, moreover the diachronic orientation, but underlying to this is the idea of language change being perceived as decay. This essentially links with the issue of language standard, where decay casts the idea of ‘standard’ as the arresting of change. In writing the first grammars for e.g. the English language during the 18th century, the notion of standard went in tandem with “the general belief that language

was of divine origin and hence in its beginnings perfect, but that it embarked upon a long process of deterioration as soon as man, ungoverned by reason and logic as he was likely to be, began to use it" (Pyles, 1965:293). Indeed, Hock (1991:1) states that this notion of language decay has been present since time immemorial, owing to negative connotations to language change in religious stories, e.g. the divine intervention at Babel to create confusion (from the Bible[15]) and the barbarian corruption of speech in the story of the Asuras (from Hindu mythology). Both these religious accounts render language change as being highly undesirable (Hock, 1991:1).

According to this approach, language was seen as being perfect at one state in time (synchronic) and on a constant downhill path of decay (diachronic). Already these chronological approaches to language seemed totally incompatible. Pyles states that by the time that the 18th century grammarians established themselves as authorities to keep language in check, because they felt obligated to prevent its degeneration, it never occurred to them that English had progressed much during the course of about eleven centuries, without any regulation[16] other than that of usage (1965:293). Hock largely agrees with this, and also provides a glimpse of how the progression of English, as noted by Pyles (1965), actually occurred:

> "...language change is not completely random, unprincipled deviation from a state of pristine perfection, but proceeds in large measure in a remarkably regular and systematic fashion." (Hock, 1991:2).

Indeed, as the exposition on corpus linguistics (§ 2.2.2) will later reveal, this thesis selects the usage-based approach to language, as mentioned by Pyles (1965:293) and seeks to investigate the "systematic fashion" of language change, as postulated by Hock (1991:2) above.

In stark contrast to the very dismissive 18th century attitude toward language change as a possibility of study, the view generally held in the 19th century, before the rise of structuralism, was that linguistic explanation was "simply to be equated with historical explanation" (Greenberg, 1979:277-278). The word *synchrony* was not even used before Saussure, and during this period the possibility of synchronic explanation

[15] Genesis 11:1-9 (Bible, 1911).

[16] Except of course for spelling, which has been subjected to a measure of standardisation since William Caxton initiated printing in England.

was simply ignored (1979:278). It is however important to note that this period also saw the differentiation between the ideas of linguistic explanation and linguistic description, of which the former was perceived to abide with an exclusively diachronic approach and the latter with the older approach of the 18th century (1979:278). In the early 20th century, the Neogrammarian Paul (1909:20) declared that the older, merely descriptive approach to grammar gave rise to the study of historical grammar, and also famously stated the following (translation taken from Greenberg, 1979:278):

> "It has been alleged that there is yet another scientific view of language than the historical. I must reject this. What is described as a non-historical, yet scientific, view of language is basically nothing but an incompletely historical one."

Despite this emphasis on historical views during the late 19th century, the coming of structuralism in the early 20th century (e.g. Saussure, 2011 [1916]) reversed this idea completely and foregrounded the autonomy of synchronic structure from language history, which entailed that synchrony was the primary object of linguistic theory (Greenberg, 1979:287; cf. Kawaguchi *et al.*, 2011:7).

This separation was however not as categorical for all structuralists, as at least the Prague and American structuralist schools did not deem the two fields completely separable, but the dominant view on this was determined by which structuralist school was the most prominent at a certain stage. Greenberg indeed describes the situation as 'noblesse oblige' – the most authoritative school was "in the position of offering new synchronically derived insights to the study of linguistic change" (Greenberg, 1979:287). Despite these fluctuating views and new discoveries, it was not yet clear what relevance historical linguistics as a study field had to linguistic science and its most basic theoretical considerations until the mid 20th century (1979:287).

Academic interest in language change reappeared roughly during the 1960s and especially the mid 1970s (Greenberg, 1979). During the same period, another linguistic discipline rose to prominence, namely sociolinguistics, which contributed to the interest in processes of change (1979:276). In the rising sociolinguistic approach, the Saussurean dichotomy of *langue* (the abstract system of language) versus *parole* (the linguistic performance of individuals) (Saussure, 2011 [1916]) has undergone a weakening, as the focus shifted to variation within a language community, usage and the process of on-going change (Greenberg, 1979:276). This consequently

strengthened the inclination toward a diachronic approach to language. William Labov, a prominent figure in the rise of sociolinguistics, already related language change to synchronic variation during this time (e.g. Labov, 1963; 1966).

As mentioned previously, Saussure, the father of structuralism, drew a clear division between the synchronic and diachronic approaches to language. He posited that the study of language at a given point in time should be kept strictly separate from the study of language change, which correlates with his general notion that a language should be studied as a system of oppositions (Saussure, 2011 [1916]; Croft, 2002:75). He describes the objective of one part of his study in *Course in General Linguistics* as follows:

> "The ideal would be for each scholar to devote himself to one field of investigation or the other... ... Whichever way we look in studying a language, we must put each fact in its own class and not confuse the two methods. The two parts of linguistics respectively, as defined, will be the object of our study. Synchronic linguistics will be concerned with the logical and psychological relations that bind together coexisting terms and form a system in the collective mind of speakers. Diachronic linguistics, on the contrary, will study relations that bind together successive terms not perceived by the collective mind but substituted for each other without forming a system."
> (Saussure, 2011 [1916]: 99-100.)

However, by the mid to later 20th century linguistic theory was slowly moving away from structuralist ideas (Greenberg, 1979:276), and the Saussurean dichotomy too was becoming blurred; in 1976 Traugott (1976:502) stated that "(i)n recent years the prospect that the assumed relationship between diachrony and synchrony might change has been shadowy but present". By this time it was possible to discern two aspects of the renewed diachronic emphasis: the first aspect is the extensive application of a dynamic framework for synchrony[17], as foreshadowed by notions of both the Prague school and some Neogrammarians, and the second aspect is the emphasis on processual generalisations, which is related to the interest in linguistic universals (Greenberg, 1979:276-277).

The growing idea of compatibility between these two chronological orientations was nonetheless met with some resistance. David Cohen (1974) published a whole series of papers advancing the assumption that a synchronic

[17] Kawaguchi *et al*. (2011:8) note that "a [dynamic] conception of synchrony assumes that similar patterns of usage can coexist in a community during a certain period and that their mutual relations are not static but conflicting enough to give rise to a future systematic change through symptomatic synchronic variation".

grammar of a language is in fact a theory of that language (Greenberg, 1979:277). In his volumes Cohen (1974) endeavours to investigate the explanatory role of synchronic grammars in linguistic behaviour, among other things, and consequently, the diachronic dimension of language is simply ignored by him (Greenberg, 1979:277). This particular point of investigation contrasts directly with the previously stated differentiation between linguistic explanation and description, of which explanation was seen as being an exclusively diachronic approach to language, as noted by Greenberg (1979:278). Cohen (1974) instead sees the synchronic approach as the only means of explaining linguistic phenomena – to him this was indeed an issue of applying synchronic investigations as a theory of language (Greenberg, 1979:277).

Another very prominent linguist argued along the lines of Cohen's idea of synchronic explanation and indeed took the idea of synchronic 'theory' to the next level: Noam Chomsky. In Newmeyer's (1986) publication, *Linguistic Theory in America,* there is only one mention of historical linguistics, as is quoted here:

> "Chomsky attributes his early interest in explaining linguistic phenomena, as opposed to simply describing them, to his childhood exposure to historical linguistics. In a period when leading theorists tended to look upon the desire for explanations as a sort of infantile aberration, historians of language like his father, either ignorant of or indifferent to the contemporary "scientific" wisdom in the field, clung to a 19th century desire to explain why a particular distribution of forms existed at a particular point in time." (Newmeyer, 1986:28-9)

Watkins (1989), in a comment on this statement, discloses that he personally "checked this statement with Professor Chomsky, and he assured [him] it was 'completely accurate'" (1989:783). It can therefore be deduced that Chomsky altogether dismissed the notion of diachronic explanation, but, surprisingly, drew inspiration from it to fuel his explanatory, synchronic 'theory' of sorts[18]. Watkins (1989:783) maintains that, by this account, the destiny of historical linguistics was "to influence the young Chomsky, and once that function was discharged it could sink back into the decent oblivion whence it had come."

---

[18] Indeed Chomsky's notion of universal grammar is inspired by the *Port Royal Grammar* (the *Grammaire générale et raisonnée* by Arnauld & Lancelot, 1660), which he claims was influenced by the (rationalist) Cartesian approach to language (Chomsky, 1966). Chomsky (1966) perceives the *Port Royal Grammar* as an explanatory work, which arose in reaction to previous descriptivist work by e.g. Vaugelas (1647). Percival (1976) disagrees with this, and argues that such 20th-century terms do not seem to apply to either these 17th-century intellectual products.

Despite strong resistance like Cohen and Chomsky's, the idea of complete separation between diachrony and synchrony was becoming more and more obsolete during the last half of the 20th century. Pyles, in the 1960s, defines a historical grammar of a language as a "consecutive and concatenated set of descriptive grammars of that language" (1965:295). Thus, he argues that it is "impossible to regard historical (diachronic) studies and descriptive (synchronic) studies as mutually exclusive" (1965:295). Baugh and Cable (1978:xiii) are convinced that the "soundest basis for an understanding of present-day English and for an enlightened attitude towards questions affecting the language today is a knowledge of the path with it has pursued in becoming what it is." Hereby Baugh and Cable propose that synchronic questions could be answered via diachronic revelations.

In the late 1970's Greenberg (1979:286) seeks to emphasise the scientific contribution that a full and systematic consideration of diachronic factors can make to solve linguistic problems in the broader field of linguistics, i.e. synchronic studies. Lass (1987:24) argues more strongly for continuity between the two approaches, suggesting that the one cannot be isolated form the other in describing language:

> "We generally describe the structure of a language at a given time in terms of a set of fixed categories and rules; and this level of description does not have to take account of the fact that languages exist in time, and are always changing. In a sense we cannot describe processes of change without prior description in terms of fixed categories. ... A 'complete' description of a language (if such a thing were possible) would take into account both its time-bound and stable properties." (Lass, 1987:24)

Here Lass suggests that synchronic study can more easily be done without considering diachrony, than diachronic studies can be done without the acknowledgement of synchrony. Yet, he deems both approaches essential in order to investigate the nature of language (Lass, 1987:24). It is interesting to note the word choice in the above quote, as Lass refers to "description" (1987:24) – a term which has generally only been used in connection with synchronic analysis, as noted in previous paragraphs, but in this quote is used to emphasise how both synchrony and diachrony would aid in complete linguistic 'description' as such.

Similarly, Southworth (1990:25) states the following:

> "Linguistic change manifests itself synchronically as variation, which can be defined in a somewhat general way as a relationship between different ways of expressing the same thing."

In speaking about synchronic manifestations of linguistic change, Southworth indeed consolidates the two orientations into a 'relationship' – a kind of continuum of temporal linguistic materialisation. This thesis is aligned with the views[19] of e.g. Pyles (1965), Baugh and Cable (1978), Greenberg (1979), Lass (1987) and Southworth (1990), in viewing language from both synchronic and diachronic angles – the one never being totally independent of the other, but rather functioning as two sides of the same coin. In the quote above, Southworth essentially states that linguistic variation is a synchronic indication that changes have occurred in the past (1990:24). Similarly, Denison (1998) suggests that language variation provides a glimpse of language change in progress, by highlighting the present, ongoing changes in the synchrony of a language (cf. Kawaguchi *et al.*, 2011:8). Linguistic variation is naturally a very important constituent of this thesis, and does indeed strongly abide with sociolinguistic ideas, as will be elaborated on in sections to come.

The very notion of variation however conflicts with the previously stated 18th century ideals of elevated standard, language deterioration and, hence, language errors and debasement, as noted by e.g. Pyles (1965) and Hock (1991). The term 'variation' implies a sense of tolerance and acceptance of supposed anomalies in language, as sections to come will discuss. On this note, Pyles argued that the historically trained linguist is bound to be "tolerant of the supposed blunders of unsophisticated speakers, while realizing perfectly well that the sole criterion for determining the status of a given linguistic form at a given time is, as it has ever been, current usage" (1965:297). The speakers of a language, and not some exalted linguistic authority, can therefore be said to determine the face of a language – language users indeed propel language change, which then becomes manifest in their current usage.

The discipline of historical linguistics has continuously met with some resistance in this particular matter. Pyles expressed his personal experiences and feelings on the challenges of the historical linguist in his academic environment regarding attitudes toward language variation:

> "...when the historical linguist becomes involved ...in controversies over what is 'good English' and what is 'bad', his disinterested ...attitude is practically always

[19] This also includes the general view adopted by Labov (1994; 2001; 2010) regarding diachronic and synchronic interplay, but other sections (§ 2.2.1.2 and § 2.3.1) will deal with Labov's views more specifically, although his concerns about historical linguistic methodology will receive attention later in the current section as well.

> misunderstood, not only by the general public but frequently ...by those of his own literary colleagues who have forgotten, or more likely insufficiently learned, what should have been obvious from their minimal training in historical linguistics. He is stigmatized as a 'liberal' (a smear-word nowadays in this particular context), as 'permissive', as one who 'has no standards', all because he conceives it to be his primary function to record, study, and, if he is lucky, to make some significant comment on the facts of linguistic change." (Pyles, 1965:297-8)

Apart from these earlier stigmas that e.g. Pyles experienced in academia, even as early as 1965, it is the methodological approach of the historical linguist that perhaps deserves mention as a greater issue of concern. Not only does his methodology attribute to the linguist's pleasure in conducting interesting research, but also arguably poses his/her greatest challenge. The main methodology of historical linguistics of course involves the study of older texts, and given the dynamics between synchrony and diachrony, contemporary ones as well.

Lehmann indeed states that "(m)aterials of concern in historical linguistics are available primarily through written records called texts" (1973:59) and that the fact of language change becomes most apparent when reading a text that is a few centuries old (1973:1). Texts can therefore be seen as both the initiators for historical linguistic study (inspiring or prompting the linguist's interest), and the primary source of the actual linguistic investigations or analyses that are to follow. The historical linguist, as much as the general linguist, has a "time-honoured technique for dealing with such texts" (Watkins, 1989:785). This technique is called *philology*, which, like pragmatics, involves "the study of the meaning of language forms as these depend on the linkage of signs to the context in which they occur (we call this the 'indexical' meaning of signs)" (Watkins, 1989:785; cf. Silverstein, 1981). A dictionary definition of philology is "the branch of knowledge that deals with the structure, historical development, and relationships of a language or languages" (Concise Oxford English Dictionary, 2006:1077).

The text as primary source and basis for investigation ties in with the empirical framework of this study, namely corpus linguistics, which will receive attention later in this chapter and in Chapter 3, when the methodology is described. However, despite the textual perspective of both philology and corpus linguistics, the latter is not so much interested in considering signs in context, but rather in finding similarities (or differences) across many contexts. Seeing that the corpus used for the current research is a historical one, it can be foreseen that certain problems, as

mentioned above, may arise in its compilation and use. The main challenges involving texts will briefly be discussed in the next few paragraphs, and in addition to this, other problem areas of the discipline will receive attention.

Methodologically and generally speaking, Labov foresees two specific problem areas for the historical approach to linguistic study. Relevant to the text as means of study, the first, methodological problem involves what can be called the 'bad data problem' (Labov, 1972:100). This entails that no spoken data are available for earlier periods and, as Labov states, written data are usually "fragmentary" and "corrupted" (1972:100). On the other hand, Romaine (1988) notes that the lack of older spoken data should not prevent the reconstruction of a language in its social context and that written material can and should be studied in its own right. Labov however further asserts, as stated above, that the discipline of historical linguistics is both driven and restricted by its data, which are "rich in so many ways" but "impoverished in others" (1994:11), as historical documents survive solely by chance and not by design. The data available for the linguist to work from are therefore the "product of an unpredictable series of historical accidents" (1994:11). Labov (1994:11) elaborates on this:

> "The linguistic forms in such documents are often distinct from the vernacular of the writers, and instead reflect efforts to capture a normative dialect that never was any speaker's native language. ... Furthermore, historical documents can only provide positive evidence. Negative evidence about what is ungrammatical can only be inferred from obvious gaps in distribution, and when the surviving materials are fragmentary, these gaps are most likely the result of chance. Historical linguistics can then be thought of as the art of making the best use of bad data."

Rauch (1990) on the other hand, argues for the usefulness and credibility of texts, as he deems data "untouchable" to the extent that "no amount of manipulation can alter evidence; it is what it is" (1990:36). Rauch therefore suggests that textual evidence is reliable enough to conduct linguistic investigation and support hypotheses, as it remains the same throughout the ages and, in a sense, is in plain sight for all to see (1990:36). He thereby shuns the idea that data can be manipulated to present skewed evidence to prove certain predetermined ideas, or can be interpreted differently by different linguists or even misinterpreted by individuals (1990:39). Rauch (1990:39) states:

> "...despite one's predilection for method, whether philological or abstract, whether scientific or humanistic, the evidence always withstands individual conviction. Language change is more readily traceable from evidence in past changes than from evidence from contemporary change."

In this quote Rauch further suggests that the older the data are, the better the evidence for language change will necessarily be. It can therefore be deduced that it is better to employ texts a from as wide a temporal range as possible, to ensure a truly comprehensive data set for the detection and analysis of language change. Chapter 3 will reveal that this endeavour was carried out for the collection of the corpus used in this study. Furthermore, in the more recent past (from about the late 19th century onwards), written texts do become increasingly rare, which to some extent solves Labov's (1972) earlier-mentioned 'bad data problem' in terms of the data for this thesis.

The second problem Labov foresees for historical linguistics, which has to do with more general considerations, lies within what Nevalainen (2009:599) calls the 'historical paradox'. Labov (1994:9) asserts that we know that language was different in the past, but we do not know how it was different. He states:

> "The fact of language change is a given; it is too obvious to be recorded or even listed among the assumptions of our research. Yet this fact alone – the existence of language change – is among the most stubborn and difficult to assimilate when we try to come to grips with the nature of language in general as it is reflected in the history of a language."

When considering this complex task of following and describing the history of a language, contextual factors – those conditions outside the text (containing evidence for language change) – are of course pivotal in understanding the text itself and the reasons for such detected change (Lehmann, 1973:240). On this note, Baugh and Cable (1978) distinguish between internal history (the structural development of a language) and external history (those political, social and intellectual forces "that have driven the course of that development at different periods" (1978:xiii). A text itself can therefore not be regarded as an isolated piece of data, but the world outside each text should be taken into consideration when conducting historical linguistic investigations.

The present study focuses exactly on this – both the elements of text and context – in order to investigate and explain linguistic phenomena within a linguistic

community. Within the framework of historical linguistics, this would naturally include the use of historical texts, while considering the historical events and conditions surrounding those texts. It is when we move to this textual-contextual approach to investigation that the lines between the theoretical frameworks relevant to this thesis become blurred as they intersect at various points, and, additionally, those theoretical frameworks in turn interact with the chosen empirical framework for the study. This text-context relationship therefore individually encompasses and universally links all areas discussed in this chapter.

Various forces are constantly at work within the contextual space surrounding the texts, which constitute ongoing language change. As a more complete discussion of such forces will be undertaken in the section on sociolinguistics, only a few overtly related to language history will receive attention here. As already mentioned above, Baugh and Cable (1978:xiii) suggest that external linguistic history or context can be shaped and influenced by political, social and intellectual forces. In linguistic change, such external factors can include highly perceptible elements like great historical events involving political climate, social conditions and widely-held perceptions of a certain period, but it can also include less obvious forces like individual creativity, ingenuity and, on the other hand, human stubbornness (Watkins, 1989:793). Watkins holds that, if one were to devise a 'proper' linguistic theory, one "must be able to account for the creativity of human language" and "must also account for the possible long-term preservation of surface[20] formulaic[21] strings[22] in the same or different linguistic traditions over millennia" (1989:793). This suggests that certain aspects of a language may change suddenly, whereas others may remain the same across many generations of users. The opposite can also occur. Croft (2003:239) writes that

[20] The surface structure is "[a] central theoretical term in transformational grammar, opposed to deep structure" (Crystal, 2008:466). Crystal (2008:466) adds that "[t]he 'surface structure' of a sentence is the final stage in the syntactic representation of a sentence, which provides the input to the phonological component of the grammar, and which thus most closely corresponds to the structure of the sentence we articulate and hear".

[21] This essentially refers to linguistic formulae, e.g. formulaic expressions in prayers (cf. Watkins, 1989:793) and also politeness formulae, which can remain in use over extended periods of time.

[22] Crystal (2008:456-7) defines 'string' in terms of its use in generative grammar as "a linear sequence of elements of determinate length and constitution." He adds that a "[f]ormal analysis also permits the notion of a string consisting of just one short element, and also one consisting of no elements... – the empty or null string." An example of a string of elements can be seen in the following sentence from Crystal (2008:457): "*the+cat+sit+Past+on+the+mat*."

"'[r]elatively unstable' language states may persist for a long time, perhaps centuries", but also that "[s]tability is ...controlled by nonlinguistic forces..., such as the enforcement of a written standard" (cf. Croft, 2000).

Labov's (2010) argument is much the same. He emphasises the sometimes sporadic character of language change, because although linguistic (internal) forces usually operate with some regularity, social (external) forces, for example the population, migration, contact, etc., on the other hand, can lead to sudden changes in circumstances that obstruct the normal, slow and gradual change (cf. Labov, 1994; 2001). Here he distinguishes, as extremes of language change, gradualism from catastrophes. He states:

> "It goes without saying that that any given state of a language is the outcome of a previous state of that language, and so on – back in time a far as our knowledge van carry us. ... In an endless chain of causes, every state of the language is a triggering event for the one that follows. Even if there is no detectable change in a given system, the system itself has a cause: the state of equilibrium that was reached in the preceding period." (Labov, 2010:90)

Labov (2010:90) adds that the "sequence of preceding causes" of a triggering event is by no means "smooth and uniform", but that there "are bends in the chain of causality at which the triggering events are located". These kinds of events (comparable with the extinction of the dinosaurs in terms of the development of mammals) could then account for a (catastrophic) disruption in the slow, gradual change in language (2010:90-1). Linking with the kinds of change related to temporality, Labov also proposes a method for detecting changes across time. Labov (1994) describes a model of language change through 'real time' analyses – that is through observing "a speech community at two discrete points in time" (1994:73): it relies on the pattern of how "individuals change or do not change during their lives, how communities change or do not change over time, and what may result from combinations of these possibilities" (1994:83).

It appears from Watkins's (1989), Baugh and Cable's (1978), Croft (2000; 2003) and Labov's (2010) opinions above that there are multiple contextual forces simultaneously at work in the process of language change, which can sometimes happen slowly over a period of time (gradual), or with quick bursts because of human creativity or external events ('catastrophes'). A degree of stability in language can be present when change happens gradually, but stability can also be gained more quickly

with e.g. standardisation. On the other hand sudden changes can lead to instability in language, but instability can also occur over extended periods. Language change can therefore be said to be a dynamic and ongoing process (Lehmann, 1973:1), at constant interplay with its own linguistic (textual) and extra-linguistic (contextual) environment – both at a specific point on the continuum that is linguistic 'usage time' (synchronic) and across many years, decades, centuries and even millennia (diachronic). Croft (2003:283; emphasis original) sums up this view (cf. Croft, 2000):

> "Language is fundamentally **dynamic**, at both the micro-level – language use – and the macro-level – the broad sweep of grammatical changes that take generations to work themselves out... Synchronic language states are just snapshots of a dynamic process emerging originally from language use in conversational interaction."

In this dynamic approach to the study of language change, the historical linguist can neither ignore evidence procured from textual data, nor the milieu that surrounds that data, be it synchronic or diachronic. If he or she neglects one of these elements in a linguistic study, the consequential lack of interplay between the aspects will render the investigation one-dimensional and wanting. Indeed to leave the entirety of the historical linguistic discipline out of general linguistic investigations will be an even greater, more unthinkable loss than omitting one or more of the elements listed above, as Watkins (1989) states in the quote given in the opening of this chapter.

2.2.1.1.1 Grammaticalisation theory

An important sub-theory within the field of historical linguistics is grammaticalisation theory. This theory is at the core of contemporary studies of language change. Heine (2003:575) describes it as follows:

> "Grammaticalization theory is neither a theory of language nor of language change; its goal is to describe grammaticalization, that is, the way grammatical forms arise and develop through space and time, and to explain why they are structured the way they are. Grammaticalization is defined as a process which is hypothesized to be essentially unidirectional."

This theory is therefore not a general theory about the nature of language, but focuses on a specific process at work in language change, namely grammaticalisation. Heine

(2003:575;578) further suggests that grammaticalisation theory is a theory to the extent that it offers both a descriptive and explanatory account of how and why grammatical categories arise and develop.

Grammaticalisation is generally described as the process of transforming lexical categories to grammatical (functional) categories (Hopper, 1996:217; Heine, 2003:575). Heine (2003:575) however argues that the process is not confined to only the development of lexical into grammatical forms, but that grammatical forms can frequently give rise to even more grammatical forms. He further asserts the following:

> "Since linguistic items require specific contexts and constructions to undergo grammaticalization, grammaticalization theory is also concerned with the pragmatic and morphosyntactic environment in which this process occurs." (Heine, 2003:575)

This quote would suggest that grammaticalisation theory could include various intra-linguistic and extra-linguistic elements and environments on its path, and is not confined to the lexical and grammatical domains. To be sure, this process has both a synchronic and a diachronic dimension, but its natural foundation is built around the diachronic aspect (2003:575).

In the development of grammaticalisation theory, three stages can be recognised, as described by Heine (2003). Phase one begins with the work of 18th century French and British philosophers, including that of Étienne Bonnot de Condillac (1793), who claims that "grammatical complexity and abstract vocabulary historically derive from concrete lexemes" (Heine, 2003:575-6), and the work of John Horne Tooke (1857), who argues that "language in its 'original state' is concrete, and abstract phenomena are derived from concrete ones" (Heine, 2003:576). Phase two is mainly associated with German linguists from the 19th century, e.g. Bopp and Windischmann (1816) and Bopp (1833). To Bopp (1833), grammaticalisation is a key notion in linguistic study; he asserts that the change from lexical to grammatical forms is an essential component of a comparative grammar (Heine, 2003:576).

Heine mentions that no major developments in the field occurred during the course of the 20th century prior to 1970 (2003:576), but this is not surprising given the previously stated resistance of e.g. Chomsky to consider the historical component of language, no-doubt including sub-theories relevant to diachronic linguistics, worth studying (see § 2.2.1.1). Phase three ensued in the 1970s with the work of Talmy Givón (1971), who argued that one must have a knowledge of the earlier stages of

development for a language in order to understand language structure (Heine, 2003:576). Givón is perceived to have opened a new perspective for understanding grammar with his slogan "today's morphology is yesterday's syntax" (Givón, 1971:12), which he considered to be part of a general cyclic evolution (Heine, 2003:576). The notion of evolution in language will receive further attention in subsequent paragraphs in this section, as well as in § 2.3.1 when general reasons for language change are discussed.

During this third phase, a number of studies centred on grammaticalisation in specific domains of grammar (2003:578), e.g. the studies on modality by Traugott (1989) and Bybee and Fleischman (1995). Heine (2003:578) mentions that grammaticalisation studies during this phase have been very helpful in understanding language change under conditions of language contact and language transmission. These are also very important aspects to be considered in this study, as will be seen in § 2.3.

It is only during the past two decades of linguistic studies that the process of grammaticalisation has established itself as a major research focus area in typological and historical linguistics (Lindquist & Mair, 2004:ix). Towards the end of the 20th century, a new mutually beneficial relationship between grammaticalisation theorists and corpus linguists arose, expanding the research potential and methodological possibilities of both the study fields through their interaction (2004:ix). This is because the empirical methods of corpus-based studies enabled the phenomenon of grammaticalisation to be analysed and systematised via real linguistic data, whereas grammaticalisation in turn opened up an exciting new realm of prospective research topics for corpus linguists.

Grammaticalisation, apart from being the motor behind many contemporary linguistic studies, is indeed the driving force behind language evolution itself (which Croft [2000, 2002] suggests is similar to evolution in biology), introducing new constructions into a language and hence changing its grammatical system (Hopper, 1996:217-8). The key notions behind the process of grammaticalisation necessitate time travel to the dawn of language itself. Tomasello (2000:162) argues that the process of grammaticalisation was the starting point of, and is the driving force behind, language evolution. At some point in history, humans evolved the ability to communicate with one another using symbols; this is extended into longer symbolic discourse interactions, which gradually take on structure, i.e. they become

grammaticalised (2000:162). He states the following:

> "These transformations of linguistic structure occur as a result of social-interactive processes in which (1) speakers try to abbreviate linguistic expression as much as they can, and (2) listeners try to make sure that speakers do not go so far in this direction that the message becomes incomprehensible." (Tomasello, 2000:163).

Similarly, Heine (2003:578) stresses that grammaticalisation theory is based on the assumption that the "main motivation underlying grammaticalization is to communicate successfully". He adds:

> "To this end, one salient human strategy consists in using linguistic forms for meanings that are concrete, easily accessible, and/or clearly delineated to also express less concrete, less easily accessible, and less clearly delineated meaning contents. To this end, lexical or less grammaticalized linguistic expressions are pressed into service for the expression of more grammaticalized functions." (Heine, 2003:578)

According to Heine, then, grammaticalisation is a process whereby expressions for concrete or source meanings are used in specific contexts to encode grammatical or target meanings (2003:578).

Consequently, grammaticalisation involves the 'birth' of new grammatical forms or constructions, resulting in fundamental change and development into the appearance and interpretations of a language's grammatical system (Hopper, 1996:218). Grammaticalisation can therefore involve semantic or functional extension in the use of a construction (Lichtenberk, 1991:475), and is not restricted to a linear type of lexical to grammatical expansion (as stated above) (Heine, 2003:575). Mair (2004:125) believes that, "given the right context, sudden and otherwise-unaccounted-for increases in discourse frequency may be regarded as symptoms of incipient grammaticalization". For already discernable grammaticalisation processes, discourse frequency is however expected to rise "in proportion to the increasingly grammatical nature of a word, morpheme or construction" (Mair, 2004:125). A gradual pattern is therefore more likely to be found as grammaticalisation progresses steadily on the path of evolution (Mair & Leech, 2006).

Mair and Leech further (2006) state that syntactic change differs from lexical change: firstly, it generally unfolds more slowly (sometimes taking centuries to complete) and secondly, it is disposed to "proceed below the threshold of the speaker's conscious awareness" (2006:318). Where such syntactic change is involved,

two general approaches exist: the rapid and isolated "development of grammars as decontextualised linguistic systems" (2006:318), as found in the generativist tradition of e.g. David Lightfoot (1979), and the gradual evolution of recorded performance data in its linguistic and social context, i.e. the usage-based approach (Mair & Leech, 2006:319). The latter is the general approach applied in historical linguistics and grammaticalisation theory and is elucidated in e.g. Hopper (1996). Tomasello (2003:14) sums up the usage-based approach based on Croft (2000): "...it is a historical fact that the specific items and constructions of a given language are not invented all at once, but rather they emerge, evolve, and accumulate modifications over historical time as human beings use them with one another and adapt them to changing communicative circumstances". Drawing on this approach, the current study intends to find evidence of gradual changes in the SAfE modal system by means of a historical corpus, while taking social contexts into account.

Deutscher (2005) in his book on evolutionary linguistics gives a similar account to the above-mentioned idea of gradualism in the extension of functional elements beyond their lexical context, as is suggested in its title: *The Unfolding of Language.* However, in addition to this view, Labov (2010) proposes that this general gradualism of linguistic-internal forces of change can also be disrupted by the discontinuous effects of external forces, such as society (cf. Labov, 2001), as is also mentioned in the previous section. One such external, non-gradual force related to society (see e.g. Labov, 2001) is language contact, which will receive attention in later sections in this chapter.

Leech (2003:223) indeed believes that significant processes of grammaticalisation could be apparent within as little as one generation of speakers, but only a glimpse of the full course of the process would be witnessed through this, as the gradual syntactic changes within a language usually emerge during an extended period – as in the case of the modals and quasi-modals in British English and American English (Warner, 1993). Although a period of about 190 years (which is the scope of this thesis) may not be such a large interval in the evolution of a language, it might nevertheless yield some noteworthy changes signalling incipient processes of grammaticalisation.

The question of when incipient grammaticalisation can be recognised is therefore very relevant to this study. Hopper and Traugott emphasise the importance of studying "patterns of usage, as reflected by the frequency with which tokens of

[the] structures may occur across time" (1993:59), which means that shifts in textual frequency are significant indicators of early grammaticalisation. Grammaticalisation does however not proceed because of frequency, but because "successive generations of speakers perceived [a] phenomenon and developed it further" (Mair, 2004:139). The fact that a construction does not become more and more frequent will therefore not absolutely and necessarily point toward no grammaticalisation whatsoever, as this may also involve a 'static' process independent of statistics (2004:138). Mair (2004:138) argues:

> "For the 'dynamic' instances of grammaticalization (that is the diachronically directed ones which leave a statistical imprint in discourse) ...grammaticalization is not accompanied by a simultaneous across-the-board increase in discourse frequency. Such instances ...should rather be seen as a delayed symptom of earlier grammaticalization... Changes in relative or proportional frequencies, however, will usually be part of the central phase of the process of grammaticalization itself."

It can therefore be deduced that the process of grammaticalisation is gradual and dynamic in nature. Traugott and Trousdale (2010:19-20;39) emphasise the gradual nature of grammaticalisation, but also highlight the relationship between diachronic gradualness and synchronic gradience (or blurred lines between linguistic categories or structures) in the process of grammaticalisation. Synchronic gradience can be perceived as language variation and variation over time involves the emergence of grammatical constructions, which Traugott and Trousdale suggest is "a gradual, global process involving discrete local micro-reanalyses" (2010:39).

The mechanisms behind such dynamic change will be explored in the next few paragraphs, as found in Heine (2003:578-9). The four interrelated mechanisms in the grammaticalisation of linguistic expressions are as follows:

(i) desemanticization (or 'bleaching', semantic reduction): loss in meaning content;

(ii) extension (or context generalization): use in new contexts;

(iii) decategorialization: loss in morphosyntactic properties characteristic of the source forms, including the loss of independent word status (cliticization, affixation)

(iv) erosion (or 'phonetic reduction'), that is, loss in phonetic substance.

As can be seen, the first mechanism of Heine relates to semantics, the second to pragmatics, the third to morphosyntax and the fourth to phonetics (2003:579). He

notes that despite the fact that (i), (iii) and (iv) involve a loss in properties (losing semantic, morphosyntactic and phonetic substance), they can also gain in "properties characteristic of their uses in new contexts... sometimes to the extent that their meaning may show little resemblance to the original meaning" (2003:579). These mechanisms are however not confined to grammaticalisation, but they can be said to constitute distinctive components of the same general process to the extent that they are jointly responsible for grammaticalisation to occur (2003:579).

The mechanisms described above give rise to evolution in language (Heine, 2003:579), which Heine (1993:48-53) presents in a three-stage model – the so-called 'overlap model'. The three stages in this model unfold as follows:

(i) There is a linguistic expression A that is recruited for grammaticalization.

(ii) This expression acquires a second use pattern, B, with the effect that there is ambiguity between A and B.

(iii) Finally, A is lost, that is, there is now only B.

Within these stages, the process of grammaticalisation reveals a chain-like structure, but indeed not all instances proceed to stage (iii), as the process might be hindered at stage (ii) (2003:579). However, Heine (2003:579) states, "once stage (iii) is reached, B tends to be conventionalized, that is, it turns into a new grammatical category".

The description of grammaticalisation as the loss of one structure and the gradual birth of another – either a change from lexical to grammatical status, a change from one grammatical structure into another, or the semantic extension in the use of a construction (cf. Lichtenberk, 1991; Hopper, 1996; Heine, 2003) – is however perceived as problematic by Boye and Harder (2012) in their recent article in *Language*. They argue that these kinds of explanations make the process of grammaticalisation difficult to define. This is caused by "the problem of defining in a satisfactory manner the distinction between lexical and grammatical expressions" (2012:2). This section has considered the views of e.g. Mair and Leech (2006) and Traugott and Trousdale (2010) in favour of the gradual nature of grammaticalisation, but Boye and Harder (2012:1-2) perceive that such views of a lexicogrammatical continuum add to the confusion.

As a solution to this problem, they propose a discourse-based theory of grammatical status and grammaticalisation, which conceptualises the distinction

between lexical and grammatical expressions – a distinction that, they argue, should be maintained (Boye and Harder, 2012:2). This claim is however "not at variance with the existence of gray zones, or clines (including grammaticalization clines)", as "the fact that a clear distinction cannot be made ...does not entail that the distinction is invalid" (Boye & Harder, 2012:6; cf. Putnam, 2001:38). The theory is described as being 'panchronic', in that it "covers both grammatical expressions considered synchronically and grammaticalization conceived of as the type of change that diachronically gives rise to grammatical expressions" (2012:6). They state:

> "The central idea is that while grammatical expressions and grammaticalization cannot be defined in terms of specific phonological, morphosyntactic, or semantic features, alone or in combination, they can be defined in terms of the ancillary status that grammatical expressions by linguistic convention have in relation to other other expressions. Whereas lexical expressions may or may not, in actual communication, convey the main point of a linguistic message, grammatical expressions (morphemes, words, constructions) are conventionally specified as noncarriers of the main point, serving instead an ancillary communicative purpose as secondary or background elements." (Boye and Harder, 2012:6-7)

Boye and Harder illustrate this with an analogy: "a second violin is not inherently different from a first violin – it just has a secondary function in relation to that of the first violin" (2012:7).

Their theory is based on the notions of discourse prominence (which involves the concepts of discursively primary vs. discursively secondary status), as well as linguistic convention (Boye & Harder, 2012:7). Discourse prominence involves the idea that "some aspects of linguistically conveyed mental content are more prominent than others" (2012:7), which links with notions of cognitive grammar (cf. Langacker, 2008). The phenomenon of linguistic prominence is described as being scalar, in that "a given element can be ranked as higher or lower on a scale of prominence than another expression" (Boye & Harder, 2012:7), which is where primary and secondary function or status comes in. They note:

> "...to be discursively primary is to have higher discourse prominence that all other syntagmatically[23] related elements in the utterance; to be discursively secondary is to have lower prominence than one or more syntagmatically related expressions in the utterance." (Boye & Harder, 2012:8).

[23] 'Syntagmatic' denotes "the relationship between two or more linguistic units used sequentially to make well-formed structures" and is contrasted with 'paradigmatic' (Concise Oxford English Dictionary, 2006:1461).

Linguistic convention[24], on the other hand, can be understood as "reflecting a state of coordination between members of a community that goes beyond individual instances of linguistic communicative interaction" (Boye & Harder, 2012:8; cf. Croft, 2000:95-99). Boye and Harder (2012:8) add that usage and conventions do interact, although they are distinct: conventions are dependent on usage (they gain stability or instability through usage), and usage is dependent on conventions (conventions constrain usage events). Linguistic convention is therefore the concept that "underlies all language-specific properties, from phonetics via structural relations to meaning", yet it also "applies to social meaning ...in cases like positively vs. negatively charged words such as *thrifty* vs. *stingy*" (2012:8-9). This corresponds with Bowerman's (2004b) description of *must* in SAfE as having less illocutionary force than in other varieties (i.e. it is less negatively charged by convention), as noted in Chapter 1.

Boye and Harder (2012:9) link the ideas of discourse prominence and linguistic convention as follows:

> "Prominence assignment as [a] ...usage phenomenon takes place all the time in all speakers, irrespective of what the precise conventions about prominence may be – but if the ...conventional prominence properties function efficiently, they reduce the percentage of cases in which speaker and hearer assignments do not co-incide. ... ...our theory depends on understanding discourse prominence both as a substance ...and as a conventional ...phenomenon. ... This view of discourse prominence as a conventional phenomenon forms the background for our claim that the contrast between lexical and grammatical expressions is one way in which prominence properties are conventionalized."

From the link of discourse prominence with linguistic convention, they propose the following definitions to point out general similarities and differences between lexical and grammatical meaning (Boye & Harder: 2012:7).

(i) Lexical expressions are by convention capable of being discursively primary

(ii) Grammatical expressions are by convention ancillary and as such discursively secondary

(iii) Grammaticalization is the diachronic change that gives rise to linguistic expressions that are by convention ancillary and as such discursively secondary

[24] The role of convention is also noted in Heine's above-mentioned stage (iii), whereby a new grammatical category is born (Heine, 2003:579). But the underlying notions of stage (iii) differ overall from Boye and Harders's idea of linguistic convention when linked with discourse prominence.

Boye and Harder argue that with these definitions their theory therefore offers a "functional link between conventional discursively secondary status and structural differentiation" (2012:39). They further note that the prevailing focus in literature "on development away from full lexical status casts the development of grammatical status in the role of form of loss", and suggest that when we focus "on the new relation that comes into being whereby one linguistic expression assumes the role of structural sidekick for the other, it becomes clear that the output is functionally enriched" (2012:39). To conclude their argument, Boye and Harder (2012:39) state:

> "The point is to look not at the single expression but at the combination of the grammaticalized expression and the expression to which it is attached. ... ...the development whereby elements acquire grammatical status has to be understood by combining discourse and structure, recognizing at the same time that structure emerges from usage and that once emerged, structural facts are not reducible to usage facts."

In summary of the above views, grammaticalisation can be described as the driving force of language evolution through the gradual birth of a new construction (below the threshold of awareness) (Hopper, 1996; Tomasello, 2000; Mair & Leech, 2006). This involves the transformation of one form (either lexical or grammatical) into another (usually grammatical) or the semantic/functional extension of a construction into the domain of another (Lichtenberk, 1991; Heine, 2003). This process can be seen as dynamic and cyclic (Givón, 1971), and can be disrupted by external forces (Labov, 2010). In contrast with some of these views, grammaticalisation is more recently described in terms of the functional 'enrichment' of one form with another, instead of the 'loss' of one form and the rise of another (Boye & Harder, 2012). The above discussions on grammaticalisation will be considered when embarking on the research on modals and quasi-modals in SAfE. First, however, the social aspect of language needs attention, in order to complete the discussion on sociohistorical linguistics.

2.2.1.2 Sociolinguistics

Sociolinguistics, the other constituent of the main theoretical framework adopted in this study, is the branch of linguistics concerned with social factors influencing language variation (Hudson, 1996:11-12). This field in its core studies the extra-linguistic (socio-cultural) context surrounding language and the influence of that context on language (1996:4). This thesis is, as mentioned earlier, largely based on extra-linguistic context. The interaction between linguistic and social variables is at the forefront of sociolinguistic studies. Social variables can include factors like e.g. region, class, age and gender, which "determines variation in language" (Stockwell, 2007:3), and linguistic variables can include anything from a language or a dialect to a style or a register, to a specific linguistic feature in the fields of e.g. lexico-grammar, phonology, semantics or orthography (cf. Stockwell, 2007:3). Where certain linguistic variables frequently coincide with certain social variables, co-variation is established (cf. Kroch, 1978:17).

Sociolinguistics itself is not a one-dimensional field of study, but includes many facets and sub-theories such as, for example, politeness theory, which will receive attention in the next section. Perhaps the reason behind the complexity of the field lies within the complexity of the relationship between language and its ubiquitous socio-cultural context. Watkins states:

> "Language is linked to culture in a complex fashion; it is at once the expression of culture and a part of it." (Watkins, 1989:785)

When considering this seemingly straightforward statement more carefully, it becomes clear that it very concisely expresses the two key concepts of sociolinguistics, relating to the reciprocal nature of language and society. It reveals that language is (a) part of a culture and (b) an expression of it: thus (a) involves the way in which a language is dependent on its external social surroundings and (b) involves the way in which a certain society is dependent on language in order to express their cultural identity. A mutual influence is therefore manifest between language and society. In this section, therefore, the field of sociolinguistics will succinctly be clarified in terms of its definition (as foreshadowed in (a) – the study of

social influence on language), its development as a discipline and the above-mentioned underlying concept of (b), identity.

Hudson (1996:4) defines sociolinguistics as "the study of language in relation to society" and through this implies that this field is primarily interested in language and secondarily in society. This is different from the 'sociology of language', the field that studies "society in relation to language" and where society is consequently emphasised before language (1996:4). Naturally, this study is firstly interested in language and secondly in the society surrounding and influencing language, and hence sociolinguistics is a field of great value to this thesis. Indeed Hudson suggests that the value of the discipline lies in "the light which it throws on the nature of language in general, or on the characteristics of some particular language" (1996:4) and that the findings of sociolinguistics are "highly relevant to the theory of language structure" in relation to the nature of meaning and of grammar (1996:3). It can therefore be anticipated that the characteristics of SAfE, which may be manifest in the grammatical and semantic findings of the analyses to come, can be illuminated by studying the social contexts of the variety.

The development of the discipline of sociolinguistic study transpired along much the same lines as the field of historical linguistics, as described in § 2.2.1.1, as the interest in language and society emerged simultaneously with the increasing interest in language change and variation (as described earlier in terms of diachrony and synchrony) (Le Page, 1998:16). In its initial stages, studies interested in society and language arose from 19th century interests in dialectology (1998:17). Dialect forms were written down on a large-scale during the latter part of the 19th century, either in dictionaries (e.g. Joseph Wright's *English Dialect Dictionary,* 1896-1905) or on maps (Le Page, 1998:17). The early methodology of this field was somewhat underdeveloped, but nonetheless as empirical as possible (even if the data were not entirely natural) for an age without modern technical aids such as computers. This is evident in the example of Wenker, a German dialectologist, who "collected his data between 1876 and 1887 by asking 50 000 village schoolteachers to write down their local dialect equivalent of each of 40 sentences, set out in a questionnaire in High German" (Le Page, 1998:17).

During the mid 20th century, the field of contact linguistics arose, of which the primary concern was the influence of languages on other languages as they made contact (Le Page, 1998:19). This was essentially the 20th century foundation for

sociolinguistics. Le Page (1998:19) notes that and the term 'sociolinguistics' (as a discipline) is first referred to by Hodson (1939) and then later also by Martinet (1953), whereas the term 'sociolinguistic' (as an adjective) is first used by Eugene Nida (1949). Sociolinguistic concepts such as variation, dialect and contact were of course dismissed in the most famous formalist or generativist[25] works of Chomsky, e.g. *Syntactic Structures* (1957), but especially in his *Aspects of the Theory of Syntax* (1965), which was "predicated on the knowledge of that famous phantom, the idealized speaker-listener" (Le Page, 1998:19). Chomsky (1965:3) asserts:

> "Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogenous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance."

This view clearly leaves no space for variation of any kind – either linguistic or extra-linguistic.

Real momentum in the field was gained with William Labov's MA thesis, *The social motivation of a sound-change* (1963), which was focussed on Martha's Vineyard, and which emphasised social identity as a factor in language change (cf. Le Page, 1998:22). Labov's contribution to the field of language change has already been mentioned in § 2.2.1.1 of this chapter, but it is in sociolinguistics that he has made the biggest strides. Three years later, Labov published his PhD thesis, *The social stratification of English in New York City* (1966), where he stressed core sociolinguistic concepts like variation, prestige, social stratification and stereotypes (cf. Le Page, 1998:22).

These concepts relate to the social variables of language, which are later described by e.g. Milroy and Milroy (1998), denoting those elements influencing a language that are liable to vary in different social individuals or groups within various social environments, as also touched on earlier. Milroy and Milroy (1998:54) list the following two composite and other simplex variables: the first composite variable is socioeconomic (social) class, which includes income, trade, profession and educational level; the second complex variable, social network, is dependent on "indicators of density and multiplicity in a speaker's social relationships" (1998:54);

[25] Today referred to as the 'minimalist program' (cf. Radford, 2003).

other, simplex social variables include gender and age, as also listed in the beginning of this subsection.

Of these, social class is by far the most widely used and emphasised variable in sociolinguistic study, and incorporates the notion of prestige (an important concept for this thesis), as also stressed by Labov (1966), according to Milroy and Milroy (1998:55). As mentioned in Chapter 1, it is beyond the scope of this study to engage in micro-level investigations of all sociolinguistic variables, but the corpora are compiled as being representative of a balanced mixture of these variables to ensure comparable data sets. In the analyses and interpretations, however, some of the core sociolinguistic notions behind these variables may arise in discussion, e.g. that of prestige.

Apart from the social variables associated with a speaker, "much language variation is a result of differences in the social *situation* of use" (Stockwell, 2007:8 [emphasis original]). The situational variable is therefore another factor to consider in sociolinguistic studies, with the concept of register at the forefront. Stockwell (2007:9) writes that register is "defined ...by the circumstances and purpose of the communicative situation", and not "by the individual user or ethnic/social group using the variety". He adds that "the definition must [therefore] be a non-linguistic one, against which particular linguistic features can then be set" (2007:9). In other words, linguistic variation is systematic, and is constrained by extra-linguistic forces, such as the social and situational context. This strongly links with the later section on register studies (§ 2.2.2.1), which will explore the nature of register in more depth.

The discipline of sociolinguistics therefore encompasses extra-linguistic factors that impact on linguistic features, and the scope of its study can be said to reach as far as the subject of human society itself. Human society, in all its complexity, as well as each individual within a specific society, carries a sense of cultural and social identity[26]. Le Page (1998:19) indeed states that both the complexity of and the means to clarify the field of sociolinguistics lie therein that a "feeling of linguistic allegiance (will) largely determine the responses of every individual". Hudson (1996:11-12) also emphasises the element of identity when he gives insight into the relationship between the various facets of language and society:

[26] Schneider (2003) indeed suggests that identity, especially identity reconstruction, is the main condition for the formation of a new variety, which will be more thoroughly examined in the next section, along with the opposing view of Trudgill (2004).

> "...society is structured, form a sociolinguistic point of view, in terms of a multi-dimensional space. One need only think of the rather obvious ways in which people can be classified more or less independently according to the dimensions of age, region of origin, social class (or profession) and sex, to see an example of a four-dimensional space, each dimension of which is relevant to language. Each utterance of speech can be seen as an act of identity in a multi-dimensional space."

Elements like social identity and variation are key in understanding the discipline of sociolinguistics, but it is in the application of these elements in various sociolinguistic paradigms that the true relevance to investigating SAfE grammar can be found.

2.2.1.2.1 Frameworks for English varieties

The global spread of English during colonial times and the influence of that spread on the many Englishes of today are the general subjects of modern sociolinguistic conjecture. It has been a particularly debated issue where various Englishes, mainly arising from colonial spread, fit into linguistic models postulated by various sociolinguistic paradigms, which adhere to classifications of English users in different contexts. Three sociolinguistic frameworks for English varieties, as well as their three corresponding linguistic models will receive attention here. This will firstly include the World Englishes (WE) paradigm and Kachru's (1992) three-circle model, secondly Southern Hemisphere Englishes (SHEs) and Trudgill's (2004) model of new dialect formation and thirdly the paradigm of Post Colonial Englishes (PCEs) (Schneider, 2007) and, as a model drawing on various facets of this paradigm, Schneider's (2003) dynamic model of the evolution of New Englishes. In § 2.2.3 the relevance, workability and accuracy of their application to South African English will be discussed and evaluated.

The first framework applicable to this study is the paradigm of World Englishes (Kachru, 1985; 1992), which is concerned with the characteristics of and relationships between different English varieties across the globe. The three-circle model of Kachru (1992:356) is perhaps the most widely acclaimed and scrutinized model of Englishes around the world: it strives to fit three levels of English usage into categories represented by concentric circles (placing native, first-language (L1) speakers in the 'inner', second-language (L2) speakers in the 'outer' and English as a foreign-language (EFL) speakers in the 'expanding' circle). South Africa is not

mentioned in this model, and there is some debate as to whether SAfE fits into this model at all, or into all three circles at once – an issue that will receive more attention later in this chapter.

The second sociolinguistic framework to consider is the paradigm of the Southern Hemisphere Englishes (SHEs), which concerns itself with those Englishes used in areas within the southern hemisphere today. These locales include larger countries like New Zealand, Australia and of course South Africa, as well as smaller South-Atlantic islands such as St. Helena, the Falkland Islands and Tristan da Cunha. The key proponent of this framework is Peter Trudgill. His model of new dialect formation (2004) claims that it can predict the outcome of dialect mixture (a common trait of all SHEs) by considering information about the dialects and the initial contact situation. The key consideration is that Trudgill perceives this outcome is being mechanically determinable by input (2004:26).

In order for a transplanted, colonial English to become a new dialect, he proposes that six key processes (mixing, levelling, unmarking, interdialect development, reallocation and focussing) take place over three chronological or generational stages (rudimentary levelling and interdialect development; variability and apparent levelling; determinism in new-dialect formation). These processes and stages would apply to all SHEs because of their similar situations of dialect input and the chronological correlation between their developmental periods (18th or 19th century). Trudgill notes that the context from which all the SHEs emerged was of such a nature that English was transplanted to a territory where no previous population of native English speakers was present – a *tabula rasa* context (2004:26-7). This context will receive attention in § 2.3.2.1. Furthermore, the SHEs all start out as consisting of many various dialect mixtures, of which the features are levelled out within three generations due to subconscious mechanical accommodation, involving the survival of majority forms (Trudgill, 2004).

In describing these stages Trudgill only deals with what happens in the dialectal development of the original settlers, reinforcing the idea of a 'founder effect' as described by Mufwene (2001), which will receive specific attention in § 2.3.2.2 of this chapter. This approach is clearly suited to countries like Australia, as there is only one official language in the country – English – and the majority of its speakers are in fact remnants of the colonial population and hence first-language English speakers,

but the complex linguistic situation found in South Africa poses many challenges for Trudgill's model, which will receive attention in sections to come.

The third relevant sociolinguistic framework is the paradigm of the Post Colonial Englishes (PCEs). These Englishes include the SHEs, but expand their territory to the northern hemisphere in order to include all those English varieties around the world that developed from a colonial source. These supplementary locales include areas within North America, namely the USA and Canada, as well as the Indian sub-continent (India) and parts of East Africa (Tanzania and Kenya). For these, Schneider prefers the term *New Englishes* (2003:234-5).

On top of their different scopes in territory, Schneider's view of New English development is quite different from that of Trudgill. But, as Van Rooy (2010:6) notes, "the difference in opinion between Trudgill and Schneider relates less to the origin of ...linguistic features and more to the processes by which they spread". The main foci of Schneider's dynamic model of the evolution of New Englishes include the factors of identity and ecology (2003:239-240). The reconstruction of socio-cultural identity as expressed through language is especially emphasised in this model, and will be explored in subsequent paragraphs. In terms of ecology, this model draws attention to the different speech communities involved, including minority groups (settlers as well as indigenous people): this is expressed via the terms STL-strand (settler strand) and IDG-strand (indigenous strand) (2003:242). This manner of division contrasts with Trudgill's focus on the native speaker community and allows for a more comprehensive view. SAfE is in its essence an STL-strand, and will be treated as such in this thesis.

Schneider's focus on ecology and identity makes his model very different from the three-circle model of Kachru (1992), seeing that countries are not the defining units of categorisation in his model, as they are in Kachru's, but rather different strands of linguistic communities within and across national borders. Kachru's grouping into the inner, outer and expanding circles, does not, however, implicate that one circle is superior to another, seeing that he does not credit native speaker authority – an aspect which sparked a famous debate with Quirk (see Quirk, 1990; Kachru, 1991). Schneider (2003), like Kachru, challenges native speaker authority and proposes other factors that drive the evolution of varieties of English. Seeing that this thesis is interested in a native-speaker variety, and not in the influence

of a native-speaker variety on a second-language variety (but rather vice versa), I will not address this issue here.

The three prevalent parameters in the evolution of a new English, according to Schneider (2003:244), include sociopolitical background (together with identity construction), sociolinguistic conditions and linguistic consequences (structural changes). All of these parameters interact dynamically throughout Schneider's five stages of evolution, namely 'foundation', 'exonormative stabilization', 'nativization', 'endonormative stabilization' and 'differentiation' (2003:243). These parameters and stages are meant to be abstract and prototypical in nature, subject to variability, but the parameters and stages are nonetheless universally applicable to all New Englishes. Schneider's model is therefore meant to serve as a 'prototype' framework for prediction (2003:254), which leaves some space for variation at least, unlike Trudgill and his belief in mechanical processes. A brief exposition of Schneider's five stages will now be offered, with the particular goal of sketching the background against which SAfE will be evaluated according to this model. As stated above, the STL-strand is the most important for the purposes of the present study and will therefore receive most of the attention, and the IDG-strand will be described only where it is relevant to the settler strand.

The first phase, foundation, is initiated when a complex contact situation arises due to colonial annexation, which may be followed by settlement (2003:244). This contact situation may exist between both different languages (STL- and IDG-strands) and different dialects of the same language (group-internal dialects within a strand) (2003:244). Through this contact the process of koinéization or structural nativization, where speakers "mutually adjust their pronunciation to facilitate understanding" (2003:244), takes place within the STL-strand, while they still have limited contact with the other strand. Such limited contact will result in lexical borrowings into the English of the STL-strand from indigenous languages and marginal bilingualism arising within the IDG-strand (2003:245). The events of Schneider's Phase One overlap largely with Trudgill's (2004) focus on input and contact. As noted above, Schneider and Trudgill's models do not so much differ in terms of the origin as in the spread of linguistic features (cf. Van Rooy, 2010:6.) The spread of linguistic features, for Schneider, occurs after Phase One.

The second phase, 'exonormative stabilization', proceeds when political stability is gained and English becomes more widely spoken. The STL-strand tends to

retain its (British) identity at this stage and therefore still adheres to external norms from the home country. Yet, "concurrently with this consciousness of being representatives of British culture on foreign soil, adjustments to the local environment start to creep in and gradually modify the English being spoken in the new country" (Schneider, 2003:245). The STL-strand therefore "begins to move toward a local language form" and expands its vocabulary even more (Schneider, 2003:245-6). This slow 'movement' toward a local form is however still overshadowed by external norms.

The third phase is that of 'nativization', wherein the STL-strand usually gains political and therefore linguistic independence, as an identity shift consequently ensues (2003:247). The settler strand will by now have gained a sort of semi-autonomy by incorporating local culture into their own, thereby closing the gap between the STL- and IDG-strands to some extent. It is usual for a complaint tradition to surface during this phase, despite growing readiness to accept localised forms. New word formation can enrich the growing localised lexicon even more during this stage (2003:248).

After this has occurred, the fourth phase, namely 'endonormative stabilization' begins. Here the settler strand moves toward the positive adoption and general acceptance of the new localised norms, with a minority still remaining linguistically conservative (2003:249-250). This phase can only really be initiated and realised when a certain great political event, the so-called 'event X', has taken place; such an event will trigger comprehensive identity revision and the redefinition of position due to a sense of isolation from the home country (2003:250). The resulting new language norm will incorporate IDG-usage traits, causing the still prevailing gap between the strands to narrow yet. It is during this phase when a small number of grammatically deviant patterns might emerge in the newly formed English, when literature rooted in the new cultural identity begin to blossom and the New English receives a label to call it 'X English', e.g. South African English: an English with its own internal norms (2003:250-3).

The fifth and final phase of the dynamic model is that of differentiation, which includes internal diversification usually resulting in dialect birth within the New English itself (2003:253). During this stage group identity based on region and community gain more salience than a broad national identity. A counter-reaction may ensue during and after this process of differentiation as an attempt to reinstate

previous norms, as dialects become ethnic and social markers (2003:254), bringing with it the sociolinguistic aspects of pride and prestige, as well as stereotypes. In discussions to come, these five phases will be discussed in terms of SAfE.

Despite many smaller differences like terminology (e.g. formation vs. evolution), the main distinction between Trudgill and Schneider's models lies in the one element or factor that is the most prominent and universal in the formation or evolution of a new dialect or New English. The main issue that drives Schneider's model is identity, moreover, social identity:

> "Based on regional and social histories but also value orientation and customary modes of behaviour, a line is drawn between 'us' (those who share essential parts of that history and orientation, those we wish to socialize and be associated with) and 'others' (who don't share these qualities), and these attitudes and socialization patterns usually find symbolic expression – including (and perhaps most readily and rapidly) by means of linguistic variability." (Schneider, 2003:239)

This means that if a person feels affiliated with a certain community, he will express that feeling of belonging (social identity) through the way he speaks (accent), as to distinguish him from another social group that he regards as not associated with his identity. The five phases described above are all dependent on this element of identity.

Trudgill perceives matters differently. In his book *Investigations in Sociohistorical Linguistics: Stories of Colonisation and Contact*, Trudgill again emphasises that it is dialect contact and thus input which determines new dialect formation, and not identity (2010:192). Trudgill seeks to solve "linguistic mysteries" (2010:xiii) relating to English in e.g. the British Isles, North America, the Caribbean, and the Pacific Ocean (but not South Africa). These mysteries essentially include the reasons and processes behind dialect formation in these territories. In his prologue, Trudgill anticipates that the solution for these mysteries is, quite simply, contact. This can sometimes involve contact between different dialects of the same language and sometimes contact between different languages altogether (2010:xiii). In his conclusion, he indeed maintains this stance favouring deterministic principles, by which new dialects are always formed:

> "My suggestion is therefore that it is this innate tendency to behavioural coordination, not identity, that is the very powerful drive that makes dialect mixture an almost inevitable consequence of dialect contact, to an extent that factors connected with

> identity would not and could not. The actual linguistic characteristics of any new mixed dialect result form the relatively deterministic principles outlined in Trudgill (2004)...; and it is the new mixed dialect to which the founder principle ...then applies. In fact, all the evidence is that large-scale and prolonged dialect contact *always* leads to dialect mixture, and therefore in a sense requires no explanation, and certainly not one in terms of identity." (Trudgill, 2010:192)

Contrary to this, Schneider (2003:235) explicitly states that language contact situations "are not the prime determinants of the outcome of the process of new dialect emergence". Trudgill's notion of 'behavioural coordination' does however not link with identity. Behavioural coordination is related to the accommodation between speakers during face-to-face interaction (giving rise to dialect contact and mixture), which Trudgill does not regard as a social act, but as an innate, biological drive (Trudgill, 2008; cf. Van Rooy, 2010:5). In Trudgill's view, accommodation is the key force "responsible for the stabilisation of a new variety" (Van Rooy, 2010:5). Trudgill (2008:252) adds that accommodation is an inevitable, deterministic and mechanical process; new dialect formation therefore occurs automatically and subconsciously. Identity, according to Trudgill (2008) is rather a consequence of dialect convergence than the driving force thereof. He states:

> I ...claim that new mixed colonial varieties can and do come into being without identity factors having any involvement at all. We do not need this as an explanatory factor at any moment in human history. Of course, since the heyday of European colonialism, new identities most certainly have developed in most of the colonies. French Canadians are no longer French; Australians are certainly not British; Afrikaners are very definitely not Dutch; and these new identities do have a strong linguistic component. But my suggestion is that if a common identity is promoted through language, then this happens as a consequence of accommodation; it is not its driving force. Identity is not a powerful enough driving force to account for the emergence of new, mixed dialects by accommodation. It is parasitic upon accommodation, and is chronologically subsequent to it. Identity factors cannot lead to the development of new linguistic features... (Trudgill, 2008:251.)

Van Rooy (2010) however proposes that Trudgill conflates the ideas of feature 'innovation' and feature 'propagation' (also called 'diffusion') in his argument against the influence on identity (cf. Croft, 2000). Croft (2000:3-6) describes the innovation of features as the adding of a potentially new feature to the feature pool, whereas the propagation or diffusion of features involves changes in the probability of occurrence of features. Van Rooy (2010) points out that when Trudgill (2008: 244;250) argues that linguistic features are neither 'created' nor 'selected' due to identity, he, by implication, does not distinguish between the two processes, but treats

them "on the same plane" (Van Rooy, 2010:10). Schneider (2007:109-112), on the other hand, explicitly differentiates between innovation and propagation/diffusion, and allows for propagation to be driven by social factors like identity. Schneider (2007:110) does however not "strictly observe a contrast between innovation as linguistic-functional in nature, vs. diffusion as social in nature", but he does state that "external (social) forces play a bigger role than internal ones" (Van Rooy, 2010:10). Trudgill, on the other hand, rejects social forces in feature selection, as shown in his above quote (Trudgill, 2008:251).

To evaluate Trudgill's (2004) application of his model to New Zealand, especially in terms of accommodation, Baxter *et al.* (2009) implement his model mathematically and consider mechanisms of propagation from the evolutionary approach to language (e.g. Croft, 2000), which applies biological evolutionary concepts and terms to language (cf. Hull, 2001). Baxter *et al.*'s investigation shows that convergence and stabilisation did not occur within two generations, as Trudgill (2004) argues, which implies that other factors than accommodation must have contributed to the stabilisation of New Zealand English (cf. Kerswill, 2010). These may include certain social events that increased interaction among the population, but Baxter *et al.* (2009:291) note that their mathematical model "cannot decide what social factor is involved – or indeed if it is a social factor at all that determines the weighting of linguistic variants in language change". Nevertheless, they do believe "that the factors that are necessary are probably social" (2009:291) Ultimately, the results of Baxter *et al.*'s study show that "determinism cannot be a sufficient mechanism for the emergence of a new dialect" (2009:257).

More recently, Blythe and Croft (2012) set out to evaluate the mechanisms of propagation noted by Baxter *et al*. (2009). These include 'replicator selection' ("the differential weighting of the competing variants in a change" [Blythe & Croft, 2012:269]), 'neutral interactor selection' (the frequency of interaction in a social network structure), 'weighted interaction selection' (based on an influential social group propagating a feature) and 'neutral evolution' or 'drift' (change resulting from random fluctuations in a population) (Blythe & Croft, 2012:272-6). Blythe and Croft define each of these mechanisms mathematically "in such a way that the effects of different mechanisms in the trajectory of a change can be modeled" (2012:269). Their model not only suggests that language change occurs in the empirically observable pattern of an S-curve temporal trajectory, but also that this pattern can only be

captured "if the mechanisms for propagation include replicator selection" (2012:269). Blythe and Croft describe replicator selection by means of an experiment:

> "Suppose two variants, A and B, are present with some prescribed distribution of usage frequencies across members of a speech community at some initial time. After some number of interactions between speakers, these frequencies will change, and there will be some probability distribution of usage frequencies of the two variants at that time. Now consider a parallel speech community, identical in every way to the first, other than that the initial frequencies of the two variants are reversed. We can now compare the distributions of usage frequencies at the same point in time between the two communities. If after exchanging the usage frequencies of the two variants in the first, there is the same probability distributional of usage frequencies in the second community, then there is no differential valuation of one variant over the other. The variants are replicated with the same frequency when used in the same social and linguistic context: that is, they are symmetrical." (Blythe and Croft, 2012:272)

Blythe and Croft (2012:273) propose that the "replicator selection mechanisms for propagation" may be directly associated with social values (cf. Labov, 2001), but, like Baxter *et al*. (2009) note that their "mathematical model cannot decide among [the] various social, functional, structural, or phonetic weightings of variants, or indeed any other differential weighting of variants", which may include prestige (cf. Labov, 2001:24; Trudgill, 2004:153), or identity (cf. Schneider, 2003). Blythe and Croft however state that all of these factors "are instances of replicator selection". Therefore, Blythe and Croft (2012:294-295) find that replicator selection is the mechanism that determines the overall trajectory (S-curve pattern) of propagation, but that the "microstructure of the process – how it is diffused through subparts of linguistic and social context – may be driven by social network structure and weighting of the language use of different subgroups of the population".

To summarise, the main difference between Trudgill and Schneider's views lies in the processes by which new features spread or propagate (see Van Rooy, 2010:6): Trudgill (2004; 2008; 2010) argues that the influence of input, dialect contact and mixture, as well as accommodation will mechanically determine the outcome of a new variety in a *tabula rasa* situation. Although he does not mention it overtly, Trudgill's model is also dynamic, like Schneider's, seeing that its processes are divided into three chronological stages (cf. Van Rooy, 2010:5). However, Trudgill does not consider social process within these three stages. On the other hand, Schneider argues for the pivotal role of social forces in the evolution of a New English: these are most notably identity reconstruction and the ecology of the STL-

and IDG-strands. In considering Trudgill and Schneider's views regarding variability in World Englishes, Van Rooy (2010) suggests that both views may be somewhat incomplete in the factors they consider, but generally accepts Schneider's model. Van Rooy rightly states that Trudgill's view lacks consideration of social forces and that Schneider's model lacks some emphasis on input to language contact settings (2010:3;16). I agree with this view that Schneider might not consider input to be important enough, but I do accept his emphasis social factors, above those of Trudgill and his deterministic approach independent of many social factors like identity.

These two opposing models and their primary factors in the formation or evolution of a new form of English will be evaluated in terms of their importance and applicability to SAfE in § 2.3.3.3. The three paradigms, and especially the latter two, described in this section are the cornerstones of the sociolinguistic portion of this study, but a certain sociolinguistic sub-theory also requires attention to complete the sociolinguistic discussion.

2.2.1.2.2 Politeness Theory

Certain notions about politeness arise from e.g. Bowerman's (2004b:477) claim that *must* "has much less social impact in WSAE than in other varieties of English and often substitutes for polite should/shall", as mentioned in Chapter 1, and therefore politeness theory needs some clarification. In its core, politeness theory, as put forward by Brown and Levinson (1987), rests on the contextual (socio-cultural) aspect of language, which is why it can be seen as a sub-theory of pragmatics, which can in turn be considered a sub-theory of sociolinguistics (the combination of these fields is often called 'sociopragmatics') (cf. Nevalainen & Raumolin-Brunberg, 2012). Brown and Levinson (1987:55) perceive politeness as a phenomenon of a special sort, as it is both universal and culturally entrenched at the same time. They emphasise "the extraordinary parallelism in the linguistic minutiae of the utterances with which persons choose to express themselves in quite unrelated languages and cultures" (1987:55). They furthermore believe firmly that message construction (language usage) is "the very stuff ...social relationships are made of" – in other words, it is a crucial part of "the expressions of social relations" (1987:55).

Indeed Brown and Levinson (1987:280) suggest that the main concern in sociolinguistics lies in the origin and quality of social valence attached to linguistic form, and hence believe that the discovery of the principles of language usage would be concurrent with the discovery of the principles of interactional social relationships (1987:55). This is similar to Hudson's (1996:3-4) description of sociolinguistic ideas and concerns in the above section on sociolinguistics. Brown and Levinson further state that the most profound interrelations between society and language can be found in action and interaction (1987:280). These interactional social relationships involve special, shared elements, according to politeness theory. They distinguish between qualities universal to all language users, namely those elements of 'rationality' and 'face' (1987:55;58). Face is the more relevant element to this thesis: it involves the want to be unimpeded by others (negative face) and the want to be approved of or seen as desirable in certain respects (positive face) (1987:58-9;62).

The term 'face', as Brown and Levinson use it, is derived from that of Goffman (1967) and from the English folk expression of 'losing face', which relates to "notions of being embarrassed or humiliated" (Brown and Levinson, 1987:61). They define 'face' as "the public self-image that every [adult] member [of a society] wants to claim for himself" (1987:61). Although the phenomenon of face and the social need to be sensitive to it in interaction are universal, it is the content of face that is distinctive in different cultures, in terms of "what the exact limits are to personal territories" and "what the publicly relevant content of personality consists in" (1987:61-2). The link between the universal and cultural nature of face is perhaps emotion:

> "...face is something that is emotionally invested, and that can be lost, maintained, or enhanced, and must be constantly attended to in interaction. In general, people cooperate (and assume each other's cooperation) in maintaining face in interaction, such cooperation being based on the mutual vulnerability of face." (Brown and Levinson, 1987:61).

This mutual vulnerability lies therein that one's 'face' is expected to be defended if threatened (1987:61), which leads to the next important concept in politeness theory, namely the face-threatening act.

The term 'face-threatening act' is self-explanatory, as it is an act (speech act) in verbal or non-verbal communication that by its nature "run[s] contrary to the face wants of the addressee" (1987:65). Such a face-threatening act (henceforth FTA), is

an intrinsic threat to either the addressee's negative or positive face, but it is the FTA that targets the negative face of the receiver that is more relevant to this study. As mentioned above, negative face is embodied in the want of an adult member of society to be unimpeded: this includes the basic claim to territories, personal preserves and rights to non-distraction – in short, it is the desire for "freedom of action and freedom from imposition" (1987:61).

Brown and Levinson (1987:65-7) generally assign three categories of FTAs to negative face and two to positive face. In anticipation of the semantic interpretations of this thesis regarding the strength of obligation (cf. § 4.3), one particular category of threats to this negative-face want is the most relevant. The relevant category includes those acts from the speaker or sender (S) that predicate some future act (A) of the addressee or receiver (H) (in Brown and Levinson's terms (1987:65)). S therefore pressurises or obliges H to do or to refrain from doing A. These include:

(a) orders and requests (S indicates that he wants H to do, or refrain from doing, ...A)

(b) suggestions, advice (S indicates that he thinks H ought to (perhaps) do ...A)

(c) remindings (S indicates that H should remember to do ...A)

(d) threats, warnings, dares (S indicates sanctions against H unless he does A)

Essentially, this particular category of Brown and Levinson contains four different levels of threats to negative face, in terms of the degrees in which they threaten the face of H: here (d) is the most threatening, (a) the second most threatening, (c) is the middle ground and (b) is the least threatening. It worth noting Brown and Levinson's use of modals in this list, e.g. *should* in (c) and *ought* in (b), which correlates with the traditional view that these modals have a lower degree of social force (cf. § 2.4.1.1).

In order to reduce impact when performing FTAs out of necessity, the speaker can follow a number of strategies, as Brown and Levinson describe (1987:68-71). One of these strategies, and the one most relevant to this thesis, is that of negative politeness, which is "orientated toward partially satisfying (redressing) H's negative face" (1987:70). This strategy is based firstly on avoidance and secondly on the recognition of the addressee's negative-face wants. Hence, the characteristics of negative politeness are self-effacement, formality and restraint, which can be done by offering apologies, by hedges on illocutionary force, with impersonalising mechanisms (e.g. passives) and with other softening mechanisms (1987:70).

Illocutionary force or obligative force, as manifested in modals, will receive more attention in later sections of this chapter and the next.

As there is a natural tension in negative politeness between going on record to 'pay face' and going off record to avoid imposing, the compromise is reached in conventional indirectness, as can be illustrated by a conventionalised phrase such as *can you pass the salt* (which is perceived as a request, with no possibilities left for a viable alternative interpretation, e.g. a question about the addressee's potential abilities) (1987:70;133). By using the strategy of conventional indirectness, the speaker pays "a token bow to the negative-face wants of the addressee", which shows that S "is aware of and honours the negative-face wants of H" (1987:71). Such conventionalised, indirect speech acts are regarded as universal phenomena by Brown and Levinson (1987:132).

Despite the universal character of the above, the degree of seriousness of an FTA is culturally determined, whereas the assessment of the seriousness of an FTA, however, again involves universal factors for consideration. These factors, which has an empirical basis according to the authors, include:

(i) the 'social distance' (D) of S and H (a symmetric relation)
(ii) the relative 'power (P) of S and H (an asymmetric relation)
(iii) the absolute ranking (R) of impositions in the particular culture.

Based on the above, Brown and Levinson (1987:76) suggest that the weightiness (W) of an FTA can be calculated as a numerical value by using the following formula:

$$W_x = D(S,H) + P(H,S) + R_x$$

This formula entails that W_x (the numerical value of the weightiness of an FTA) is determined by the combination of three dimensions or factors, as quoted above (1987:76). Brown and Levinson elaborate on this to some extent, but this thesis does not necessarily wish to calculate scores or values of FTAs later in its analyses. This formula should merely illustrate the fact that "all three dimensions P, D and R contribute to the seriousness of an FTA, and thus to a determination of the level of politeness with which (...) an FTA will be communicated" (1987:76).

All communicative intentions indeed have underlying social implications and therefore may often be of a threatening sort, as is embodied in the FTA. It is the FTA, as well as the constraints of communicative intention, expressed by means of "the pragmatic resources of a language in the form of constructed messages", that are the core of politeness theory (1987:281). It is therefore the universal phenomenon of politeness that helps to add constraint to face-threatening acts – constructed and strategised differently in diverse cultures – which will aid the social relationship of the speaker and addressee. In conclusion, Brown and Levinson (1987:281) argue:

> "[T]he social valence of linguistic form has two especially important sources: the intrinsic potential impact that a specific communicative intention may have on a social relationship, and the ways in which by modifying the expression of that intention participants seek to modify that impact – such modification measuring for participants the nature of the social relationship. (...) In short, language usages are tied to strategies rather than directly to relationships, although relationships will be characterized by the continued use of certain strategies."

Politeness theory is therefore part of the bigger picture sketched in sociolinguistics. The social aspect of language, its social context, also ties in with the historical aspect thereof, as was seen in previous sections – the result of which is the main theoretical framework of this thesis, namely sociohistorical linguistics.

Section 2.2.1 has explored this theoretical framework in terms of the development of the fields of historical linguistics and sociolinguistics, as well as the main ideas that form their foundation. The discussion on historical linguistics considered issues involving diachronic and synchronic description, historical texts as a means of investigation and the role of external (e.g. cultural) factors in language change (linking with the field of sociolinguistics), as well as the nature of the grammaticalisation process as the driving force behind language evolution. The section on sociolinguistics discussed the various social factors that influence language variation in a society, including identity, and the three main frameworks for English varieties as postulated by respectively Kachru (1992), Schneider (2003) and Trudgill (2004), as well as politeness theory. The theoretical framework and its constituents being clarified, the next step is to discuss the main empirical framework adopted in this study to perform analyses, ensure credible results and aid reliable interpretations based on sociohistorical theory.

2.2.2 Main empirical approach: corpus linguistics

The general empirical approach of this study is that of corpus linguistics, as detailed by e.g. Biber *et al*. (1998). Corpus linguistics involves the method of using corpora, which consist of real linguistic data (texts of spoken or written language), as the basis for an evidence-based strategy to investigate language phenomena, i.e. it is essentially rooted in the usage-based approach to language. Biber *et al*. (1998:1) state the following about corpus linguistics and corpus linguists:

> "From this perspective, we can investigate how speakers and writers exploit the resources of their language. Rather than looking at what is theoretically possible in a language, we study the actual language used in naturally occurring texts."

The corpus-based approach to linguistic analysis holds the following four characteristics, according to Biber *et al.* (1998:4): it is empirical, it utilises a corpus as the basis of analysis, makes use of computers for analysis and depends on quantitative and qualitative techniques.

The empirical nature of corpus linguistics is underpinned by the corpus itself, which is, ideally, a direct source of natural language, as actually used by speakers and writers of a language. A corpus, which is defined as a "large and principled collection of natural texts" (1998:12) can be either synchronic or diachronic in nature, depending on the chronological range of the text production within the corpus. It is therefore a dynamic tool for the linguistic researcher, as a publicly available corpus can be chosen or a new one could be collected to suit particular research aims, e.g. sociolinguistic variables (e.g. a corpus of texts produced by dialect speakers, a certain age group, or gender) and historical linguistic variables (e.g. texts from different periods in history, etc.).

The selection of data can be manipulated to meet different needs, so long as the data remain in their authentic state, untampered and therefore natural. This does not apply to the original written documents or spoken language, which would be essentially useless in its physical, scattered and sometimes fleeting form. The data need to be digitalised before they can be properly analysed (e.g. electronic written texts or transcriptions made from audio recordings of speech), and it is in this process that the natural linguistic character of the texts should not be tampered with in order

to ensure its empirical value. Various techniques aid in achieving this and these will be discussed in Chapter 3.

Modern technological aids such as computers, scanners, transcription tools, text-converting software, text-analysis or parsing software, electronic corpora, online-corpora, etc. facilitate and optimise the accessibility, speed and accuracy with which the corpus-based approach can aid in the gathering, storage, preparation, filtering and analysis of corpora. The complex analysis of corpora is particularly promoted by these electronic analysis tools, because such tools provide "consistent, reliable analyses" (1998:4) – much more than the human eye and hand can. However, the human analyst remains an essential interactive companion to the machine where it comes to difficult linguistic judgements, according to Biber *et al.* (1998:4).

As soon as the corpus is ready, a search for linguistic features within the corpus can be embarked on and e.g. concordances or word lists can be made via analysis tools. The raw frequency results then need to be normalised if they include counts from texts of different lengths, which entails adjustment of the counts for accurate comparison (Biber *et al.*, 1998:263). These normalised frequencies can now serve as a statistical data set ready for further interpretation; ergo, the quantitative analysis can be meaningful on its own. It is, however, the deeper, qualitative analysis that follows, that will explore the importance of the statistical findings and their significance to the research topic. By means of extracting samples from the concordances, close explanation, exemplification and interpretation can be done, which is, according to Biber *et al.* (1998:5) an imperative part of the study, in order to learn about the patterns of language use. To summarise, both quantitative and qualitative techniques of analysis should be employed by the analyser. Both are essential in sketching a complete, functional picture, as straightforward numeric counts of the occurrence of linguistic features, although always essential in corpus research and meaningful in itself in simpler cases, can however result in one-dimensional research in the case of complex association patterns if not used in conjunction with statistical and interpretational analyses and evaluations (1998:8-9). Only after this step, grounded conclusions about linguistic patterns can be made and the corpus-based study has reached its goal.

The corpus-based approach and its techniques therefore play a valuable and indeed essential role in modern linguistic analysis, where the full potential of the user-friendliness and accuracy thereof should be exploited to the researcher's benefit. This

approach to investigation does however not come without challenge or effort. For example, some of the challenges of using texts as investigative basis have been noted in § 2.2.1.1. Chapter 3 will explain this thesis's application and use of this empirical framework and its methodology, as well as its challenges. Despite its challenges, corpus linguistics has become increasingly common and appreciated; corpora have become more accessible and the empirical results sprouting from corpus-based investigations have triggered many new research questions (1998:ix).

Many modern and major lexicogrammatical descriptions of the English language rely on corpus-based principles and methodology, such as the *Concise Oxford English Dictionary* (2006), of which the lexicon is 'powered' by the Oxford corpus. On the other hand Biber *et al.* (1999) employ a corpus-based approach in order to describe English grammar across spoken and written English. This thesis is only interested in describing one constituent of the grammar of SAfE (modality), and only in its written form. Written texts will therefore constitute my corpus, as is also described in Chapter 3. Within the written register, however, many sub-registers can be distinguished, each with different characteristics and intentions. Hence, the branch of corpus linguistics interested in register variation, i.e. register studies, will be discussed next.

2.2.2.1 Register studies

Within the field of corpus linguistics, the sub-discipline of register studies, as explored by e.g. Biber (1988, 1992) and Biber *et al.* (1998), focuses on variation in different textual genres. In essence, register studies are concerned with language variation, the latter of which has received some discussion above in terms of e.g. Pyles (1965) and Southworth (1990), as well as in the earlier section on sociolinguistics (§ 2.2.1.2). Variation in language is highly systematic, according to Reppen *et al.* (2002:vii), where the speakers of a language have to make choices which depend on non-linguistic factors, including the purpose of the communication, the relationship between the speaker and hearer (interlocutors), the circumstances of the communication, etc. This strongly links with the sociolinguistic concept of interaction between extra-linguistic (social and situational) and linguistic forces (see §

2.2.1.2). The term *register* hence denotes "(a) set of features of speech or writing characteristic of a particular type of linguistic activity or a particular group when engaging in it", according to Matthews (2007). This firstly implies the division of language into the macro-registers of spoken and written language, and secondly implicates a situational or circumstantial element. This suggests that the language of a particular group can be characterised in terms of a particular register that is appropriately formal or informal according to the relevant circumstances.

Linking with this definition, Biber (1995:1) suggests register to be a "cover term for any variety associated with particular situational contexts or purposes". Biber (1995:1) furthermore states that although register distinctions are usually defined in non-linguistic terms (linking with Reppen *et al.*'s (2002) statement above), there are also prominent textual differences between registers; they can be named varieties of texts within a culture (including e.g. letters, novels, sermons), and can also be defined according to level of generality (where e.g. academic prose is a general register and methodology sections in psychology articles are a very specified register). This suggests that, on top of the macro-registers of speech and writing, there are also many micro-registers, for example, conversation, public oratories, letters, novels (fiction) and academic papers, which form part of a bigger group.

Most studies on register variation, like those of Biber (1988, 1992, 1995) for example, adopt a stylistic focus, investigating the differences between written and spoken language. Furthermore, such studies tend to follow the methods of a multi-dimensional analysis (as postulated by Biber, 1992, 1995) whereby the textual dimensions of each register (in terms of e.g. information density) are investigated and compared to other registers (e.g. Biber, 1995:141).

Geisler's study (2002:268) is one example that adopts this multi-dimensional methodology to investigate register variation in nineteenth-century English, for which he found that those registers that share situational characteristics tend to be clearly defined and are becoming more specialised or distinct. On the other hand, Biber and Finegan's (1989) multi-dimensional study on the historical development of registers from the 17th to the 20th century, which also employs the multi-dimensional analysis, is concerned with the general written register and, moreover, the evolution of the specific registers of fiction, essays and letters. As part of their findings, Biber and Finegan (1989:487;516) report a trend or a 'drift' toward the increasing use of oral, more informal and colloquial styles in written English over the last three centuries.

This linguistic phenomenon, known as colloquialisation, is also posited for the 20th century by Mair (2006) in general terms, as well as by Westin (2002:163) regarding the growing informality of newspaper editorials. Colloquialisation will receive more attention in § 2.4.2.1, which will specifically consider the phenomenon in terms of diachronic trends in the modal and quasi-modal auxiliary groups.

The main focus of this section is on those written registers relevant to this thesis, such as letters, fiction, non-fiction and news. They require description in terms of their textual and contextual characteristics. The first register to receive attention in this regard is letters, whereafter fiction and non-fiction will be discussed, before moving on to the final register of news. Each of these registers will be explored in terms of their form and function.

The practice of writing letters is indeed an enduring method of written correspondence since ancient times and can vary from very formal to very informal letters, depending on various factors (Barton & Hall, 2000:1). Contemporary letter writing is rapidly moving away from the almost bygone era of the pen and paper, to be replaced by electronic mail, and, on the more informal end of the continuum, social-networking correspondence via all sorts of electronic devices (including computers, tablets, mobile phones, etc.) and networks. Up until the late 1990s, however, these kinds of 'new' electronic modes of correspondence were not very common in South Africa, and the letter remained an integral part of everyday life. It is therefore clear that the letter text has been a very important register throughout those periods this thesis is interested in: from the early 19th century to the late 20th century.

Letter writing during the nineteenth century was indeed a very formal, and socially very significant affair filled with convention, ceremony and ritual. During this era the act of writing letters fulfilled three main purposes, the first being that it was a household chore for women to read and compose letters of personal and household correspondence on a daily basis (Barton & Hall, 2000:2). The second purpose is that it was a social practice and the third purpose is that letter writing was positively an art form during this period, where the writing tools, materials, media and methods matched that of an artist (Hall, 2000:91). Apart from these general practical purposes described above, letters have many various textual purposes as well, and assume different forms. Barton and Hall (2000:1) describe the nature of letters as follows:

> "A letter can be used to mediate a huge range of human interactions; through letters one can narrate experiences, dispute points, describe situations, offer explanations, give instructions and so on. Letter writing occurs in many forms, letters, postcards, memos, electronic mail, dialogue journals, fax, etc."

The letter is therefore a flexible and multi-faceted register, serving many different purposes and adapting to different contexts. Indeed the flexibility of this form of social or official communication has, "historically, allowed other genres to emerge from letter writing" (Barton & Hall, 2000:1), rendering it not only a strong basis of communication in itself (see also Pahta *et al*., 2010:8), but also a platform from which other genres have emerged and are emerging, like today's ever-growing range of electronic correspondence. When describing the register of letters in more detail, two broad kinds of forms (conventions) and matching functions are perceptible.

Firstly, as described by Barton and Hall (2000:2-3), the social letter is a personal form of correspondence, tending to be more informal and hence less ritualistic, which some exceptions as to opening and closing, especially during earlier periods. These letters have a propensity to centre on matters of family or local news, adopting a narrative style in some cases, and therefore tend to, contain requests for favours or express feelings between familiar people such as friends or family (2000:2-3). In the second place, Barton and Hall (2000:2-3) describe the business letter as being an official form of correspondence, which usually includes formal address and is highly conventionalised in form. This letter usually serves as correspondence between people who are unfamiliar with each other and may include letters of information or request to e.g. schools, companies, politicians or even the local newspaper (2000:2-3). In addition, an administrative act such as 'paying the bills' is seen as being an extension of the act of business letter writing (2000:3).

Pahta *et al*. (2010:8) adopt a similar view, but add that a layer of formality may also be added to personal letters in certain cases. They maintain the following:

> "Linguistic variation within the letter genre, for example, is evident as business correspondence in general tends to be more formal than private correspondence, but private correspondence may also be formal particularly if the correspondents are socially distant and/or unequal in status."

This adds a layer of social roles to the act of letter writing, which links with the concepts of social identity between correspondents. As will be seen later, this thesis includes both social and business letters in its data for analysis, so these above

concepts and conventions will be taken into account where letters are concerned.

Two related textual genres or registers that have also been highly conventionalised in terms of form and function are fiction and non-fiction. However, these two registers share one trait that letters do not: they are usually published forms of writing, in which the editing process by an independent editor or editors play an integral role. The next few paragraphs will relate the forms and functions of fiction and non-fiction, as well as describe their editorial processes very tersely.

The primary function of fictional or creative writing, including novels, novellas, short stories or sketches, is to entertain by means of 'reading for pleasure', whereas non-fiction, which includes academic writing, popular non-fictional writing about special skills and hobbies, autobiographies and biographies, pamphlets, textbooks, etc., tend to serve an informative, explanatory or argumentative function (cf. Biber *et al*., 1999:20-23). Of course these functions are reflected in the various forms these registers can assume.

The two main forms of language conventionally found in fiction are narration, as well as speech or thought representation (Banfield, 1982). Narration can be defined linguistically "with reference to the behaviour of person and tense", according to Banfield (1982:142), in that, firstly, it is the primary means of the author to relate the story, conventionally, via either a third- or first-person narrator, and secondly, it is usually written in the past tense (Banfield, 1982:142; 169-170). As briefly mentioned earlier, social letters can also include narration, which means that it is not confined to fiction[27]. Furthermore, narration can form part of other text types, such as reminiscences, first-hand accounts of events, etc. Terblanche (2011), in a corpus-based study on the features of the narrative in East African English, found many more such features on top of these (core) features noted above. Most relevant to this thesis, Terblanche (2011:36;39;61;83) deems modals and semi/quasi-modals peripheral features of the narrative (but not part of the core features), and places these auxiliaries in the evaluative[28] feature category together with the emotional stance verb *feel*,

[27] As will be seen in Chapter 3, other narrative texts were also included as part of the data in this register, especially in the 19th century, for which fiction is not always easily available.

[28] Other feature categories of the narrative, apart from the evaluative, include agency (core: first and third person pronouns, proper nouns for persons, activity verbs, e.g. visit, etc.; peripheral: second person pronouns), causation (core: non-finite causative clauses, e.g. *verb (e.g. get/afford) + to-clause*; peripheral: causative verbs, e.g. *make/get*, causative preposition verbs, e.g. *lead to/call for*, causative subordinator *because*) and contextualisation (core: past tense, perfect aspect, time and place adverbials, e.g. *afterwards/downstairs*; peripheral: present tense) (Terblanche, 2011:29-39; cf. Biber *et al*., 1999;

evaluative adjectives, e.g. *nice* (core features) and epistemic stance adverbials, e.g. *maybe/sort of* (peripheral feature) (cf. Biber *et al*., 1999).

I of course do not wish to calculate narrativity scores for e.g. letters and fiction in this thesis based on feature counts, but the simple fact that these registers contain narration is significant to consider, as this fact could aid in the explanation of some frequency patterns in the modals and quasi-modals I analyse (see Chapter 4). The narrative text itself differs from speech and thought representation firstly in that the latter form is a direct presentation of a character's words or thoughts, without the interference of a narrator, except for reported speech surrounding the quotation (e.g. 'she said'), and secondly in that it is usually written in the present tense (Banfield, 1982:141).

Non-fiction, on the other hand, usually excludes direct speech or thought representation, except perhaps on occasion in texts like biographies, but, like fiction, can also include long episodes of narration (albeit of factual rather than fictional events). Fontaine and Glavin (1987:170) argue that non-fiction does not just aim to inform, explain or illuminate a set of facts or a situation, but that it can also include narration in order to help lessen the distance between the reader and those facts or situations. However, non-fiction is a broad register with many potential forms and functions, from where a general intent to inform can be established, which can not only adopt a narrative tone, but also an argumentative (rhetorical) tone to persuade the reader of something (Russell, 2002). One of the most prominent forms of non-fiction is academic prose, but this register, which harbours both argumentative and informational writing, can be seen as a relatively new one, when considering the mostly oral nature of such language acts before the nineteenth century.

Russell (2002:35) suggests that writing instruction beyond that of elementary school was deemed unnecessary until about the last third of the nineteenth century, because writing was seen as being ancillary to speaking. It is indeed difficult to imagine, from a modern print-culture perspective, "a time when education, indeed public life, was so dominantly oral" (2002:35). This however began to change by the 1870s[29], when a subtle but profound shift in the role of writing began to manifest. This shift mainly occurred due to the host of new professions arising during this

---

my emphasis).

[29] This coincides with the start of Period 1 (1870s-1900s) of the historical corpus used in this thesis.

period, whose members communicated through texts, which were never meant to be mere substitutes for oral communication, but to represent something entirely new, including reports, memoranda, specifications, scholarly articles, etc. (Russell, 2002:4).

Russell describes student writing (an unpublished form of academic writing) in America as a notoriously problematic new register during this period (2002:7). As university students were required to produce academic essays, late nineteenth-century educators were increasingly concerned with ways to solve the 'problem' of students' bad writing skills – eventually settling on a course for freshmen, which acted as the successor to a very different rhetoric course in the old curriculum (2002:7). Today, with the main purpose of academic writing being to argue, persuade, inform and discuss the results and implications of research, the language of argument (although toned down since the rhetoric of previous eras) is used to convey these ideas, for example, the present tense is mainly used to show the universal truth of a proposition (cf. Biber *et al*., 1999:456).

To describe the forms and functions of fiction and non-fiction without considering the role of editing would be to miss an important part of the nature of these registers, seeing that both kinds of texts have to pass though sometimes-rigorous editorial processes before their publication. According to Gunther & Sharpe (1994:11; 15) the chief editor for a work of fiction or non-fiction is not primarily concerned with an author's language use, but rather seeks holistically to evaluate content and the presentation of the text, whereas the line editor concerns himself with the actual manuscript – evaluating the issues of consistency in tone, dialogue representation (in fiction), style, logical development of ideas (in non-fiction), etc. The language use of these two registers can therefore be anticipated to be much more 'polished' and to contain fewer 'errors' (in a conservative, formal sense) than the unpublished register of letters. Apart from fiction and non-fiction, news is also dependent on editing, which will be discussed next.

Apart from the editorial process, Conboy (2010:1) maintains that newspapers have always been subject to other external factors, seeing that they had to "accommodate social, political and technological changes throughout their history". He further states:

> "Despite their underlying commercial imperative, this need to provide a distinctive language in which to give a coherent editorial expression to readers' tastes has had both conservative and radical implications at different moments in the history of the newspaper." (Conboy, 2010:1)

The genre of news writing in its original state was simply the narrative report, but it has since evolved beyond this form due to a specific set of socio-historical processes (Conboy, 2010:14). Conboy (2002:15) relates that by the early sixteenth century, more than a century after William Caxton had introduced his Westminster Press in 1476, the first (narrative) news pamphlets started to appear, outlining the progress of the English king and his army travelling north, their battle strategies, etc., and that the social and political impact of print was only starting to be appreciated during the accession of Henry VIII.

In modern times, a very different style of news writing is emerging, seeing that newspapers in the contemporary world experience more rapid responses to their articles, as well as more enthusiastic interactions with readers' views and opinions (Conboy, 2002:136). Although this is an unavoidable part of the newspaper's evolution, especially with the rise of electronic versions of newspapers with space for comment and discussion, a "casualty of these processes has been the prime function of 'news'", which is increasingly becoming replaced by e.g. a range of views and lifestyle commentary (Conboy, 2002:136). Another aspect of the newspaper which makes it a unique register is the impact of editing style, as mentioned previously, which may differ from newspaper to newspaper, and causes the register to become a stylistically diverse one due to variance in in-house style, although keeping within the functional objective to inform, relate and communicate (2002:137). All these processes, which involve engaging with the reconfiguring social setting, have continued to shape the language of newspaper (2002:136).

In summary, this subsection has described letters, fiction, non-fiction and news in terms of their forms and functions. It shows that, apart form the concept of register involving notions of text, it also rests on notions of context, including especially social factors. Such social factors are indeed an integral part of not only the process of defining and describing registers, but also when considering language in general, especially the history of English and the impact of such factors on language change.

Section 2.2 has discussed and evaluated the composite theoretical framework

of sociohistorical linguistics (including the fields of historical linguistics and sociolinguistics) and the empirical approach (corpus linguistics), which together form the basis of this thesis, and has shown how the various notions of theory and method intersect. In this thesis, the empirical approach, which underpins the methodology of the study, is informed by the theoretical framework as a means of explanation, and the theoretical framework likewise relies on the empirical approach as a means of investigation. The next section will discuss the impact of social history on language change. In other words, the theoretical framework catalogued above will be applied to specific social factors involved in language development, as well as to the South African context.

2.3 SOCIAL HISTORY AND LANGUAGE CHANGE

In keeping with the general theoretical framework of this thesis, sociohistorical linguistics, this section aims to explore the general sociohistorical reasons behind language change. I will therefore cast light on the various historical circumstances that have driven the evolution of language, and indeed English, across the ages, by impacting on various social factors, and ultimately lead language on a path of change. History and society interact on various levels and will therefore be described in relation to each other in this section. This section will ultimately show that migration, language contact and social identity are the most prominent sociohistorical factors of language change and especially for English, where language contact serves as the critical link between the other factors.

2.3.1 General reasons for language change

When surveying the history of language itself – investigating how it first came into being – one can immediately notice that social factors have always been at the forefront of motivation for its development in history. Croft (2000), Deutscher (2005) and Tomasello (2008) argue for an evolutionary perspective on language genesis and language change, the former of which broadly hypothesises that language has

developed progressively (by means of selection and adaption) since its humble beginnings of signing and the making of sounds, due to a need for human communication in every society. Tomasello (2008:2) summarises his argument as follows:

> "[The] evolutionary hypothesis [is] that the first uniquely human forms of communication were pointing and pantomiming. The social-cognitive and social-motivational infrastructure that enabled these new forms of communication then acted as a kind of psychological platform on which the various systems of conventional linguistic communication (all 6,000 of them) could be built."

Here Tomasello refers to the "social-cognitive and social-motivational infrastructure" that underpins the development of language. The role of society can therefore be deduced to be a central aspect of human communication, impacting cognition and internal motivation to not only develop a language, but also to keep developing it (in the broader sense of ongoing language change). These macro-factors, that have shaped and continue to shape language today, are expounded on in a series of volumes by Labov, where he explores the internal (Labov, 1994), social (Labov, 2001) and cognitive and cultural (Labov, 2010) factors or principles of linguistic change.

In his most recent volume, exploring cognitive and cultural factors, Labov (2010:4-5) argues that a need for social cohesion and a sense of community (even if this occurs below the level of consciousness), as well as the separation of communities, contact with other communities and the influence of old and new social identity, are the main external reasons for language change (in terms of divergence and convergence). He further asserts that change is usually unidirectional, which, in lexicogrammatical terms, occurs through grammaticalisation (2010:120). Labov however focuses on phonological changes to illustrate this unidirectional nature of change, which is especially notable in vowel chain shifts (2010:140), and although it is not particularly relevant to the main focus of thesis as such, it is nonetheless a prominent part of language change in general, and also in English, as will be seen below. Hence, from Labov (2010) the three factors of migration, contact and identity in terms of language change in general are already discernable; it is indeed this general view that I adopt in this section.

When considering English, the social history of this language provides the reasons and factors behind its birth, development, transfer and relocation (diachronic),

as well as its contemporary state (synchronic). The 'shape of English'[30] as it is today, is the result of various external factors, including invasions, conquests, wars, imperial ventures, colonial settlements, geographical migrations, and, of course, contact with various other peoples, their cultural identities and their languages along the way. All of these frequently operate on top of internal factors by e.g. arresting changes, changing the course of change or introducing new variability in language (see e.g. Baugh & Cable, 1978; Lass, 1987), although internal factors can certainly motivate language change independent of external factors (cf. Labov, 1994; 2010). The above-mentioned sociohistorical factors are the most prominent external forces at work throughout the evolution of the English language, as will be shown below.

2.3.1.1 Change in English: migration, contact and identity

In this section I argue that the special social factors of migration, contact and identity are very prominent in the historical evolution of English in general, before turning the attention to the World Englishes of today, including SAfE, in sections to follow. Migration has been a part of the human story since long before the earliest times of Neolithic, nomadic hunter-gatherer groups, moving, for example, from the region spanning the Near East to south-eastern Europe and further into the continent after the last Glacial Period (around 5 000 BC, following the Palaeolithic era spanning much further back to around 50 000 BC and according to some around 250 000 BC) (Baugh & Cable, 1978:42). According to Clackson (2007:2), among the earliest of the Indo-European (henceforth IE) languages was Hittite, which is attested nearly 4 000 years ago.

The Germanic (Teutonic) peoples, speaking an early language called 'Proto-Germanic' by historical linguists (e.g. Lass, 1987:20), comprised of tribes dispersing across Europe in thee branches, to the north in Scandinavia (North Germanic), to the north-eastern parts of Europe (East Germanic) and the most western branch of these people migrated to central Europe (West Germanic) (Baugh & Cable, 1978:30-32; Lass, 1987:9;17-20). This migration was the beginning of the pre-history of the English language. The western branch, the West Germanic people, had carried the

[30] With recognition of the title from Lass (1987).

Proto-West-Germanic language with them, which later developed into the Ingvaeonic branch consisting of Anglo-Frisian (spoken around the region of the North-Sea) and Low German (Lass, 1987:17-20).

The Anglo-Frisian group of Germanic peoples, while living in what is today Denmark and the low countries (The Netherlands and Belgium), experienced their first prominent linguistic contact with another IE language of the Italic branch, classical Latin, during the invasion of the region surrounding the Rhine river by the Roman Empire around the time of Christ, which resulted in the borrowing of some imperially useful words (Lass, 1987:36). Roman identity during this era of 'civilisation' was at the forefront, with Roman buildings, methods of administration, battle tactics and lifestyles being introduced to conquered peoples. This group, after their contact with the Romans, was to experience another, more momentous migration – one that would mark the arrival of the English language in its current homeland, England, in the 5th century AD (Baugh & Cable, 1978:42).

According to the most famous account of early English history[31] by the 8th century historian, the Venerable Bede, three strands of Germanic warrior tribes, including Jutes from Jutland, Angles from contemporary Schleswig-Holstein and Saxons from Old-Saxony, crossed the English Channel to conquer and settle Britain in roughly 449 AD (under false notions of military aid to the leader Vortigern during intertribal disputes after the Roman withdrawal) (Lass, 1987:42.) These tribal movements, according to Lass (1987:35), are known as part of the *Völkerwanderungen*, which affected various European tribes during the first half of the Christian era on the continent and resulted in shifts and reorganisations of socio-cultural identity.

Once again, contact ensued between the original Celtic inhabitants of the British Isles (speaking various Celtic [IE] languages, which also came into contact with Latin before the collapse of the Roman Empire) and these Germanic warriors, now called the Anglo-Saxons, who founded many kingdoms on the island and spoke various dialects of their original languages (Baugh & Cable, 1978:42-53). The emerging 'standard' language among the new, collective group of Anglo-Saxons, with a separate identity from their relatives on the continent, is now referred to as Old English (henceforth OE), which consequently borrowed many words from the Celtic

[31] *Historia Ecclesiastica gentis Anglorum*, c. 731 AD.

languages. Shortly after, another contact situation arose between Latin and the dialects of OE in 597 AD, but this time due to the spread of Christianity and the conversion of the Anglo-Saxons, resulting in a new religious identity expressed by the borrowing of ecclesiastical terms (Lass, 1987:46-50).

It was however not long before another nation would migrate into Britain to conquer and settle among the Anglo-Saxons around 865 AD, following various attacks and raids – the Norsemen from Denmark, who were distant cousins of the Anglo-Saxon tribes (Baugh & Cable, 1978:42-50, Bragg, 2003:18). The Norse brought with them Old Norse, a North Germanic language, which came into contact with OE, resulting again in many borrowings during war, as well as during peace, as a new social identity of a merged people grew under the Danelaw (Lass, 1987:50-54). Another, more infamous invasion ensued with the Normans (a conquering nation from Scandinavia who settled in northern France beforehand) this time conquering England from Earl Harold Godwinsson under the leadership of Duke William ('the Conqueror') in 1066 AD (1987:54-61).

Due to many complex social factors, for example Norman French prestige at court after the settlement of the Normans in the country, Old English (now regarded as low status) was influenced profoundly by this historical event; for example, it liberally borrowed scores of French words (Baugh & Cable, 1978:107-118). After the later fusion of these two nations, first causing bilingualism and later the re-emergence of English and indeed English identity, especially among the upper class and the court by the mid-thirteenth century and everyone by the fourteenth century, the Middle English (henceforth ME) period was at its height (from about 1150 to 1500) (Baugh & Cable, 1978:118;136;152).

The re-emergence of English is greatly attributed to the hundred-years war between England and France (1337-1453) influencing English patriotic identity on a massive scale with such a sense of hostile antagonism toward their neighbours only attainable through war (1978:140-1). However, when English was firmly established as the language of the court, and of law, education and literature, it continued to have linguistic contact with French, due to its accompanying notions of chivalry and cultivation remaining a lingering cultural factor in England and indeed across Europe (1978:133). By now the English had therefore acquired a distinctly English identity, even if it was often characterised by means of 'not being French' (1978:141).

These situations of linguistic contact described above are indeed the most

manifest in lexical borrowings, but the lexis of English is not the only element affected by migration and contact. However, Katamba (2005:162) deems borrowing "the main external cause of language change" in English, but also states that despite the large number of foreign acquisitions, the core of the language has remained Germanic in essence (see also e.g. Bragg, 2003:7). Croft (2009:76), for example, argues for the influence of contact reaching beyond lexis:

> "...communicative isolation is never complete. People do talk to speakers of other dialects and other languages. … ...speakers of distant languages can communicate and borrow words, and even morphology and syntax."

Grammatically, OE gradually evolved from a fully inflected, synthetic language (expressing grammatical role via morphology, like its Germanic relatives) to an analytic one (expressing grammatical role via syntax) before re-emerging as ME. This event is ascribed to various reasons including the simultaneous levelling of inflections across England (a by-product of phonological change in slurring over word-final inflections due to root syllable stress) and the influence of French prestige rendering English without a notion of standard (from court), and, hence, unchecked. As a result, many grammatical forms gradually began to simplify from the more complex OE forms, including the diminishing of noun and adjective declension, as well as strong verb forms (very few original OE forms survive today). This general process of simplification in English is also noted by e.g. McWhorter (2007). Contact with the French had the further consequence that English spelling was altered, owing to the writing preferences of the Norman monastic scribes, who believed in 'simplifying' the spelling in favour of internal consistency, as well as correspondence with French spelling.

During the Renaissance period in England (about 1500-1650 AD), with the rise of new technology, discovery and scientific advance, English saw many Latin (sometimes via French) and Greek (sometimes via Latin) borrowings of scientific, administrative and legal terms (Lass, 1987:222-226). During the transitional period into Early Modern English (EME), the Great Vowel Shift affected the short and long vowels of ME, which either became long vowels or diphthongs into EME (1987:129-131)[32]. However, with the invention of the printing press and the resulting interest in

[32] Vowel shifts like this are noted to be the most prominent kind of phonological change in language by Labov (2010:140), as also mentioned above.

standardisation, the spelling of words tended to remain constant even if their pronunciation changed, causing poor correspondence between the two areas.

During this period, colonial expansion (imperially sanctioned migration and settlement by the British Empire) was to be the one sociohistorical factor most important for the new and complex contact situations which arose throughout the world (Lass, 1987:197-294), which gave rise to the face or many faces of Late Modern English (LME) and Contemporary English (CE) and their multi-faceted identities. Linguistically speaking, the 18th to 19th centuries (LME) were characterised by prescriptive attitudes toward especially grammar, whereas the 20th and 21st centuries (CE) are known for their growing interest in variation, as well as cultural identity. The prescriptive attitude however remained prominent in South African into the latter part of the 20th century, as seen in the example of the study by Beeton and Dorner (1975) in § 1.1.1.

Although this is a very concise account of the development of English from the IE branch of languages to its contemporary state, it reveals the pivotal nature of migration, contact and identity in the social history of the language. Various tribal movements and settlements, conquests and invasions gave rise to contact between different IE languages, some Germanic, some Italic and some Celtic, all of which caused the remodelling of identity and hence language change. The many lexical borrowings that followed these events are the most visible evidence of the contact, but, as e.g. Lass (1987) asserts, phonology, grammar (morphosyntax) and semantics are affected by contact as well, as is also shown above (cf. Croft, 2009). The next section will follow this argument, maintaining that migration contact and identity have shaped and continue to shape the Englishes that emerged during and after the colonial era.

2.3.2 The World, Postcolonial and Southern Hemisphere Englishes

The linguistic paradigms of World Englishes (e.g. Kachru, 1992), Postcolonial Englishes (Schneider, 2003) and Southern Hemisphere Englishes (Trudgill, 2004), as discussed in some detail in the theoretical framework in § 2.2.1.2.1, owe their existence to the development of different forms of English around the world mainly

due to colonial migration and expansion. Various concepts surrounding the formation of New Englishes and the main factors behind such formation will be explored in this section, retaining the argument that migration, contact and identity are the sociohistorical keys to language change – even in a Postcolonial setting. The contextual concepts of a *tabula rasa* situation, Founder Effect and colonial lag will receive attention in the first place, before moving to the factor of contact (which is caused by migration and influences identity) in the emergence of the many Englishes of the globe.

2.3.2.1 A *tabula rasa* context

A specific initial colonial situation (and the most common among the Southern Hemisphere Englishes) will be considered here, namely *tabula rasa*, which is Latin for 'scraped tablet', meaning literally a 'clean slate' (Concise Oxford English Dictionary, 2006:1464). The explanation of this context by Trudgill (2004) will be highlighted here with reference to the Southern Hemisphere Englishes (SHEs).

With the transplantation of English to new territories across the world through colonising and settling, a new context arose for these migrating people. By the time of the colonists or settlers' first landing on foreign soil (be it e.g. South Africa, North America, Australia or India), the English language had never reached the ears of the indigenous inhabitants. This kind of situation, a *tabula rasa* context, therefore occurs where no prior population of native speakers existed to represent a target for newcomers (2004:24;26). Trudgill (2004:23-6) asserts that this kind of context is not present in all situations of colonial settlement[33], but that is the most common circumstance in the SHEs, of which English in South Africa forms part.

Indeed Trudgill's model of new dialect formation is supported by the concept of a *tabula rasa* environment in its (deterministic) approach. It indeed appears to be an essential ingredient in his model, when he states:

[33] As in e.g. Northern Ireland, where migration from an entire Scottish population (due to colonial transplantation across the Irish Sea) resulted in continuity with Ulster Scotts (see e.g. Hickey, 2007:431;434;438).

> "My claim about determinism is made purely with respect to the unusual type of situation in which colonial varieties develop, in tabula rasa environments, out of dialect mixtures." (Trudgill, 2004:27).

The fact that there is no reference dialect present in the new area therefore creates an optimal platform for dialect contact, resulting in the mixing of the different dialects from the input language and later in koinéization; and indeed therein lies the key implication of the *tabula rasa* context. Trudgill emphasises the "absolutely pivotal role" that the young children of the settlers (and especially the second generation) play in new-dialect formation (2004:27). The original settler communities and the children in such a *tabula rasa* context can therefore be said to find themselves in a very influential position regarding language, as the following section will reveal.

2.3.2.2 The Founder Effect

This section considers the linguistic impact of the original settlers in a colonial setting, mainly according to the views of Mufwene (2001; 2008), who often makes reference to the Founder Effect (also called the Founder Principle). Before Mufwene's application of this effect to English development across the world is discussed, it is useful to note that 'Founder Effect' is an evolutionary term applied not only to English, but also to the development of language and culture in general. In his recent article in *Science*, Atkinson (2011:348) draws the following conclusion after his investigations into the expansion of the human race (and their languages) from Africa:

> "To the extent that language can be taken as an example of cultural evolution more generally, these findings support the proposal that a cultural founder effect operated during our colonization of the globe, potentially limiting the size and cultural complexity of societies at the vanguard of the human expansion."

The Founder Effect therefore has far-reaching implications for language and culture in general, but it is also pivotal to the understanding of the English spread around the world in the more recent past.

Mufwene (2001:26-27), who also adopts a language evolution perspective and emphasises the principles of competition and selection of linguistic features, refers to

the definition of Zelinsky (1992 [1973]) for his 'Doctrine of First Effective Settlement', which is equivalent to the Founder Principle. It reads:

> "Whenever an empty territory undergoes settlement... the specific characteristics of the first group able to effect a viable, self-perpetuating society are of crucial significance to the later social and cultural geography of the area, no matter how tiny the initial band of settlers may have been... ...in terms of lasting impact, the activities of a few hundred, or even a few score, initial colonizers can mean much more for the cultural geography of a place than the contributions of tens of thousands of new immigrants generations later." (Zelinsky, 1992 [1973]:13-14)

Similarly, in an interview with Michael Collins (Collins & Mufwene, 2005), Mufwene suggests that the Founder Principle centres on the large influence of the earliest settlers in a new territory on the evolution of the new variety, which "can be disproportionate to their size, because every new instalment of newcomers will find it more practical to speak like the locals than to speak like outsiders" (2005:453). He further states that children, not adults (similar to Trudgill's (2004) view), are usually the catalysts for this process, as the children want to associate themselves with the new language and culture and consequently tend to "learn everything local, including the way of speaking" (2005:453).

The Founder Effect prevails in the majority of cases, but a new layer of immigrants can impede this effect if they are suddenly more numerous, socio-economically more powerful or regarded as more prestigious than the founders, according to Collins and Mufwene (2005:453). However, if this does occur, at least some traces of the older systems are likely to survive in future generations, which means that the Founder Effect is a phenomenon worth considering when investigating any colonial or Postcolonial variety (2005:453).

The Founder Effect is therefore a significant element to consider in language evolution, not only of language and culture in general, as noted by Atkinson (2011), but especially in the colonial contexts of English, as posited by Mufwene (2001), seeing that it is at play even in the smallest settler community. Ergo, this phenomenon is highly relevant to this thesis, since the 1820 Settlers (a relatively small group) is part of the study population of the empirical study, as Chapter 3 will show in more detail. The colonial context, apart form the concepts of a *tabula rasa* situation and Founder Effect, also incorporates the concept of colonial lag, which will be illuminated in the next section.

2.3.2.3 Colonial lag

The third issue of interest to social history and language change in the paradigms in question is that of colonial lag. In the 6th chapter of his book, Trudgill (2010) discusses this phenomenon, which he deems part of the natural process of language change within a transplanted linguistic society. The term *lag*, as Trudgill uses it, was first introduced into linguistic literature by Albert Marckwardt (1958) on the subject of American English, who described it broadly as the delay that occurs before a newly transplanted language becomes adapted to its new environment (1958:80). Trudgill supports this notion to an extent, but does not perceive the adaptation to a new environment as the greatest causal factor of colonial lag (2010:131). The main examples Trudgill uses to demonstrate his stance stem from phonetic data from the Origins of New Zealand English (ONZE) project, which is based on audio recordings of elderly people in the 1940s (2010:131,137).

Trudgill provides a list of five contemporary phonetic features of NZE (New Zealand English) and AusE (Australian English) that differ from BrE in that they are more conservative, such as the retroflexing of /r/, the use of /æ/ in e.g. *dance* and the lack of innovative pre-glottalisation in e.g. *match* (2010:137). Despite this, there are also some endogenous features that NZE has developed, that have not occurred in BrE, suggesting that NZE and other SHEs are only conservative with respect to some (regular) features (2010:138).

It is therefore that Trudgill argues that it is not adaptation to a new environment that creates colonial lag in a transplanted linguistic community. He asserts:

> "'Lag' in its linguistic sense has nothing to do with adaptation to a new environment. It is simply a delay in the normal progression and development of linguistic change which lasts for about one generation, and which arises solely as an automatic consequence of the fact that there is no common peer-group dialect for children to acquire in first-generation colonial situations involving dialect mixture, and where therefore the speech of the older generation provides the model." (Trudgill, 2010:142).

Colonial lag, which is here treated as an automatic consequence of a *tabula rasa* situation, therefore usually lasts for one generation (see also Trudgill, 2004:34), which will be considered later in the interpretation of the empirical results of this

study. This happens because peer groups usually serve to propagate ongoing change, but when a stable peer group is not present, as in a *tabula rasa* situation, the older generation, which is not that advanced in terms of the change, becomes the reference group and, hence, colonial lag transpires. With colonial lag being described succinctly, the next section will discuss the role of language contact on English within the socio-linguistic paradigms.

2.3.2.4 Contact in sociolinguistic paradigms

Mufwene (2008:25) maintains the following:

> "Both contact and migration influence the ways competition and selection affect variants differentially in the colonial ...populations."

In this section I will show that contact resulting from migration has influenced the Englishes of the world and their accompanying identities to such an extent that contact is the most prominent of the three previously-mentioned factors in the development of these Englishes. Here the views of Kachru, Trudgill (2004; 2010) and Schneider (2003) will be examined again, but not before some general observations about language contact and its impact on English are made.

It has been noted in the section above that English has lost some of its grammatical complexity over time in its earlier history (e.g. the loss of inflections). On this point, McWhorter (2007:59) notes:

> "In the emergence of Modern English, simplification dominated complexification to a greater extent than in any other Germanic language. ... ...this simplification was no [sic] a happenstance peculiarity, but due to a sociohistorical factor hindering the full transmission of the grammar across generations."

McWhorter later reveals that the sociohistorical factor he refers to here is in fact language contact (2007:60). Apart from the loss of inflections, which is noted above, English has shed many other Germanic features connected with analyticity, for example reflexivity marking, which rendered it less complex than its sisters, suggesting that something other than an unbroken internal development is at play in its development (2007:60). He argues that a contact-based, external explanation

provides a principled account for the simplification of English (2007:60) and further suggests that this qualitative difference in general structural complexity between English and its sister languages is caused by incomplete (non-native) acquisition in contact situations that arose for English in the more recent future (2007:252).

The latter statement implies that the non-native speakers of English, who have learnt it under colonial circumstances, for example, are the driving force behind the simplification process of English in more recent times. Yet it also implies that this phenomenon is not a new one; it has indeed been central to the history of English since the Anglo-Saxon migration to the British Isles. This section is not so interested in the simplification process of English, but it is very interested in the fact that non-native acquisition of English in contact situations can have such a major influence of the structure of the language itself. This non-native influence can therefore be anticipated to continue in its impact on English today.

Similarly, Trudgill (2010:xiii) proposes that contact is the key solution to the riddles surrounding language change due to colonisation (albeit with a focus on dialect contact). On the other hand, as was shown in sections above, Schneider emphasises identity as the key factor for the formation of a New English, but also considers contact in the following statement:

> "...what counts here is not the colonial history or the former colonial status of a given country per se, and also not the specifically British connection, but rather the type of language contact situation caused by ...historical circumstances, the expansion and relocation of the use of a single language to new territories where a characteristic type of language contact situation evolves." (Schneider, 2003:235)

Despite the fact that identity creates an 'enabling frame' for the adaptation of variants across groups (i.e. the STL- and IDG-strands), it is contact that remains essential to the introduction of such variants. Hence, language contact induces innovation (cf. Van Rooy, 2010; Croft, 2000). Yet, for Schneider (2003), identity remains the main driving force for propagation to take place. According to this approach, language contact situations therefore have the potential to drive the 'birth' of New Englishes, but identity enables the spread or diffusion of new features.

A similar view on contact is adopted by Mesthrie (2003) when he argues that the field of contact linguistics is the way forward for studies in World Englishes, seeing that these New Englishes "are after all contact varieties" (2003:449). He further asserts that the historical input to each individual World English should not be

ignored and proposes that more sociohistorical work needs to be done to uncover the roles of e.g. missionaries, settlers, soldiers, traders, teachers, etc. in providing input to a contact situation (2003:459). This stance links with Trudgill's (2004) view to some extent, regarding emphasis on contact and input, but it is also adheres to the views of Schneider (2003), especially with regards to the importance of sociohistorical considerations. This indeed is the stance of this thesis.

It is however important to note that not all linguistic change in the Englishes of the world will necessarily occur due to contact, and it is therefore useful to consider when this kind of interpretation is relevant and when it is not. Thomason (2007:42) defines a contact-induced change as "a particular linguistic change [that] is caused at least in part by language (or dialect) contact if it would have been less likely to occur outside a particular contact situation". Levey and Poplack (2010) propose that alternative criteria to those of Thompson (2007) are needed in this case, and state the following:

> "A candidate for contact-induced change in a contact variety is present in the presumed source variety and either 1) absent in the pre-contact or non-contact variety, or 2) if present (e.g. through interlingual coincidence), is not conditioned in the same way as in the source, and 3) can also be shown to parallel in some non-trivial way the behaviour of a counterpart feature in the source." (Levey and Poplack, 2010:398)

These criteria will be carefully considered when interpreting the results of this thesis in Chapter 5.

Apart from this, Mufwene (2008) argues that linguistic contact should not be understood in terms of whole communities, but rather on an individual level (cf. Mufwene 2001), as seen below.

> "Contact plays an important role in my approach to language evolution, the most significant part being contact between individuals rather than between populations... ...the coexistence of two populations in the same geographical area is not a sufficient condition for language contact. They must interact with each other. At the level of linguistic communication, this is made possible by interacting individuals, who can spread features from the other language among monolingual members of their respective languages." (Mufwene, 2008:17)

When individuals interact, as foregrounded in the quote above, their social identities also interact – and an additional factor of language change is therefore added during

contact between different people. Thus, contact is a pre-condition for interaction at the level of identity.

To summarise my argument for contact at this point: migration inevitably gives rise to language contact situations, and those situations assist in shaping not only the language, but also the social identities of individuals and societies. Seeing that the language contact situation is clearly the key link between society and linguistic change in the Englishes of the world, the role of contact in creating the World Englishes, the Postcolonial Englishes and the Southern Hemisphere Englishes will be explored next, according to the views of the primary proponents and contributors to these paradigms.

Firstly, Kachru considers contact as a widespread ingredient in the evolution of the World Englishes across various continents, including Africa. He suggests that throughout the diaspora of English, four major phases in the development of a new variety can be discerned, the first being the implanting of English within a *tabula rasa* context by a relocated community and secondly "contact with genetically and culturally unrelated major languages: in Africa with the Bantu and Niger-Congo languages, in Asia with the Dravidian languages, and in Southeast Asia with the Altaic languages, to give just three major examples" (Kachru, 1996:136). He therefore mentions the Bantu and other Niger-Congo languages with reference to Africa, but no mention of Afrikaans, in fact a genetically related language to English, is made, which will receive attention later in the chapter. Kachru's third phase involves a diversity of various contexts and inputs in "imparting English education" (1996:136), and the fourth the "sociolinguistic reincarnations" and acculturation of English to its environment, including the "liberation (...) from the traditional canons associated with English" (1996:136-7).

Kachru (1996:144) further states that this process of 'opening up the canon' in English has occurred to an unprecedented extent, and that language contact has initiated the processes of its change in terms of grammar, lexis, style and discourse. Kachru adds that various language contact situations are contexts which spark interest in the field of World Englishes in general, but also beyond that, to other academic interests as well, as seen below.

> "I believe that world Englishes provide a challenging opportunity to relate several areas of academic interest: language, literature, methodology, ideology, power, and

> identity. The contexts for inquiry involve diverse cultures and varied situations of contact and creativity. There is a cross-cultural arena with one linguistic constant, English." (Kachru, 1996:149)

English, with its multi-faceted academic interest field, has therefore no doubt been influenced by not only culture and creativity, but also by extensive interaction, i.e. contact with speakers of other languages during its evolution in various locations.

The role of contact is also pivotal in the formation of a new variety of the SHEs, according to Trudgill (2004; 2010), but here the emphasis is not so much language contact as it is contact between the various English dialects of the settlers (Trudgill, 2010:181). Trudgill (2004:23) devotes much attention to this factor of dialect contact, which he regards as having been, "in most instances of colonial new-dialect formation, the most important". Even though this consideration is relevant when later discussing the formation period of SAfE, it is language contact in which this study takes the greatest interest, which is why Schneider's view is perhaps more apposite in this case.

Schneider (2003) also stresses the role of contact in the evolution of New Englishes or Postcolonial Englishes, as is seen in an above quote in this section. Contact situations can have many forms and arise from different types of scenarios[34], which are surveyed by Schneider (2003:235); these include "the movement of one group into another group's territory, immigration of small groups or scattered individuals, importing a labor force, or cultural contacts through long-term neighborhood" (cf. Mufwene, 2001:204-6).

Along the same lines, and as Schneider also notes (2003:235), Gupta (1997:53-6) recognises five different patterns for English across its territories, including the following:

(i) monolingual ancestral English (e.g. the USA and Australia)
(ii) monolingual contact varieties (e.g. Jamaica)
(iii) multilingual scholastic English (e.g. India)
(iv) multilingual contact variety (e.g. Singapore)
(v) multilingual ancestral English (e.g. South Africa and Canada).

[34] Of course dialect contact and language contact can and often do take place simultaneously in the environment of a New English.

It is evident from Gupta (1997:49;53-55) that different combinations of either mono- or multilingual scenarios for English have arisen with their primary driving forces either being ancestral (transmission from generation to generation), scholastic (taught in schools and primarily enjoys classroom support [associated with prestige]) or contact-based (arising from contact situations; usually used as a lingua franca [with primarily community support], but varieties can have major syntactic differences [e.g. creoles]). Gupta (1997:53-54) uses the terms 'monolingual' for those countries where "most people speak only English as a native language", and 'multilingual' for those countries where "people are seldom native speakers of English". It is important to note that Gupta emphasises that these labels should be seen in the light of predominance, where exceptions may and do exist. South Africa, for example, is classified among the multilingual ancestral group, which essentially entails that the descendants of the British settlers are "not a sufficiently large majority to absorb the majority of other groups in the country" (Gupta, 1997:56). Indeed such descendants are a minority in South Africa, whereas the majority of the population finds itself in a scholastic English setting (1997:56).

I generally agree with Gupta's short mention of South Africa, but, I would like to suggest that the country additionally displays traits of the third kind of scenario in the 19th century, mostly for individual Afrikaans speakers learning English as a second language. Acquisition was mostly reliant on community support before the Afrikaners received widespread scholastic instruction for English (see § 2.3.3.4.1), which suggests that it was a kind of contact variety that later transformed into a scholastic variety[35] (see e.g. Branford, 1996:38; Steyn, 1980:127-128).

Schneider makes further mention of language contact as being part of the events in phase one of his dynamic model (foundation). He states:

> "In almost all cases indigenous languages are spoken in [the] area, so a complex contact situation emerges. In fact, contact operates on two levels, independent of each other at first, involving dialect contact and language contact, respectively: it concerns both the group-internal communication among the English-speaking settlers and the interaction between these settlers and the indigenous population – two different types of linguistic ecologies." (Schneider, 2003:244)

[35] Contrastingly, the native African population enjoyed a primarily scholastic scenario due to the mission schools.

Here Schneider, like Kachru (1996), mentions contact between indigenous populations and their languages, but unlike Trudgill he does draw specific attention to language contact on top of dialect contact. It is however important to note that Schneider does not necessarily rank contact and identity against each other, but rather suggests that contact 'enables' or serves to 'activate' the factor of identity, as also mentioned earlier in this section.

Van Rooy (2010:8) indeed notes that a change first has to be invented before it "can spread and become established as a feature of [a] speech community", and therefore that (to use Croft's [2000:3-6] terms, as described in § 2.2.1.2), innovation and propagation/diffusion do not necessarily occur simultaneously. He further states that "[i]t is more likely that a time delay can be observed between the innovation and propagation of a feature, allowing for the strong possibility that the processes of innovation and propagation result from different causes" (Van Rooy, 2010:8-9). One of these causes can rely on contact-induced factors, contributing "to the feature pool of innovations" (Croft, 2000:145-8) that "result from connections made between two systems by bilingual speakers" (Van Rooy, 2010:9). Labov (2007:380) indeed notes that "contact across communities involves learning, primarily by adults, who acquire the new variants of the originating community in a somewhat diluted form" (this happens because adults learn slower than children and tend to lose "much of the fine structure of the linguistic system being transmitted" [2007:380]) and that the "diffusion of specific linguistic structures is [therefore] one of many changes that spring from adult language contact" (2007:382).

To summarise, it is important to note that Schneider (2003) does not consider contact as the prime determinant of the formation of a new English variety, but he does deem it an important 'enabling' factor nonetheless. He considers contact as part of the relationship between his two strands (STL and IDG), which is an underlying association present in all the phases of his model, albeit not the principal consideration, which he believes to be identity (being 'enabled' by e.g. contact). Apart from Schneider's view, this section shows that Trudgill (2004) emphasises dialect contact rather than language contact. Kachru (1996:144), however, considers the contact situation as a separate and critical phase (the initiator) in the evolution of the grammar, lexis, style and discourse of World Englishes, which share some aspects of Schneider's notions regarding the 'enabling' role of contact.

The next section will discuss South African English with reference to its social history and the impact of migration, contact and identity in its evolution into a new variety of English.

2.3.3 A sociohistorical overview of SAfE

Seeing that the aim of this thesis is to investigate changes in the modal system in SAfE with some emphasis on its context, it is necessary to situate it against the broader canvass of varieties of English in the country (as there are many, e.g. Black and Indian South African English (BSAfE and ISAfE), as well as Afrikaans English (AfkE). Thus, in this section I will confine my description to SAfE, but not without consideration of its broader context. The history of English in South Africa is an intricate one – indeed Schneider (2007:173) states that the level of complexity that South Africa possesses in its "sociolinguistic constellations" exceeds that of any other country studied by him in his book entitled *Postcolonial English: Varieties around the World* (2007). Thus, it would be practical to start from the beginning, as history and language are entangled right from the start in the complex tale of the formation of SAfE[36].

2.3.3.1 Migration

English was transplanted to South Africa in the typical colonial, *tabula rasa* context, where no existing population of native speakers existed to represent a target for newcomers, as also noted above (Trudgill, 2004:24,26). This transplantation occurred during the time of British rule in South Africa, more specifically during the period of British occupation of the Cape, firstly from 1795 to 1803, then during the second occupation of 1806, and thirdly after the Cape was officially made a British colony

[36] The historical account of WSAfE, and of Afrikaans (to follow), is largely based on the structure, and some prominent facts outlined by Van Rooy (in press). However, even though many of Van Rooy's sources were also used in these sections, the original sources, with the addition and subtraction of a few, where also surveyed independently by the author. I therefore give acknowledgement and recognition to Bertus Van Rooy (in press) with many thanks.

around 1814-1815 (Hockly 1966:2; Giliomee and Mbenga 2007:85). The British were however not the first to have annexed the Cape, as the Dutch (Afrikaans) community had already settled there about two centuries before them, apart from the indigenous population – the details of which will be discussed later on in this main section.

After the period of occupation, the first significant group of plus-minus 4 000 British Settlers mainly from London and the Home Counties (Mesthrie, 1993:27; Lanham & Macdonald, 1979:14), but also some from the West Country, Yorkshire, Scotland and Wales (Lass, 2004:371), arrived in the Eastern Cape between April and June in 1820, out of a group of about 80 000 applicants (Butler, 1974; Lanham, 1996:20; Giliomee and Mbenga, 2007:85). Their chief motivation was of an economic nature, as the conservative government in England offered an opportunity for emigration as a scheme of poverty relief after the Napoleonic Wars. The reality was, however, that most settlers were citizens of some financial means (however not rich enough to start their own businesses in England), as the government did not intend to make the Colony a 'dumping site' for the poor of England. Half of these settlers were urban artisans with no intention of becoming rural farmers, and consequently focused on trading or producing goods (Giliomee and Mbenga 2007: 85-86).

This Eastern Cape settler strand therefore comprised English speakers from all over Britain, although the majority came from the South-East of England, speaking regional British dialects (Lanham 1996:20; Bowerman 2004a:931), and served as input to the broad contemporary (native) variety of SAfE. According to Bowerman (2004a:931-932) the regional dialects levelled out within two generations, forming a single dialect known as Cape English (CE). Lanham and Macdonald (1979:72) indeed also note that Cape English originated from the vernacular that developed from "the first childhood peer groups of over 1000 under-twelve children of the 1820 Settlers, and in the first generations born in Africa".

Bekker (2012:129) notes that the "prior presence of the Afrikaans population" complicates the *tabula rasa* context to a certain extent, but that, when it comes to English in itself, "the Eastern Cape settlement constituted a *tabula rasa* colonial context". Therefore, the first Western Cape occupants, as well as the first wave of settlers in the Eastern Cape both had their own *tabula rasa* context, owing to the geographical distance of up to 1 000 kilometres between them (see also Lanham, 1967:104). Bekker (2012:129) further suggests that this situation is not construable as

"constitutive of a Founder Effect that nullifies the *tabula rasa* status of the later settlement", here denoting the 1820 settlement.

A second wave of 5 000 settlers followed in Natal between 1847 and 1851, who were mostly from the northern counties of England, especially Yorkshire and Lancashire. They included a very large proportion of people from the upper-middle class, as well as retired officers and even some aristocrats, according to Lanham (1996:20-21) and Giliomee and Mbenga (2007:148). These settlers were enticed by the promise of farmland, and indeed between 1854 and 1866 sugar cane farming became the major economic resource of Natal, with plantations increasing from 217 to 5 170 hectares as the period progressed (2007:148-149).

The English of the Natal settler strand exhibited less regional variation than that of the Eastern Cape strand and remained fairly stable and contained within its own borders, and was perceived as the 'upper-class' form of English in South Africa (Bowerman, 2004a:932), maintaining a British identity (Lanham, 1996:23). The Natal dialect became the basis for the emerging provincial standard of SAfE in the early decades of the 20th century (Lanham and MacDonald, 1979; Lanham, 1996).

Bekker (2012:131) suggests that the circumstances surrounding this second settlement are not as unambiguous in terms of a *tabula rasa* context and the presence of a Founder Effect. He states:

> "...while not as clear-cut as the Eastern Cape scenario, there is evidence to suggest that the Founder Effect, in the form of a pre-existing population of migrants mainly form the Cape Colony, was minimal in Natal; and that the context was effectively *tabula rasa*." (Bekker, 2012:131)

However, Bekker (2012:131) further states that this 'minimal' Founder Effect (of an early-Cape English koiné) has nonetheless affected the new variety residually, seeing that there is "some evidence of migration from the Cape before the main influx into Natal".

A third wave of immigration during the 1870s and 1880s, as well as migration on an intra-national level, occurred after the discovery of diamonds in Kimberley and gold in Johannesburg. During this era, more than 400 000 fortune seekers, workers and capitalists were lured to the diamond and gold fields from all over Britain, Eastern and Western Europe and the southern parts of South Africa (Giliomee and Mbenga, 2007:185). Lanham and Macdonald (1979:71) give a similar account:

> "Mainly on the Witwatersrand (Johannesburg and its environs) was the SAE community's heterogeneous past woven in to the more homogenous present; it was here that the different ethnic, descent and settlement groups became socially integrated more than they had ever been before. … This was the 'new society' which attracted the British 'colonials' from the Cape and Natal; the Afrikaner – reluctantly at first and in small numbers until large-scale urbanization brought him into the cities; the new immigrant from Britain and the Empire together with the European immigrant, mainly Jewish, from Eastern Europe and Germany."

Perspective is gained on the extent of immigration and migration during the mining era when one considers that the total white population of South Africa (both Englishmen and Afrikaners) only amounted to around 340 000 before the discovery of precious minerals (2007:185, 200).

In the case of Johannesburg, gold was discovered in 1886, and what was previously farmland grew to an urban mining area with 44 mines within two years. By the early 1890s the new city of Johannesburg already had 50 000 white inhabitants, of which only 7 000 were Afrikaners (Lanham, 1996:23; Giliomee and Mbenga, 2007:200). This too is put into perspective when considering that about 45 000 Dutch (Afrikaans) speakers greatly outnumbered the 5 000 British settlers at their time of arrival (1820) (Branford, 1996:38).

From this third wave of immigration and migration new complex contact settings inevitably arose, where the local English-speaking population was greatly outnumbered by immigrants, some of whom were from much higher classes in the British society of the time, and succeeded to a fair degree to maintain the British class structure (Lanham, 1996:20-24). Lanham (1996) maintains that class-consciousness remained a feature of the English-speaking colonial population throughout the 19th century and into the 20th century. In addition, and as touched on above, Lanham (1996:23) refers to a very strong sense of British identity among colonial English speakers lasting well into the twentieth century, which is somewhat contrary to the context of a Founder Effect in a traditional *tabula rasa* situation in which native norms are usually shed within one generation to accommodate acculturation (see Trudgill, 2004; Mufwene, 2001; Kachru, 1996).

Bekker (2012:133) argues for the following:

> "...for all intents and purposes, Johannesburg constituted a *tabula rasa* context for dialect mixing, koinéization and new-dialect formation."

Therefore the third wave of migration has constituted a third and separate *tabula rasa* situation for English in the country, and, according to Bekker (2012:133), no Founder Effect was in operation as a consequence, leaving "ample opportunity for mixing and levelling to occur" and hence for the development of a new kind of SAfE, "different in important respects to the original two colonial varieties", Cape English and Natal English. The region surrounding Johannesburg was, after all, literally another country altogether, as it was part of the Zuid-Afrikaanse Republiek[37] (also known as the Transvaal – see Figure 1 below) and hence independent of the Cape Colony and Natal[38].

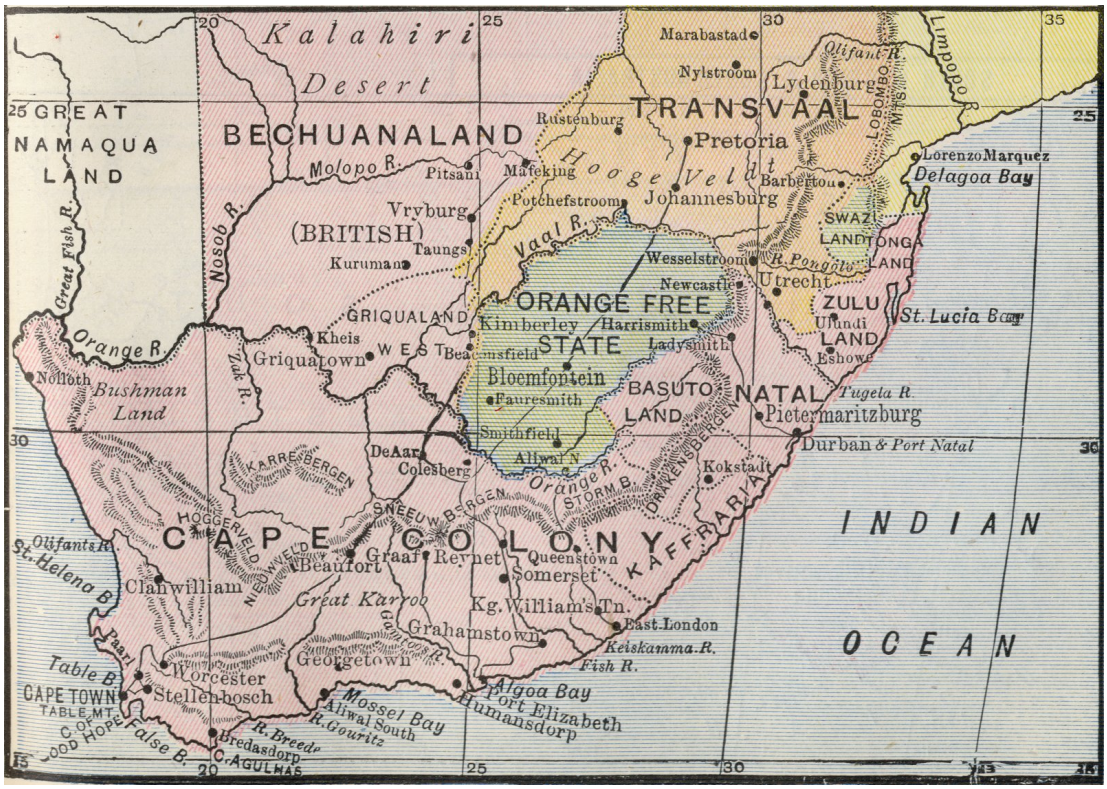

Figure 1

Map of Southern Africa[39] ca. 1898

---

[37] One of two Boer (Afrikaner) states (along with the Orange Free State) that were established in the 19th century and later toppled by the British after the Anglo Boer War (Bekker, 2012:133).

[38] Today, South Africa comprises nine modern provinces, namely the Western, Eastern and Northern Cape, KwaZulu-Natal, Free State, Mpumalanga, North-West, Gauteng and Limpopo.

[39] Map available from:
http://www.probertencyclopaedia.com/photolib/maps/Map%20of%20South%20Africa%201898.htm

Thus, unlike other colonial settings, the South African society was characterised by a stronger sense of prestige, which may have had consequences for the degree of determinism in the formation of a stable dialect, further aggravated by continued waves of settlement and migration (with three different *tabula rasa* contexts) resulting in stabilisation (or in Trudgill's (2004) terms 'koinéization[40]') taking a longer time to happen, which might point toward colonial lag taking place for a little longer than expected. These continued processes of migration have almost certainly impeded the Founder Effect of the original 1820 Cape Settlers in SAfE, but that does not mean that it had been extinguish altogether (see Collins & Mufwene, 2005:453). Futhermore, in my view, this does not necessarily have the implication that one case of colonial lag was in a sense 'dragged out' over a long period of time, but, moreover, that the very fact that there were three separate *tabula rasa* contexts and hence three opportunities for colonial lag to occur, may have lessended the influence of its effect, just as it can be said for the Founder Effect. By the time that there had already been three sets of linguistic models provided by the older generation for the first-generation children in each of the three *tabula rasa* contexs, i.e. near the turn of the 20th century, extensive contact with Afrikaans children (and adults) would already have taken place over the course of the 19th century, as will be discussed in § 2.3.3.4, which means that although prestigious British norms may have been encouraged in each of the *tabula rasa* contexts, and despite the provincial status of Afrikaans, the inevitable contact between the English and Afrikaans communities on especially individual level is a factor that would have been influential over the scope of all three *tabula rasa* contexts. Hence, the broad peer group of the first-generation children in each *tabula rasa* context, and perhaps especially during the third wave of migration after the gold rush, would have included Afrikaans children. With the onset of bilingualism (cf. § 2.3.3.4.1), this added reference group for the English-speaking children born in South Africa over the course of the 19th and early 20th centuries may therefore have obscured the effect of colonial lag to some degree, although, just as in the case of the Founder Effect, it was probably not extinguished altogether.

[40] Defined as the process by which a koiné, which might later develop into a standard language, is formed. The word 'koiné' denotes a form of Greek that became widespread in die area around the eastern Mediterranean after the conquests of Alexander the Great in the 4th century BC, and comes from a word meaning '(in) common' (Matthews, 2007).

2.3.3.2 The extended formation of SAfE and its contemporary state

When taking the above account into consideration, SAfE can be said to have had an extended formation period compared to other colonial varieties, largely due to issues surrounding status and stigma. As mentioned earlier, the first local variety to stabilise, Cape English, enjoyed a lesser status than the Natal dialect, which later became the 'standard' for SAfE (Bowerman, 2004a:933), as well as a provincial standard (Lanham & MacDonald, 1979:20;81). Cape (Colonial) English therefore became stigmatised as the variety of the working class, whereas Natal English evoked a social image of 'educated' and 'elitist' speakers (Bowerman, 2004a:933), with a strong British orientation (Lanham, 1996:23).

At the same time, it is important to note that for the largest part of the twentieth century, an exonormative Conservative SAfE[41], continued to reign as the most prestigious South African English variety. According to Lanham and MacDonald (1979:30) the speakers of Conservative SAfE were generally born before 1930, and even though they were still middle-aged in the late 1970s when they were writing texts, by today they would have become a small minority in the living population of English native speakers. Consequently, their associated exonormative prestige variety has largely disappeared from public life in South Africa.

SAfE can therefore be said to not fully conform to the model proposed by Trudgill (2004), because of the extended formation period and the continued influence of Standard British English as reference point well into the twentieth century. Trudgill (2004:23) proposes a formation period of around 50 years or two generations in colonial situations, which suggests that SAfE's formation period spanned from 1820 to 1870 (Trudgill, 2004:24; see also Lanham, 1996:19); but these above-mentioned complications suggest that extra time might be added in the process, due to the third wave of immigration, which would move the formation period even beyond this. Indeed Bekker (2012) suggests that there are three distinct periods of formation for SAfE, given the three different waves of migration noted above. His view will be considered in the next section.

[41] Bekker (2009; 2012), for example, prefers the term 'Cultivated SAfE', as noted in § 1.2.1.

In terms of Schneider's (2003) typology of five stages in the evolution of New Englishes, the variety in question in this thesis only represents the STL-strand, and the focus of this study will principally deal with Phases Two ('exonormative stabilization') and Three ('nativization'). Schneider (2007: 185) suggests that since 1994, South African English has begun to enter Phase Four ('endonormative stabilization'), but both Bekker (2009, 2012) and Van Rooy (2010) are cautious about affirming this. Schneider (2003) and Trudgill's (2004) models will receive further attention in this chapter, and the analyses of this study are anticipated to help establish how far exactly SAfE has evolved in terms of these models and where the variety finds itself in contemporary terms. The contemporary state of SAfE will briefly be described below before I undertake an exploration of a framework for SAfE in the next section.

During the first decade of the 21st century, the percentage of native English speakers in South Africa who can trace their ancestry back to Britain (Schneider's [2003; 2007] STL-strand) has slightly dropped in terms of the total population, among the ten other official languages of the country, according to the 2001 and 2011 censuses (Statistics South Africa, 2003; 2012). In 2001, white L1 speakers of English amounted to 1 687 661, which is only 3,7% of the total population estimated to number 44 819 788 (Statistics South Africa, 2003:6;18). However, by 2011 this percentage had dropped to 3,1%, seeing that only 1 603 575 white L1 speakers of English were counted among the total population estimated to now number a larger 51 770 560 (Statistics South Africa, 2012:18;26).

Despite South Africa's colonial history, the contemporary remnants of its earlier affiliation with Britain (as part of the Commonwealth), and the fact that BrE norms are generally regarded as the standard in e.g. the education sphere, contemporary SAfE is also part of a conglomerate of English varieties in an ever-growing culture of globalism – one which is influenced by the USA to a large extent. This links with the 20th-century trend toward Americanisation, for which Leech *et al*. (2009:253-4) found grammatical evidence in terms of "American 'leadership' being one of the major moving forces on BrE", not necessarily "in the sense that direct dialect contact is the reason for change", but moreover in the sense that "AmE is ahead or is changing more quickly than BrE" in a kind of "follow-my-leader pattern". Mair (2006:193-4) suggests that this influence stretches beyond BrE, and agrees with the general findings of Leech *et al*. (2009) noted above:

> "...instances of American influence on other varieties of English in the twentieth century are not hard to find. Given the global presence of the United States in what may well be referred to as the "American Century" in future historiography, they are not surprising, at all. ... However, a dispassionate look at contemporary linguistic developments shows that popular discussions hopelessly overemphasize the influence of American English on the development of the language as a whole. ... First, we need to realize that many instances in which British (or Australian or Irish, etc.) usage seems to follow American practice do not necessarily represent direct American influence. Rather, they show all varieties of English developing along the same lines and toward the same putative end-point, but at slightly different speeds. ... Second, American influence operates selectively. ...it is pervasive in the lexicon, modest in the grammar, and almost nonexistent in pronunciation. Also, it tends to be restricted to communicative domains with a global reach – from international science and research to entertainment and mass culture. It is less in evidence in local communicative domains (for example ordinary face-to-face conversation) or domains with a strong local tradition (for example, literary writing...)."

Indeed Trüb (2008:80;327) produces evidence of AmE impact on SAfE on lexical, phonological and morphosyntactic levels, despite widespread negative attitudes toward Americans and AmE among South Africans (mainly based on interviews). The channel of transmission of this impact is the American mass culture, including the cinema, television and music (Trüb, 2008:329), as also mentioned by Mair (2006:194) in the above quote. Since I will compare the findings for SAfE modality with those of BrE and AmE (e.g. Mair & Leech, 2006), the results of this thesis might cast some light on this subject – whether SAfE modality follows a similar or different pattern to that of BrE and/or AmE – which might contribute to the evaluation of AmE impact on SAfE. The next subsection explores the implications of the extended formation period of SAfE in terms of theoretical frameworks.

2.3.3.3 Toward a framework for SAfE

The material point of discussing the above-mentioned frameworks of theory and empirical techniques is of course to apply and evaluate them in terms of written South African English. The position of SAfE relevant to language theories and especially the three major linguistic paradigms of English will therefore be discussed in this section. The question here is where SAfE fits into Kachru's three-circle model (1992), as well as Trudgill (2004) and Schneider's (2003) models of SHEs or PCEs respectively; if it does, at all, fit into these models. The new Three-stage Koinéization

Model of the Formation of SAfE of Bekker (2012; 2013) will also be considered in particular, as an alternative view to the more traditional ones.

The general question in this section can be broken down into four questions: has SAfE developed along the lines of Trudgill or Schneider's models; what are the driving forces behind the 'birth' and evolution of SAfE; how far down the path of evolution does this variety currently find itself; what does this mean in terms of SAfE's 'place' in the English-speaking world? Even if these questions are not all completely answerable at this point in the study, or even in South African English studies in general, this section serves to evaluate previous applications of models to the variety based on sociohistorical information and to explore further possibilities.

The history of English in South Africa is particularly complex, and it is exactly this complex context and the many varieties of English in the country that make Kachru's tree-circle model of World Englishes (1992) difficult to apply. The fact that this model is arranged according to territories or countries (in terms of L1, L2 and EFL users) and not to individual English varieties, poses the first challenge for application to South Africa itself. South Africa is indeed not mentioned in Kachru's model[42], as also noted above. Schneider explores a reason for this:

> "...in the traditional models, countries like South Africa, with a strong proportion of English speakers but also other population groups whose native languages are recognized as official languages as well, have defied a clear categorization and have therefore frequently been sidestepped in such listings, or been classified somewhat forcedly." (Schneider, 2003:243)

Schneider (2007) argues that the linguistic situation of English in South Africa possesses a special kind of complexity, largely due to its history (cf. Kachru, 1996:137). As seen above, the many South African Englishes (including also BSAfE for example[43]) are the cause of concern here, but for the SAfE variety, at least, Kachru's model is not so wholly unhelpful. As a native-speaker variety, SAfE should naturally fall into the inner circle of this model, even though it is not mentioned. There is still debate as to which of the other Englishes of the country fall into the

[42] In Kachru (1996), however, it is mentioned as being part of the Outer Circle, which categorises it as an L2 variety. This is not entirely in accordance with reality, as, while it is a L2 for most South Africans, it is also a L1, and sometimes a L3 or even beyond for many others.

[43] For the purposes of this thesis, the other varieties of English in South Africa will not be considered in terms of linguistic models.

inner, outer or expanding circle, but this is not an issue I wish to address here. Kachru's model is therefore useful to classify SAfE among the native Englishes of the inner circle for the purposes of this thesis, but the model hides the presence of other speakers or languages that give rise to e.g. language contact.

Schneider's model of the evolution of new Englishes (2003) and Trudgill's model of new dialect formation (2004) provide further frameworks for interpreting the colonial and postcolonial history of SAfE. Trudgill's model could certainly be relevant for analysing SAfE, but, as it happens, Trudgill has used extensive evidence from one particular source, the ONZE-project (Origins of New Zealand English) to exemplify and validate his model. SAfE unfortunately did not develop under the same kinds of circumstances as New Zealand English (NZE) and the other SHEs when considering the above-mentioned complexity apparent in the formation process of SAfE. This complexity mainly lies in the extended formation of the variety, as well as in its language contact situations, which will receive attention in the next section.

A brief summary of the productivity and shortcomings of Trudgill's model with regards to SAfE will be given here. The variety did enjoy an initial *tabula rasa* situation (which is a key ingredient in the model) with the arrival of the 1820 Settlers; yet the fact that two further such situations followed this initial settlement (cf. Bekker, 2012) renders SAfE unsuited to the model of new dialect formation (2004), as indeed also noted in § 2.3.3.1. Furthermore, the role of the Founder Effect is impeded by its extended formation process. Therefore, SAfE enjoyed a prolonged, varied and rather erratic process of input unlike any other SHE: firstly mainly from London and southern England for the stigmatised Cape dialect, secondly mainly from northern England for the more prestigious Natal dialect, and thirdly from both these dialects, as well as from many different British dialects for the new mixtures arising at Johannesburg, as is also mentioned earlier. Furthermore, Lanham and MacDonald (1979:82) note the high proportion of British-born South Africans even into the second half of the 20th century, suggesting that the revitalization of British input continued well after the British influx during the mining revolution.

Even though Trudgill (2004:22;24) does recognise these "three distinct periods of immigration" for SAfE, he suggests that they only instituted "a small amount of regional variation" and maintains that SAfE was formed between 1820 and 1870, which Bekker (2012) strongly contests. Schneider (2007), although emphasising different factors than Trudgill, argues along a similar timeline to

Trudgill's notion of the 1820 to 1870 formation period, and this is indeed the traditional view of the formation of this variety. One particular factor on which Schneider and Trudgill differ is language contact. Trudgill's choice to exclude language contact among the pivotal factors in the formation of SHEs renders it particularly unproductive to analyse SAfE along the rigid lines of his deterministic model, as the next section will also reveal. Therefore, the usefulness of Trudgill's model toward SAfE lies in its emphasis on the effects of input and dialect mixture, as these were both prominent factors in the formation of the variety. However, it is my opinion that the evidence from the ONZE-project is not necessarily productive for drawing conclusions about all SHEs, and especially for SAfE.

Schneider (2007), on the other hand, analyses SAfE according to the five phases in identity construction he proposes in his dynamic model (Schneider, 2003). A brief account of his application will be given here. The first phase in the model is 'foundation' – which indeed marks the first step in the formation of SAfE (1806-1822/1870s). Schneider asserts that during this phase the STL-strand, which comprised of the Natal and the Eastern Cape varieties, came into contact with the IDG-strand, which included both the Afrikaners[44] (due to their already 'indigenized' identities or at least not a British identity, although European) and the native Bantu and San populations (Schneider, 2007:175-176). Koinéization took place among the British settlers (2007:177), but the division between Natal and the Cape still remained prominent. Within the STL-strand some place names and terms for landscape features entered their vocabulary, as the model suggests would happen during this phase (2007:178).

To evaluate Schneider's classifications, the Afrikaners, which are grouped under the IDG-strand (2007:176), cannot truly be regarded as part of the same IDG-strand as the African population. Indeed Afrikaners neither felt themselves to be part of the native populations' identity, nor of an exclusively European one – they had their own hybrid sense of social selfhood, as sections to come will also discuss (cf. Raidt, 1991:247). It is therefore plausible that the Afrikaners form a unique, supplemental third strand to the STL- and IDG-strands in the case of SAfE – one that is not present in any other native variety of English[45]. Schneider (2007:58) notes that,

[44] The Afrikaans-speaking population.

[45] Except perhaps in Irish and Canadian English, as also noted in Footnote 8 in § 1.1.3. However, Irish Gaelic has been rather obsolete since the 20th century, while the Quebecoise (French Canadians) are

"(w)ith the notable exceptions of the Boers[46] in South Africa and the Acadiens in Canada, we can largely disregard non-indigenous groups who arrived earlier than the British", which suggests that the Afrikaners do not exactly fit into the IDG strand. Schneider goes on to describe a third strand, namely the ADS-strand, or 'adstrate' (2007:58): "a linguistic input that enriches and expands an existing contact scenario". This strand comprises "large population groups migrating to a country where the English-speaking population had already established itself, usually alongside the IDG strand" (2007:58). But, as § 2.3.3.4 will show, the Afrikaners arrived long before any kind of English-speaking population did (as Schneider [2007:58] also notes). Apart from his above mention of the Boers, Schneider does not explicitly classify the Afrikaners as an ADS-strand in South Africa, but still regards this nation as a separate 'stream' in the IDG-strand; therefore, Afrikaners are part of the IDG-strand, but not on the same level as the indigenous African population (2007:178). This issue will be further addressed in § 2.3.3.4.2. However, whether the Afrikaners are part of the IDG-strand or not, it remains important to note that they enjoyed contact with the English-speaking population since the very beginning.

The second phase in the model is 'exonormative stabilization' (1822/1870s-1910) wherein Schneider deems the Anglicization of the Cape and Natal as promoting the establishment of English as a regularly spoken language (Schneider, 2007:178). The STL-strand maintained their British identity and a swift rise of bilingualism took place within the IDG-strand with the need for social contact on the mines of Transvaal (2007:179-180). Natal English however retained its status as the 'respectable accent' and its speakers therefore inclined toward British norms and a very strong British identity (cf. Lanham & MacDonald, 1979), but loanwords for fauna and flora did enter the vocabulary (Schneider, 2007:180;186).

I generally agree with this application in terms of the two settler dialects, but suspect that the influx of new immigrants into the Witwatersrand (Johannesburg area) is an important, if not the most important part of the 'foundation phase' (cf. Bekker, 2012; 2013), which means that SAfE might not have enjoyed the apparent stabilisation gained in Phase Two at this stage in time. Schneider's neglect to consider

---

quite isolated from Anglophone Canadians. SAfE can therefore be said to represent a more extensive contact situation with the Afrikaans-speaking population. This is indeed mirrored in Afrikaans-English bilingualism on both sides (Coetzee-Van Rooy, 2013:179), even if Afrikaans users are generally more eager to switch to their L2 than the users of SAfE (Hauptfleisch, 1979:42-43; see also Deumert, 2005).

[46] Afrikaners.

the input of a New English like SAfE more carefully has therefore influenced his application in this case. British identity among the Natal dialect (cf. Lanham, 1996:23) strengthens the argument for SAfE moving to Phase Two at this point in time, but with the mixture of an assortment of dialects arising from the three stages of migration, it is certainly not enough to assume that SAfE has moved into this phase based on Natal English alone. The issue of bilingualism that Schneider addresses here is indeed very important, as it underscores language contact situations between the different strands (including Afrikaans, whether it be part of a special third strand or not).

The third phase is perhaps most worthy of attention: 'nativization' (1910-1994). It is within this phase that the crucial identity shift necessary for the formation of a 'new English' occurs (Schneider, 2003:247). Schneider alleges that political stabilisation had been gained through the Union formation in 1910[47], because it was a step towards a "self-contained state" (Schneider, 2007:181). However, the Union still supposed British rule (even though it was administrated internally, as Schneider concedes) – therefore striving towards British identity for reasons of 'status' – leaving the Afrikaners feeling hostile toward the British (cf. Raidt, 1991:247). On top of this, the continuous influx of immigrants, especially capitalists investing in the mining and other emerging industries, further complicates the notion of such a self-contained state (cf. Branford, 1996; Lanham, 1996). Likewise, when the Afrikaner-based National Party came into rule in 1948 and later established the fully independent South African Republic in 1961, political stabilisation (supposed by independence) did not enter the social scenario either, as Schneider assumes (Schneider, 2007:182). But Schneider (2003:287-8) does state that nativisation can still take place even if the political assimilation in a country is not without struggle.

Yet, in my view the establishment of apartheid by the National Party led to the division of identities in the country, and not to the cultural assimilation and reduced gap between the STL- and IDG-strands that the phase of nativisation would require (Schneider, 2003:247-248) (see also Van Rooy, 2014). Large-scale second language acquisition of English was indeed eminent during this period of apartheid, but it was compartmentalized and therefore could never promote the supposed cultural or

[47] The history of South Africa regarding the eras of Union, Republic, apartheid and democracy will be considered in more detail in the next section, where SAfE contact with Afrikaans is described.

linguistic assimilation the model would require. Although the influence of Britain was reduced, the segregational system led to an increase of social disunity – leading to linguistic disunity in the broader community of English speakers in the country. A complaint tradition did arise because of this (see e.g. Beeton & Dorner, 1975), expressing dissatisfaction about the competence of, for example, the growing BSAfE-group (Schneider, 2007:183), but as to the structural nativisation of SAfE emerging (developing shared linguistic features), there is no certainty.

Schneider (2007:183) admits that these processes needed to take place within every set of closed linguistic communities (i.e. each SAfE sub-variety), but it remains doubtful to suppose broad structural assimilation took place at all – the history of apartheid and the clear differentiation between the subsets indeed suggest otherwise. It could be true, however, that the native variety, SAfE, has undergone nativisation during this stage, but the "[w]hite dialects of English [remained] most obviously socially stratified", despite the emergence of a standard based on Natal English (2007:183), so the influence of British norms thus continued to be prominent, which would of course slow down the nativisation process. Nevertheless, there is justifiable phonological, lexical and grammatical evidence that SAfE has undergone or is at least already undergoing structural nativisation (2007:184).

Schneider moves on to describe SAfE in terms of Phase Four, 'endonormative stabilization' (Schneider 2003:249), as he currently believes English in the country to be in this phase of linguistic evolution (1994-), with the final phase of 'differentiation' lurking in the distant future. He perceives the rise of democracy under the flag of the ANC (African National Congress) in 1994 as Phase Four's important 'Event X' (a prominent political event leading to large-scale identity revision), as the idea of a 'rainbow nation' redirected the previously divided social identities of the population to a collective identity, with BSAfE as the symbol of that identity (being a "powerful national unifier") (Schneider, 2007:188). SAfE, in terms of the native speaker variety, is now marginalized or becoming assimilated into BSAfE, according to this view.

Schneider thus predicts a movement into homogeneity in the near future of English in South Africa, but the identities of South Africans are not rushing toward social unity, they are still divided (geographically and socio-politically) due to the lingering influence of apartheid, as it has only been terminated 19 years ago after (formally) being in use for almost twice as long. Its effects have not disappeared in

the country, and, regardless, the ideal of a rainbow nation does not suppose such an absolute unity as Schneider would like to believe, but 'unity in diversity' as the national motto states. The entire nation do not perceive BSAfE as the national norm, despite it being the largest sub-variety, because they are a diverse people not nearly ready to surrender the expression of their identities. This view is indeed aligned with that of Van Rooy (2010:10-11), who argues:

> "...in South Africa, the STL strand is outnumbered by the IDG strand, and ...we are once again in a tabula rasa kind of context in that various groups of English users have to accommodate. The political power shift has removed the reference status of the STL strand in the community... To the extent that any convergence is observed, Mesthrie (2009) finds initial indications of IDG speakers at the upper end of the socio-economic spectrum to approximate the norms of the STL group, but little movement in the STL group to incorporate variants characteristic of the IDG group."

Although Schneider's model emphasises identity and language contact, which are important factors to consider in the formation of SAfE, and is certainly suited to other PCEs, it is perhaps premature to put English in South Africa in Phase Four. There is some linguistic evidence for SAfE perhaps moving into Phase Four at least, but some other applications of social factors to the formation of the variety remain debatable.

In contrast to these traditional views on the formation of the native South African variety, Bekker (2012; 2013) proposes the Three-Stage Koinéization Model of the Formation of SAfE, based on phonological evidence. However, being primarily interested in the (late) 19th century, Bekker (2012) does not offer a direct solution to contemporary debates surrounding the current state of the variety. His main concern (which will no doubt help to gain new perspectives on the later development of the variety) is that three different *tabula rasa* situations arose from the three different strands of migration (impeding the Founder Effect, seeing that the numbers of the third migration outnumbered those of the first two[48]), as also mentioned in earlier sections. Bekker's emphasis is on input and dialect mixture (like Trudgill), as well as language contact (like Schneider); although, unlike Trudgill and Schneider, he refers to the three strands of SAfE varieties of CulSAfE, GenSAfE and BrSAfE (as mentioned in the Introduction [§ 1.2.1] of this thesis) as the products of the formation

[48] What is implied by Bekker (2012) is however not the notion of 'swamping' due to the three waves of immigration per se, but that because there were three successive (geographically separated) koinéization moments, the outcome of one case of the Founder Effect was the input into another.

processes, instead of one unified variety. The model, depicting the three stages and their sources of input, is presented below (Bekker, 2012:136):

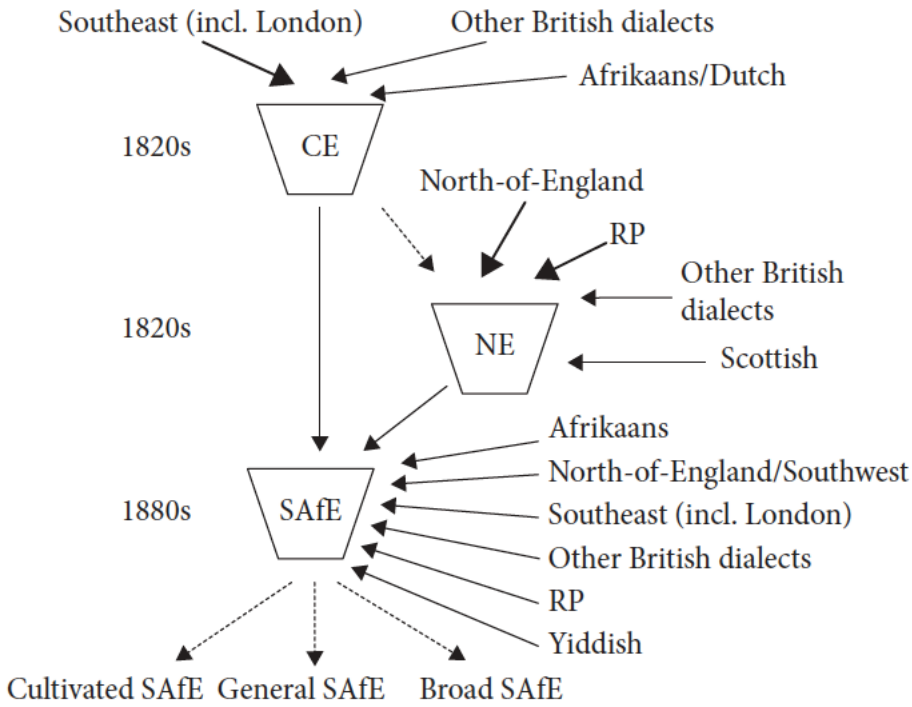

Figure 2

The Three-Stage Koinéization Model of the Formation of SAfE

This is indeed the only model that considers the extended formation period of SAfE and deems the third stage of koinéization in the late 19th century (and the role of Johannesburg) essential for the understanding of the variety's development. Bekker (2012:135) summarises his argument as follows:

> "What is being argued for is the merits of accepting that the late-19th-century birth of Johannesburg provided an appropriate context for new-dialect formation and that a full understanding of the formation (and character) of SAfE cannot be gained (particularly *vis-à-vis* the other Southern Hemisphere varieties) if this third stage of koinéization is not taken into account."

Bekker (2013:6) writes that the general view held by "commentators" on SAfE has up until now been informed by the traditional version or "Standard Model" of SAfE formation offered by e.g. Lanham and MacDonald (1979), which dismissed the

impact of the Johannesburg settlement entirely. Bekker describes the traditional view on GenSAfE (the local standard variety):

> "According to Lanham and Macdonald (1979), this sociolect [GenSAfE] is, roughly speaking, NE absorbed into Johannesburg and reanalysed as a sociolect, a position which implicitly rejects the tabula rasa status of Johannesburg and which emphasizes the role that prestige and identity played during this stage in the history of SAfE. According to this 'picture', Johannesburg simply 're-allocated' the two original colonial (regional) lects to two of the socioeconomic classes (lower and middle) of early Johannesburg society, effectively ignoring the other varieties of English brought in during this period of immigration." (Bekker, 2013:6)

Hence, Bekker's model reaches beyond the timeline of these kinds of traditional views, that proposes the formation period of SAfE to be 1820 to 1870/the 1870s, and adds another decade to the picture. Indeed he strongly argues that the 1880s were the defining period for SAfE formation, rendering the other two dialects inferior in influence. In essence, the traditional views overlooked that the entire English landscape was turned upside down by the third wave of settlement and overstated the strong role of the NE and CE settlements in this regard:

> "...I would like to reconfirm that the hypothesis put forward in this paper is meant to explicitly reject the notion that SAfE is anything other than a late 19th-century variety. If we take the koinéization process in Johannesburg seriously ...then SAfE must be defined in terms of this third and last koinéization process. While, in terms of this model, CE and NE were certainly *inputs* into SAfE, they did not define it and were almost certainly not majority inputs either." (Bekker, 2012:143)

Bekker develops his argument further when he applies Trudgill's (2004) notion of a 50-year focusing period to his model. When the formation of SAfE is perceived as taking place between 1875 (the date linked to the discovery of diamonds at Kimberley) and 1925, this already means that this variety formed far later than Australian English (AusE) and New Zealand English (NZE), slightly later than Falkland Islands English (FIE) and "overlapping to some degree with" Tristan da Cunha English (TdCE) (Bekker, 2012:143). But, when dating the beginning of Trudgill's stage three (determinism in new-dialect formation) at around 1886, the year gold was discovered on the Witwatersrand, "SAfE only becomes a focused variety in approximately 1935 and should therefore perhaps be referred to as an early 20th-century variety" (Bekker, 2012:143). Bekker (2012:143) adds that "[o]n this basis, SAfE is in fact the youngest of the colonial varieties of English", especially in the

Southern Hemisphere. Based on Bekker's account, the period between the mid-1870s and late 1880s could be regarded as the actual, in a way delayed, onset for Schneider's foundation phase (Phase One), which Schneider (2007) himself regards as already beginning in 1806, as the period of input was extended almost over the entire span of the 19th century. Furthermore, the mid-1930s could be regarded as the actual onset for Phase Two ('exonormative stabilization'), which Schneider regards as starting between 1822 and the 1870s. As noted in § 2.3.3.2, the prestige associated with British norms meant that an exonormative Cultivated SAfE (CulSAfE) (or Conservative SAfE) continued to reign for the largest part of the 20th century, when considering that the speakers of CulSAfE were mostly born before 1930 (Lanham & MacDonald, 1979:30). If the youngest users of this sociolect were middle-aged by the late 1970s, this means that exonormative standards would still prevail by this period, but perhaps not long after. Bekker (2013:6) writes that CulSAfE "is hardly used among young speakers any longer" (cf. Lass, 2002:110), seeing that GenSAfE is the contemporary local standard variety. This certainly provides a new perspective on SAfE and its formation. The results of this thesis may serve as a means to 'test' Bekker's model and its possible implications in terms of grammatical evidence, on top of his phonological evidence (START-vowel backing).

In conclusion of this section, SAfE, an inner-circle variety in its purely native form, has neither developed fully along the lines of Trudgill nor Schneider's models, but many descriptions and factors proposed in these models were found to be particularly useful, especially input (Trudgill, 2004), as well as identity and language contact (Schneider, 2003; 2007). This section has therefore also shown that migration, contact and identity are the most prominent driving forces behind the 'birth' and evolution of SAfE. It is debatable whether SAfE has entered Phase Four of Schneider's model, but based on linguistic evidence this is probable. The argument of Bekker, which states that SAfE was formed between the 1880s and 1930s, has the effect that this variety had its formation period much later than the other SHEs, due to the prolonged era of migration and the resulting conditions thereof. The next section will explore one of these conditions, which is the contact situation that arose for SAfE with a related language.

2.3.3.4 Contact with Afrikaans

Despite the issues related to English-internal contact of dialects based on e.g. location and relocation of not one, but three strands of British settlers, and the complications these issues cause when it comes to the traditional linguistic models, there is another matter of contact which is essential in considering the formation and evolution of SAfE. Branford (1996:36) notes:

> "For English in South Africa ...there are two complicating factors seldom fully matched elsewhere: firstly the presence of a substantial English-speaking 'settler' population, and secondly the survival of a second colonial language, namely Afrikaans, formerly a kind of partner in dominance with English and still influential."

The impact of extensive and continual contact with Afrikaans, a West-Germanic relative of English, indeed arises as a factor less present in other inner-circle varieties – a factor that is further complicated by the presence of a singular circumstance like apartheid, as well as other sociohistorical circumstances related to migration and identity, where contact serves as the quintessential link.

Branford (1996:37) notes that contemporary SAfE reflects "a long-standing symbiosis of English with Dutch and its successor Afrikaans". This symbiosis was not always mutualistic[49], but I believe it is in the very complexity of the contact situation between these two languages that the uniqueness of SAfE lies, together with its extended formation period. This section will provide a brief historical overview of the contact situation between English and Afrikaans among the 'white' societies of South Africa.

2.3.3.4.1 Historical overview

Before the British landed at the Cape, a transplanted variety of 17th century Dutch had been the spoken in this region. This language gradually changed into a new language,

[49] Similarly, Old Norse and Norman French relations with English were also not always only amicable, yet the influence of these languages on English are undeniable, as shown in § 2.3.1.1, and further mentioned in § 2.3.3.4.2.

Afrikaans[50], during the course of the 18th and 19th century, mainly due to the interaction between Dutch colonists and their slaves (Roberge, 2002). It is this already-transplanted language that came into contact with English for the first time in 1795, when it in turn was transplanted to the region. The Cape had in fact been an area of Dutch settlement since 1652, which was governed by the Dutch East Indian Company. Well before British rule, some immigrants from Germany, France and other European countries settled amongst the Dutch (Giliomee and Mbenga, 2007:46-47) and from around 1658 to 1808 over 60 000 slaves were imported to the Cape in about equal proportions from Africa (mainly Angola), Madagascar, India and Indonesia (2007:53).

When the British annexed the Cape, the Dutch had lost their colony to imperial newcomers, which initiated an antagonistic relationship between the English and the Afrikaans communities – one that continued to mount throughout the whole of the 19th century. The tension grew due to the Anglicisation policy of Governor Lord Charles Somerset in 1822 that English would be introduced systematically as the only language of government, and the situation was further aggravated when English was proclaimed the sole language of education (Branford, 1996:38; Steyn, 1980:127-128).

As a consequence of these and many other political factors, Afrikaners experienced a sense of prolonged exclusion and oppression under British rule which antagonised political relations (Steyn, 1980; Branford, 1996:38), but Branford (1996:39) also remarks that relationships at individual level were less strained than those on government level until the third quarter of the 19th century. Bilingualism was even common among the more educated Afrikaners (as noted earlier; cf. Schneider, 2007), which points to mixed attitudes toward their British neighbours. Hooper (1951:83) mentions that inter-marriage between the two groups often took place, usually among those British who were born in South Africa. However, the Afrikaans community did not undergo a language shift to English (Steyn, 1980:130-131) and the Afrikaner kept his own identity. This was largely owing to their isolation in rural areas after the earlier Great Trek into the interior from the late 1820s to the end of the

[50] The kind of Afrikaans I refer to here is the standard dialect spoken by those of European ancestry as described above, and is to be separated from the so-called 'West Cape Afrikaans' of the coloured communities, who are mostly descendants of the slaves of the Dutch settlers (cf. Branford, 1996:38).

1830s, which had happened in reaction to British oppression in the Cape (Giliomee and Mbenga, 2007:108). On this topic, Schneider (2007:177) writes:

> "As to the English and the Dutch, it may be presumed that their shared European ancestry and complexion offered them the option of a certain degree of an "us" reading as against the Africans. Though the anglicization of many Afrikaners and the acquisition of Afrikaans by many descendants of the English indicate some cultural and linguistic assimilation, the Great Trek makes it clear that it was the Afrikaners' desire to maintain their distinctiveness and independence. So, sociolinguistically speaking, contacts between the English settlers and the Dutch were rather extensive..."

As mentioned in § 2.3.2.4, Mufwene (2008:17) emphasises the important role of contact at individual level, which, when considering Branford's (1996:39) above statement, is very instructive for SAfE: indeed contact at an individual level with speakers of Afrikaans was much more likely than contact at this level with speakers of African languages.

Political oppression and prejudice was indeed equally present from the sides of both nations – especially in the education systems of the Cape and the Transvaal. Branford (1996:38) notes:

> "The nineteenth century interactions between English and its speakers with the white Dutch-Afrikaner community are highly complex. Much of the early official history is of acts of discrimination: by English authorities against Dutch, as in the enforcement of English-medium instruction in State-aided schools at the Cape ...or by Transvaal authorities against English-medium instruction for Uitlanders[51] in state schools in the 1890s."

The late 19th-century 'Mineral Revolution' on the Witwatersrand however created an urban platform for English, Afrikaans and many indigenous African languages to make contact on a larger scale than ever before and to co-exist to some extent, as also noted in previous sections. Bekker's Three-Stage Koinéization Model of the Formation of SAfE (2012), shown in Figure 2, also explicitly illustrates the contribution of Afrikaans to the mix of influences in the formation of SAfE during the 1880s.

English, however, generally remained the most 'powerful' language in the new urban environments. The Afrikaners were initially much disadvantaged, along

[51] A popular Afrikaans term in the 19th century meaning 'foreigner', which mainly referred to the British.

with their language, as they "lacked social attributes and socio-cultural competence to negotiate desirable positions in a typically British social system" (Lanham & MacDonald, 1979:77). This points to a lack of industrial skills, as well as education, and indeed Lanham and MacDonald (1979:77) note that the majority of the around 800 children without education in the slums of Johannesburg in 1896 were "Dutch speaking". Even though the Afrikaner endorsed or tolerated the ways of the British social system in the new cities, this did not necessarily accompany a sense of allegiance with the British (Lanham & MacDonald, 1979:81; cf. Labov, 1966).

Notwithstanding, even the more friendly individual relationships far away from the new urban areas were compromised as increasing political tensions culminated in the Anglo-Boer War and its social repercussions from 1899 to 1902 (Giliomee & Mbenga, 2007:206). The Boer army was essentially vanquished, but not at all political levels, as negotiations between the two groups, which helped end the war, brought about pivotal reverberations in the country's political landscape in decades to come (2007:216). The specifications of the 'Peace of Unification' decree stipulated that the Dutch/Afrikaans community would enjoy the right to vote (2007:216), even despite the fact that English became the dominant language of education (run by the government and churches) and still did not replace Afrikaans as the home language of this population group (McCormick, 2002:50-58). This edict further entailed that whomsoever would represent the constitution in the future, would have final say about the voting rights of non-Europeans – a serious implication for future legislation in what was later to deprive the black communities of their political voice (Giliomee & Mbenga, 2007:216).

Eight years after the war ended, the unification of all territories occurred and in 1910 became known as the Union of South Africa (Giliomee & Mbenga, 2007:232). Despite public antagonism resulting from the war, Steyn (1980:178) notes that eagerness among the poorer and less-educated Boers to learn English in the immediate post-war period was induced by the fact that many had no other choice but to work for the English (cf. Branford, 1996:39). An official bilingual policy was adopted by the government – initially between English and Dutch, but the latter was formally replaced by Afrikaans in 1925 (Giliomee & Mbenga, 2007:253). In 1948 the dominant Afrikaans party, 'die Nasionale Party' (the National Party), was voted into government and the road towards formal apartheid, the exclusive political, social and regional segregation of communities based on racial groups, was paved with the aid

of legislation, as the new governing party exercised their previously ordained constitutional right to decide the fate of black voters[52] (2007:306).

The victory of the National Party was seen as a setback for the English community, and friction between the two nations remained prevalent. On this topic, Cochrane (1987:3), while discussing the role of the English-speaking churches in the South African society from 1903 to 1930, notes the following:

> "The year 1948 – the victory of Afrikaner political forces – stands as *the* symbol of disaster for English-speaking whites as a group. But it represents far less a victory of segregationalist views over liberal than a particularly acute moment in a much longer competition for political hegemony between two groups sharing a dominant position in an oppressive society."

The struggle for political hegemony, as noted above, was therefore settled with the election of 1948 – with the Afrikaner *volk* (nation) emerging as the very proud victor, and by implication, with Afrikaans as the new dominant language. But, as Cochrane (1987:3) states, the setback for the English community occurred much more on a competitive level, than on an ideological one. Given the political history from 1948 to the late 1970s, it becomes clear that, the English community was not united in its views, and the majority of the members of the United Party, mainly white speakers of English, were just as segregationalist in their views as the National Party. This became plain when a dispute about black property rights sparked the breakaway of eleven members of the United Party (out of 53) to form the Progressive Party under Dr. Jan Steytler[53], and with Helen Suzman as the party's only parliamentary representative (Giliomee & Mbenga, 2007:326).

By the time South Africa officially became an independent Republic in 1961, the legislation implicating apartheid was already firmly established (Giliomee and Mbenga, 2007:338). This caused increased contact among the general white community (Europeans by race), i.e. the English and the Afrikaans. Although their identities were still separate, as the next section will also note, their racial connection

---

[52] The resistance movements of the Black communities from the late 1940s to early 1990s, as well as the implications of apartheid on other racial groups such as the Indian and Coloured communities, although very important in the struggle for freedom and equality which was achieved in 1994 (Giliomee and Mbenga, 2007:308-407; cf. Van Rooy, in press), are not the direct foci of this thesis. The consequences of the political climate before and after 1994 are however important factors in the creation of a unique language contact situation between Afrikaans and English, so the focus is more strongly on historical events relevant to this situation.

[53] See: http://www.da.org.za/about.htm?action=view-page&category=383

had the effect that contact was inevitable, if somewhat unnatural. About 40 years of this rather forced contact situation passed before democracy was finally achieved after the first general elections in April 1994. The kind of 'racial affinitive' relationship between white English speaker and white Afrikaans speaker was of course not bluntly severed after 1994[54], and the road toward political and racial stability has rather proved to be a gradual one (2007:433-437).

One example of a 'forced' contact situation during the apartheid years is the relationship between Afrikaans and English in the context of the South African Army (late 1960s to early 1990s). The shared quarters, ideals and tasks within this context facilitated a contact situation where English-speaking men were not only obliged to learn Afrikaans as the official military language, but also to communicate with their fellow (Afrikaans) soldiers. Therefore, Afrikaans denoted group identity within the Army, but on an individual level English and Afrikaans interacted on more equal ground, creating a rather prolific contact situation.

This section has shown that the English and Afrikaans communities enjoyed a particularly complex relationship, ever since the annexation of the Cape in 1806 and the migration of the British into a country previously settled by the Dutch. As implied by Cochrane (1987) in the quote above, the political competition between these two nations has been long-standing and far-reaching, and always entwined with language (cf. Branford, 1996:37-40). The next section will explore the linguistic implications of the complex relationship between these communities and their cultural identities.

2.3.3.4.2 Identity and its linguistic implications

> "If you talk to a man in a language he understands, that goes to his head. If you talk to him in his langauge, that goes to his heart." (Nelson Mandela)[55]

Because identity is such a complex factor and sensitive subject in South Africa, as suggested by the quote above, this section will explore this issue in terms of the contact situation between Afrikaans and English in the country, which resulted from the many migrations described previously. Firstly, it will be considered whether the

---

[54] The Democratic Alliance (DA), as the opposition party to the ANC, led to a single 'home' for many English, Afrikaans, Coloured and Indian voters since the late 1990s.

[55] Quote obtained from: http://www.sowetanlive.co.za/news/mandela/2013/07/18/67-awesome-quotes-by-nelson-mandela.

Afrikaans community is in fact a unique strand – a supplement (adstrate) to Schneider's (2003) two strands – in the New Englishes of the world, and if so, what the resulting linguistic implications for SAfE would be. Secondly, the issue of whether a language like Afrikaans could have actually had an effect on SAfE under the relevant circumstances will be addressed.

Politics arises as the most prominent social factor in the case of SAfE and Afrikaans identity; indeed Coetzee-Van Rooy and Van Rooy (2005:4) note that identity and language are highly politicised concepts in South Africa, since the earliest times of colonialism until today. This is mainly due to the 'acts of discrimination' described earlier by Branford (1996:38), which in fact stretches far beyond the circumstances of the Afrikaans and English communities in the 19th century, well into the late 20th century regarding the ideology of apartheid and its consequences on all of the country's inhabitants.

Within this context, the identity of the Afrikaner can be seen as something entirely unique to South Africa, and even though Afrikaans was only officially recognised as a language in May 1925 (Giliomee & Mbenga, 2007:253), an entirely *sui generis* and decidedly Afrikaans identity had been firmly established and distanced from its Dutch ties long before this. Edward Sapir's (1929:209) statement that "no two languages are ever sufficiently similar to be considered as representing the same social reality" is certainly true for Dutch and its descendant, Afrikaans. Indeed Branford (1996:39) notes that it is already proclaimed in the 1875 manifesto of *Die Genootskap vir Regte Afrikaanders* (The Fellowship of True Afrikaners) that it stands for *ons Taal, ons Nasie en ons Land* (our language, our nation and our country – note the order of mention). In this manifesto both *Afrikaanders met Hollandse harte* (Afrikaners with Dutch hearts) and *Afrikaanders met Engelse harte* (Afrikaners with English hearts) are condemned, and it is proclaimed that "*regte* (true) *Afrikaanders* are only those *met Afrikaanse harte* (with Afrikaans hearts)" (Branford, 1996:39). Note that, although *Afrikaanse* and *Afrikaanders* would translate into *Afrikaans* and *Afrikaans-people* from 19th-century South African Dutch to English, in this context they rather denote the meanings *African* and *Africans.* These Dutch terms do however not imply any shared identity with the native African population, even though Afrikaans is regarded as an "African language" with no other home than in Africa (Raidt, 1991:256). These terms therefore rather denote a strong connection to the place, Africa, than to its indigenous population. Furthermore, the Dutch/Afrikaans

community did not regard themselves as European colonists living away from home either, but as a distinct group of 'Africans'.

This is in accord with Schneider's (2007) suggestion that the Afrikaner is not part of the STL-strand (which comprises the British settlers), but despite their 'indigenized' identity in Africa, it does also not mean that they are to be considered part of the same IDG-strand as the many native African nations, but rather as a separate IDG-'stream', as Schneider implies (2007:178). Certainly the systematic implementation of the segregational laws against the 'non-white' population of the country by the Afrikaner regime during the mid-20th century is enough evidence of this (cf. Giliomee & Mbenga, 2007:338-345). The 'true' Afrikaner therefore had neither a 'white European' (i.e. Dutch or British) nor a 'black African' identity, but rather a 'white African' sense of self – something entirely different from anything else in the English-speaking world (cf. Raidt, 1991:256).

A few questions now arise from the above discussion. Is it possible that a stigmatised, rival language like Afrikaans could have had an effect on the SAfE in the South African context – especially during the earlier times of contact under circumstances of discrimination, exclusion and conflict, i.e. the 19th and early 20th centuries? Could this influence extend beyond that of e.g. loanwords into a domain like grammar? And is it possible that the unnatural contact situation between the two major 'white' ethnic groups of South Africa, with Afrikaans as a symbol of the bygone apartheid government, could have increased Afrikaans influence on SAfE, if such an influence exists? And finally, are there traces of such an influence in contemporary SAfE, which could support e.g. Jeffery & Van Rooy's (2004) findings, as mentioned in Chapter 1? I will attempt to answer these questions in the concluding chapter of this thesis, after the relevant results from the SAfE and Afrikaans data regarding modality are considered and compared. As for the initial question, Mesthrie and West (1995) offer some insights related to the syntax of the first input variety of SAfE – the English of the 1820 Settlers.

Mesthrie and West (1995:115) opt for a somewhat different view on the issue of Afrikaans influence on SAfE than some of the earlier studies mentioned in Chapter 1. They argue that, for the input variety, Afrikaans is not "the prime mover behind the distinctive features of SAfE syntax", and rather a "historical continuation of early 19th century input from Britain" (1995:110;115), but neither do they deny the influence of

Afrikaans altogether. Relevant to this, Lanham (1978:157-8) notes the following in terms of phonology and the imitation of phonological variables:

> "The advance of the 19th century saw Cape Settler society becoming progressively less a product of its history and more of its environment. Inter-marriage and social mingling with the Afrikaner community did not efface the English identity ...but contributed towards modifying their social values towards frontier values. ... The social forces behind such imitation was the 'accommodation' of, for example, a Settler son to the Afrikaans-English of an Afrikaner wife. ... Through more extensive contact ...Settler descendant and Afrikaner had grown to resemble one another to an extent that, ironically, they shared the same social image in the early days of the mining cities."

Here Lanham (1978:156) argues that the variety of Extreme SAfE[56] originates primarily from the vernacular developed by the many young children of the 1820 Settlers, and the first generation born in South Africa, due to their accommodation to L2 speakers (mother tongue Afrikaans speakers). This kind of accommodation is also classified by Mesthrie and West (1995:115) as the second of three types of influence of Afrikaans on SAfE. They list the three types as follows:

> (i) Transfer from Afrikaans evident in the L2 English of Afrikaners. ...
>
> (ii) Accommodation of English speakers to Afrikaans L2 English...
>
> (iii) A third, infrequently cited possibility is that of the L1 English of people of Afrikaans background – i.e. among families which underwent language shift. This is a plausible source for the passing of Afrikaans features into the L1 English of some people in the Cape, which is seldom given its due in the literature.

This basically entails that there were three kinds of linguistic relationships between SAfE and Afrikaans due to contact – for the first wave of settlers at least. The last sentence of point (iii) is striking: it accounts for the possibility that Afrikaans features may have entered into L1 SAfE in terms of Afrikaans families that underwent language shift and afterwards had further contact with other L1 speakers. Mesthrie is certainly right that this issue of Afrikaans influence is underrepresented in literature, but support for this kind of influence is found in the results of this thesis (see § 4.2.2.3 and § 4.3.2.2.3). It is interesting to note that language contact has already been part of the very first two generations of koinéisation, which complicates Trudgill's idea of a *tabula rasa* context.

[56] Bekker (2009; 2012), for example, prefers the term 'Broad SAfE', as discussed in § 1.2.1.

It is true that Mesthrie's study only relates to the initial wave of settlers, and of course the influence that the late 19th century migrations to the northern cities could have had in terms of Afrikaans and SAfE contact are not discussed in literature, except in terms of the later teaching environment (see the discussion of Hooper [1951] below). Lanham's (1978) above account regarding the "early days of the mining cities" is one of social cohesion, which supports Branford's (1996:39) notion that individual relationships were less strained, but it would be misleading to suppose a close sense of collective solidarity between these nations at this stage, seeing that the Anglo-Boer War broke out barely a decade after these "early days".

Yet it is not entirely impossible that even during a conflict situation like war linguistic influence from the opponent's side could not seep through to the other party, even if unnoticed, as is evident in the historical state of affairs between Old Norse and Old English during the Viking invasions, when OE was indeed influenced by a related Germanic language, as discussed in § 2.3.1.1. Whether a similar situation existed for English contact with Afrikaans is still unexplored. Clearly much insight is still to be gained regarding the linguistic implications of the complex contact situation between two very (proudly) distinct, but not wholly unrelated or unconnected languages and identities.

In the early 20th century, and especially after the political change brought about by the victory of the National Party into the mid-20th century (see § 2.3.3.4.1), there was "an increasingly strong influence of Afrikaans upon English", largely due to the growing urbanisation of the Afrikaners, who more often tended to live in the country (Hooper, 1951:81). This promoted bilingualism, especially among the English-speaking children in the smaller towns and among the Afrikaans children in the larger towns (1951:81). By implication, Afrikaans influence on English was in turn promoted by the teaching situation that arose due to various social factors, as Hooper (1951:81) describes:

> "In the country and smaller towns the teachers are almost all Afrikaans-speaking. And since the teaching profession does not carry the prestige among the English-speaking section of the people that it still does among the Afrikaans-speaking, chiefly because it is as poorly paid in South African as elsewhere, it attracts very few English-speaking men and not enough women. This means that even in some of the larger towns too the class teachers in English-medium schools, and even the teachers of English to English-speaking pupils, are now not English-speaking themselves, i.e. are people whose first or home language is not English."

Hooper (1951:82) adds that although such teachers were often very fluent in English, the influence of Afrikaans on the vocabulary, pronunciation and idiom of "the English of those born and brought up in South Africa" was unavoidable.

In conclusion of this subsection, it is therefore a fact that the Afrikaans and English language came into contact with each other in South Africa, but whether the complex contact situation between identities would permit the influence of Afrikaans on English reaching beyond that of lexis is yet to be thoroughly explored. An inquiry into this could reveal whether indigenized forms have entered the English and therefore induced nativisation to take place. Schneider (2003:249) states that the early stages of the indigenization of language structure mainly occurs on a lexicogrammatical level, where high-frequency items "adopt characteristic but marked usage" – "[g]rammatical features of New Englishes emerge when idiosyncrasies of usage develop into indigenous and innovative patterns and rules". It remains to be seen if and when idiosyncrasies of use have entered SAfE in terms of the modal verb class, which could contribute to establish whether nativisation has taken place.

The complexity of identity construction in this contact situation is indeed unique in the Englishes of the world, and it is therefore very likely that unique linguistic features will arise as an effect thereof, but substantial evidence of this is still to be provided. Chapters four and five of this study will elaborate on this. The complexity of the sociohistorical factors related to the formation and development of SAfE, as evident from § 2.3, is owing to a set of unique situations related to migration, contact and identity. The three strands of migration for the British and the foregoing migration of the Dutch, as well as subsequent movements by the Afrikaners, gave rise to a highly complex contact situation on many fronts (in peace or war), which in turn effectuated an influence on identity construction of these two nations and their languages. Having explored the effects of these three interconnected sociohistorical factors, migration, contact and identity in the language change of English since its beginnings to its colonial expansion over the world and finally in South Africa, the next section will turn the attention to the modal system, which is to be the linguistic object of investigation with which to unlock the diachronic and synchronic position of SAfE, which cannot be separated from the context described above.

2.4 MODALITY AND THE MODAL SYSTEM

A host of modal expressions exist in contemporary English, both verbal and periphrastic (Facchinetti *et al*., 2003: viii) – the former being the focus of this study. The verbal expression of modality is conducted by means of two specialised kinds of auxiliaries: the modal (single-word) and quasi-modal (multi-word) verbs (Biber *et al.*, 1999:483-4). The modal auxiliaries, also called modal verbs, are one of the three major verb classes, the other two being lexical verbs (e.g. *run* and *eat*) and primary verbs (*be, have* and *do*) (Biber *et al.*, 1999:358). A short history of the English modal system will be presented in this section, to provide a basis for understanding both the system itself, in terms of form (grammar) and function (semantics), as well as its general development through the process of grammaticalisation[57] in pre-colonial English and the native Englishes of the post-colonial world, before moving to describing its journey within South Africa.

After the nature of modality is described in terms of semantics and the accompanying terminology, the evolution of the English modal and quasi-modal verbs across the centuries will be recounted, in order to illustrate the originally polysemous nature of these auxiliaries in expressing meanings of e.g. permission, possibility, ability, obligation, necessity, volition and prediction (Biber *et al.*, 1999:485). The modal verbs studied in this thesis include *can, could, may, might, must, need, ought, shall, should, will* and *would*, whereas the quasi-modals include *BE able to, BE going to, (HAVE) got to, HAVE to, NEED to, BE supposed to*, *BE to, (had) better* and *WANT to* (cf. Biber *et al.*, 1999; Van der Auwera *et al.*, 2012; Collins, 2009a; Mair & Leech, 2006).

An account of the more recent trends of modality (regarding frequency of use and semantic changes) in American English (AmE), British English (BrE) and Australian English (AusE) will then be explored, with the aim of setting the stage for a comparison with SAfE. Lastly, I will give a description of the circumstances under which the modal system arrived in South Africa via migration in terms of SAfE and Afrikaans, in order to provide the grammatical background for the investigative methods in this study, which will be described in Chapter 3.

[57] See § 2.2.1.1.1.

2.4.1 The English modal system: description and development

Both a synchronic and diachronic perspective on English modality will be maintained in this subsection. A mostly synchronic view will be adopted in the description of the various approaches, terms, characteristics and meanings related to the modal system, with some references to diachronic changes where relevant. The diachronic discussion will however focus on the evolution or development of the modal and quasi-modal subsystems through e.g. grammaticalisation, as well as the semantic developments in the deontic domain, specifically regarding certain modals and quasi-modals.

2.4.1.1 Description: approaches, terminology, grammatical characteristics and semantics

This section will define and describe modality and the modal system in general terms, as well as clarify terminology issues in literature. I will primarily consider the descriptions, classifications, approaches and terminology of Coates (1983), Bybee & Fleishman (1995), Van der Auwera and Plungian (1998), Biber *et al.* (1999), Krug (2000), Palmer (2001), Huddleston (2002), Facchinetti *et al*. (2003), Traugott (2006), Collins (2009a) and Leech *et al.* (2009). As a means to approach modal semantics, prototype theory (e.g. Taylor, 1995) is selected as a useful way to deal with the gradience of modal meanings, after fuzzy set theory (e.g. Coates, 1983) is also discussed. But first the general nature of the modal system will receive attention.

Modality is usually categorised together with tense and aspect as the general kinds of variation found in the structure of the verb phrase (Biber *et al.*, 1999). Biber *et al.* (1999:483) note that finite "English verb phrases can be marked for either tense or modality, but not both". Beyond lexical means, such as stance[58] adverbs, English modality is expressed by two grammatical systems: mood, which is inflectional, and modal auxiliaries or the modal system, which is non-inflectional or analytic (Palmer, 2001:4; Depraetere & Reed, 2006). The broad study of modality has been channelled through various theoretical strategies, where the most notable distinction is between formal and functional approaches (Traugott, 2006:108).

[58] Stance adverbs are, for instance, *probably* (epistemic stance), *surprisingly* (attitude), *honestly* (style) (Biber *et al*., 1999:557-558). Halliday and Matthiessen (2004:125-132) treat a corresponding system of interpersonal adjuncts, which they divide into modal and comment adjuncts.

Formal approaches are assumed by e.g. Chomsky (1957), Visser (1963-73), Lightfoot (1979; 1991), Van Kemenade (1992) and Warner (1993), as well as Roberts and Roussou (2003). These studies tend to focus on competence or differences in individual grammars (Traugott, 2006:108). The formal approach further emphasises the classification of central, peripheral and semi-modals[59] according to formal properties, where e.g. core modals do not require *do* support (in negation, inversion, code and emphasis [abbreviated as *NICE*] contexts), can be negated by *not*, can appear in question tags, do not require third person singular *-s* marking and do not constitute non-finite forms (Radford, 2003:44; Depraetere & Reed, 2006:272-273). Also with an inclination towards the formal approach, Abraham and Leiss (2012), for example, prefer a typological, cognitive approach to modal categorisation.

The functional approach, on the other hand, is preferred by e.g. Plank (1984), Denison (1993), Bybee *et al.* (1994) and Hopper and Traugott (2003), and focuses on grammaticalisation and usage (cf. Croft, 2000; Traugott, 2006:108). I generally adopt a functional approach to grammar (cf. Halliday, 1994; Halliday & Matthiesen, 2004) and hence to modality in this thesis, since I aim to achieve a description of the use of modals in SAfE, and not to contribute to a formal categorisation based on superimposed qualities. Yet some formal background is also useful when studying modals, e.g. from Warner (1993) and Abraham (2001), but only to the extent that it aids, firstly, grammatical understanding in pre-analysis stages and secondly, grammatical classification in the very early stages of corpus analysis. The actual analysis is therefore functionally orientated, but is informed by formal categorisation to an extent. The next few paragraphs will be devoted to defining modality in terms of the modal system (as opposed to mood) and clarify the polysemous nature of the modal system with reference to the work of various proponents of the functional approach, with some formal considerations where relevant.

Bybee and Fleishman (1995:2) distinguish between modality as a "semantic domain pertaining to the elements of meaning that languages express", covering a broad range of semantic nuances[60], and mood as a "formally grammaticalized

---

[59] I prefer the term quasi-modal (cf. Collins, 2009a) and use the twofold system of modal and quasi-modal, rather than the triadic system mentioned here.

[60] Such nuances include the "jussive, desiderative, inattentive, hypothetical, potential, obligate, dubitative, hortatory, exclamative, etc. – whose common denominator is the addition of a supplement or overlay of meaning to the most neutral semantic value of the proposition of an utterance, namely factual and declarative." (Bybee & Fleishman, 1995:2). These meanings are described in terms of

category of the verb which has a modal function"[61] (cf. Matthews, 2007). Along the same lines, Huddleston notes that "mood is a category of grammar" and "modality is a category of meaning" (Huddleston, 2002:172). The modal system, as part of the domain of modality, is essentially a category of grammar that expresses modal meaning, with the modal auxiliary verbs (specialised auxiliaries), e.g. *can*, *will* and *may*, as the carriers of such meaning (cf. Biber *et al.*, 1999:483-502). A modal function (or meaning) is therefore carried in the form (or grammatical category) of modal verbs. In language in general, the functional or semantic domain of modality is vastly broad and therefore "lends itself best to investigation in social, interactive contexts" (Bybee & Fleishman, 1995:3). Additionally, each language differs in its "mapping of the relevant semantic content onto linguistic form", which means that languages have many different means by which to communicate various modal meanings (1995:3). This however does not mean that a specific language necessarily houses set forms to express an unambiguous kind of modality. Indeed Matthews (2007) suggests that even though modality may be a notional term, it certainly is not univocal. Abraham and Leiss (2012:1) note that "[m]odality can be characterized as the most complex functional category of all linguistic categories known to the human species".

Along these lines, for example Coates (1983), Bybee *et al.* (1991; 1994), Palmer (2001 [1986]), Nuyts (2006) and Van der Auwera and Aguilar (in press) emphasise the difficulty of defining modality as such. Bybee *et al*. (1994:176, emphasis original) argue as follows:

> "**Mood** and **modality** are not so easily defined as tense and aspect. A definition often proposed is that modality is the grammaticization of speakers' (subjective) attitudes and opinions... Recent cross-linguistic works on mood and modality, ...however, show that modality notions range far beyond what is included in this definition. In fact, it may be impossible to come up with a succinct characterization of the notional domain of modality and the part of it that is expressed grammatically."

---

language in general and are not confined to English in this context. The meanings I favour for English in this thesis are mentioned in the introductory paragraphs of § 2.4, as well as in Table 3 in the methodology chapter.

[61] Mood is expressed inflectionally, usually in sets of verbal paradigms (Bybee & Fleischman, 1995:2), e.g. the subjunctive in English (which is "confined to 1st/3rd person singular *were*" and usually replaceable by *was*) (Huddleston, 2002:172).

Similarly, Nuyts (2006:1-2) writes:

> "…modality turns out to be very hard to delineate in simple, positive terms. … There is …no unanimity among scholars as to how the set of modal categories should be characterized, either in terms of its outer borders – i.e. which semantic notions or dimensions do and which do not belong to it… – or in terms of its internal organization – i.e., how the field should be divided up in distinct categories and what their precise boundaries are."

It is clear from these quotations from Bybee *et al*. (1994) and Nuyts (2006) that mood and modality as semantic domains in language are, very hard, and, quite possibly, even impossible to define in typological or cross-linguistic terms. Van der Auwera and Aguilar (in press) note that the difficulty of defining mood and modality has influenced the degree of surety scholars could achieve regarding the fundamental notions of these complex fields "from Antiquity to the modern age". They state:

> "There have been thousands of grammatical discussions of mood and modality and though some linguists are self-assured and clear, the field as a whole cannot be said to have come to grips with these notions. This avowal has a negative and a positive side. The negative side is obvious: it is sad that after more than 2000 years our discipline has not reached a better understanding of what is fundamental to mood and modality. On the positive side, one gets the feeling that the subject matter of mood and modality is a fascinatingly difficult one and also that one can still learn from past scholarship." (Van der Auwera & Aguilar, in press:1)

Notwithstanding, Palmer, in his own search to define the semantics of modality, emphasises the negative side of things:

> "The first task is to identify the relevant area of meaning; this is not easy in the case of modality. … To begin with, the definitions are, in practice, vague and difficult to apply with any degree of precision, and do not lead to clearly distinct categories. The real problem with modality, moreover, is not just that there is great variation in meaning across languages, but that there is no clear basic feature. … Where precise criteria are given, moreover, the precision may be more apparent than real, and may involve special pleading or a degree of circularity." (Palmer, 2001 [original edition, 1986:4])

A typological definition of modality, its grammar and semantics indeed appears to be problematic to procure, mainly because "the notion of modality …belongs at a higher level of abstraction ...than categories such as [tense] and ...aspect" (Nuyts, 2006:1), and, more generally speaking, because of the natural degree of indeterminacy in language (cf. Coates, 1983). This kind of 'vague' theoretical semantics potentially

causes methodological difficulty for corpus analysis. Coates (1983:9) summarises the problem as follows:

> "...language is not an orderly phenomenon, and, as far as meaning is concerned, indeterminacy seems to be a feature of all natural languages. But it is one thing to recognise the existence of indeterminacy, and another to deal with it adequately. Semantic analysis conventionally consists in distinguishing one meaning from another, that is, in recognizing discrete categories, yet the acknowledgement of indeterminacy explicitly denies the existence of such discrete categories".

Indeterminacy is therefore a hazard in semantic analysis, yet the presence of indeterminacy, despite seeming a somewhat fatalistic term, should not suggest that a semantic system like modality is impossible to analyse in relatively precise terms. The modal system is prone to polysemy and gradience, and caution is necessary in its analysis, but it is my argument that a framework of categories can still be drawn within a polysemous system, even if the boundaries of this framework can sometimes vary in their degree of discreteness, as will be explained in paragraphs to come.

Nevertheless, in the case of English, some scholars have adopted a monosemantic approach to modality, such as Joos (1964) and Ehrman (1966), who assign a 'basic' meaning for each modal verb. Seeing that such a meaning would have to be "related to all uses of a modal, the monosemantic approach implicitly dissociates itself from a strict categorical model" (Coates, 1983:9). Similar to this approach, Johnson-Laird (1978:25) suggests that the meanings of modal auxiliaries are unambiguous, and multiple-interpretation possibilities are wholly dependent on the complements of the modals. Most scholars of English modality however adopt a polysemantic approach, such as Krug (2000), Palmer (2001), Leech *et al.* (2009) and Collins (2009a; 2009b), to name a few, who generally prefer to deal with (discrete) semantic categories, rather than a one-to-one modal-to-meaning categorisation. It is within the polysemantic approach that issues of gradience, ambiguity and indeterminacy arise.

Michael Halliday explores the functional diversity of language at the hand of modality, which includes notions of ambiguity (Halliday, 1970:326-8). In his systemic functional approach to language,[62] Halliday (1970:347-8) distinguishes between two parts of a unified syntactic system, namely modality and modulation, as

[62] cf. e.g. Halliday (1994) and Halliday & Matthiesen (2004).

forming part of different functions in language – respectively the interpersonal (social) and the ideational (content) function[63]. He maintains that the two syntactic systems are "formally identical" up to a certain (delicate) point (1970:350), only differing in the way that modality has no tense, but "may combine with any of the tenses of the verb" (e.g. *Smith can't be so busy / Smith can't have been so busy*[64]), and modulation has a fixed set of tenses it combines with (e.g. *You must build a gazebo*[65]) (1970:338). In a certain sense these concepts correspond with the idea of distinguishing between epistemic and deontic meanings, which is described later in this section. Halliday maintains:

> "Modality and modulation are the same system in different functions...: the one is interpersonal, the other ideational. In both cases we have to do with some kind of qualification of the process expressed in the clause, or rather of the complex of 'process + participant'; either this qualification resides in the speaker's own mind, or it resides in the circumstances. If the former, then it is interpersonal in function: it relates to the speakers' own communication role. … If the latter then it is ideational function: it relates to a particular part of the content of the clause." (Halliday, 1970:347)

Despite modality usually adhering to the interpersonal function and modulation to the ideational, Halliday postulates that there can also be an overlap between the two systems and their functions (1970:347). This means that the same expressions of e.g. modality can have both interpersonal and ideational content, depending on the angle from which the interaction between the filtered judgments of the speaker (external) and the characterisation of the participant (internal) are perceived. He defines these concepts as follows, and provides some illustrative examples:

> "Modality ...is the speaker's assessment of probability and predictability. It is external to the content, being a part of the attitude taken up by the speaker: his attitude, ...towards his own speech role as 'declarer'. It is thus clearly within the interpersonal component; but at the same time it is orientated towards the ideational, because it is an attitude towards the content that is being expressed. Modulation ...is part of the ideational content of the clause; it is a characterization of the relation of the participant to the process – his ability ...to carry it out. But while reference to ability does in fact characterize the participant in question – *Smith can swim* is a fact about Smith – reference to permission or compulsion does not. *Jones must swing* is not a

[63] Halliday (1970:326) also lists the textual (discourse) function as part of the triadic system of the interpersonal, ideational and textual function of language.

[64] Examples from Halliday (1970:337).

[65] Example from Halliday (1970:336).

> characterization of Jones' participation in the process but of someone else's judgment about Jones' participation; and that 'someone else' is, typically, the speaker. Thus the same forms can be used to express both. (Halliday, 1970:349)

Halliday (1970:349) further observes that when "the opinions a person expresses often tell us more about the speaker than about the subject he is pronouncing on it is likely to be his use of these 'quasi-modalities' that we have in mind: his *musts* and *mays* and *shouldn't*". Here the term 'quasi-modalities' refers to the quality of modality when imposed on a participant, which is essentially the modulation part (1970:338). Hence, "while modulations are incorporated ...as ideational material, they represent that part of it that is orientated towards the interpersonal – it is the content as interpreted by or filtered through the speaker that is being expressed" (Halliday, 1970:349). Because of the overlap of the two systems and their functions, Halliday (1970:350) suggests that 'semantic blends' can follow, i.e. ambiguity between modal expressions. To sum up this perspective in the simplest terms, the polysemy, ambiguity or 'blended' semantics of modality (being related to modulation) exists due to the different functions this system can fulfil in different environments and contexts, and when perceived from different angles.

The English modal system is indeed often described as a gradient notion[66] (cf. Facchinetti *et al*., 2003:viii), which is not binary like English mood, the latter being dividable into the indicative (or *realis*) and subjunctive (or *irrealis*) (Palmer, 2001:4; Portner, 2009:285). The gradience of the modal system can most clearly be perceived through the fact that one modal or quasi-modal verb may perform various functions and have different meanings within separate contexts (Palmer, 2001). The ways in which modals and quasi-modals 'behave' may therefore overlap, as is visible in the case of e.g. *will*, *shall* and *BE going to,* which can all either denote a future intention, volition or prediction. This general semantic overlap supports the notion that the contemporary English modal system is characterised by polysemy, and consists of categories, rather than a basic meaning for each modal.

These categories are traditionally described as being discrete (cf. Aijmer, 1985:11), but it is sometimes difficult to recognise clear boundaries between categories. For example, modals from the volition or prediction category can signify

[66] Cort and Denison (2005) appropriately describe modality as a shifting category, or a 'moving target'.

meanings from another category like obligation and necessity in special contexts, e.g. *will* can denote a sense of obligation in a sentence like "you *will* pick up your toys" in a certain context (similar to e.g. "you *must* pick up your toys", but with an intensified sense of obligation that is perhaps more clearly reflective of the imperative mood). These kinds of categorical overlap cases will however not be foregrounded in this thesis, as the focus will be on specific modals and their possible meanings, rather than meanings and their possible modal expressions. Still, gradience and indeterminacy are present in the modal system (cf. Coates, 1983), and it is precisely because of this that I compile microsemantic parameters for my study in Table 4 of Chapter 3, in order to deal with the indeterminacy problem in terms of high/stronger and median/weaker obligation to some extent – the concept of which will receive attention in paragraphs to come.

Coates (1983:10) feels that semantic analysis can still be achieved despite indeterminacy, even though she is cautious to support either the monosemantic or polysemantic view, and instead argues for a synthesis between the two approaches:

> "...while indeterminacy undoubtedly exists, it would be unwise to jettison the qualities of explicitness and precision typical of logical semantics. It is surely not impossible to be precise about indeterminacy. Moreover, ...it is not simply a case of adopting or rejecting discrete categorization, or of preferring a monosemantic or a polysemantic approach; ...both categorical and non-categorical approaches are relevant and therefore an adequate description of the meanings of the modals must achieve a synthesis of these two approaches."

Coates' (1983:10) above-mentioned synthesis is based on the 'fuzzy set', which is defined as "a class in which the transition from membership to non-membership is gradual rather than abrupt" (Zadeh, 1972:4). There may therefore be a gradient area between a perceived core and periphery meaning, and this intermediate area is termed the 'skirt' by Coates (1983:13). This means that, even though categories (discrete or broad) may be recognisable, it should be kept in mind that sometimes modality and its meanings might be distinct and unambiguous, and sometimes not. Linguistic and external contexts are major factors in this case. The term 'fuzzy set' illustrates the system well: the boundaries of modal meaning may often be fuzzy, yet it remains a set, which suggests that there is also something 'un-fuzzy' about it. According to this approach, modal meaning therefore operates on a special kind of continuum or cline (cf. Lakoff, 1972; Labov, 1973).

Coates (1983) describes the following traits of the three areas of the fuzzy set. The core meaning usually corresponds to the cultural stereotype that is first learned by children and would be the first response if people had to 'give an example of e.g. *must, may* or *can*' at random (1983:13). Yet, core examples occur rather infrequently in real usage, whereas the majority of examples are found in the skirt and at the periphery, where the latter frequently displays the quality of an 'emergent category', "because it is often possible to define peripheral examples by contrast with the core" (1983:13). I generally agree with the polysemantic approach to English modality and the modal system, and do not intend to pronounce on typological matters in this thesis, except on a very simple basis when certain comparisons with Afrikaans is drawn. Nevertheless, in consideration of Coates' (1983) argument for modality as a fuzzy set, I certainly acknowledge the gradience of modality, but I also recognise that some less 'fuzzy' categories may exist within this gradience, as relevant sections in Chapter 3 will show.

Although the cognitive linguist George Lakoff was optimistic about fuzzy set theory in a 1972 article, he lost interest in the idea in his 1987 book entitled *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind* (1987:196), which is rather informed by prototype theory. Lakoff (1987) indeed suggests that fuzziness exhibited by categories is only comparable in superficial terms with the prototype effects exhibited by categories (see Taylor, 1995:55).

Karin Aijmer (1985) also prefers idea of linguistic prototypes (cf. Langacker, 1987), which contrasts with the notion of discrete, non-overlapping categories – thus leaving space for gradience. She states that the "view of the modal category as a discrete unit in opposition to other categories is ...difficult to combine with the vagueness and non-distinctness of the modal categories in language" (1985:11). In order to achieve what she terms a 'weaker' definition of a modal category to better reflect the nature of language, she proposes the concept of a linguistic prototype (1985:11). The idea of prototypes in language was of course first postulated by the psycholinguist Eleanor Rosch and associates (e.g. Rosch, 1975a; Rosch, 1975b; Rosch & Mervis, 1975), in order to describe natural kind categories, e.g. 'bird'. This involved the existence of a taxonomy comprising a "'basic level' category (e.g. bird) and its hyponyms (sparrow, finch, eagle, penguin, etc.)" (Aijmer, 1985:11)[67].

[67] The idea of the prototype in general was first postulated by Wittgenstein (2001 [1953]).

'Sparrow' would be a more prototypical member of the 'bird' category when compared to e.g. 'penguin', yet the non-prototypical members, even though they have little in common on an individual level, still have a 'family resemblance' to the prototype (1985:11). Prototype theory accordingly underpins the notion of semantic change, which would account for the different directions of semantic development of the modals as well, which will be discussed later in more detail. Very generally speaking, this theory therefore corresponds with the ideas of a cline of meanings and a fuzzy set of meanings, which both emphasise gradience in meaning.

John Taylor provides a comprehensive exposition of semantics in terms of prototype theory in his book *Linguistic Categorization: Prototypes in Linguistic Theory* (1995). He presents his view on linguistic categories as follows:

> "To the extent that a language is a conventionalized symbolic system, it is indeed the case that a language imposes a set of categories on its users. Conventionalized, however, does not necessarily imply arbitrary. The categories encoded in a language are motivated, to varying degrees, by a number of factors – by actually existing discontinuities in the world, by the manner in which human beings interact, in a given culture, with the world, an by general cognitive processes of concept formation." (Taylor, 1995:viii-ix)

The physical world, the culture of a society and human communication as realised in cognition therefore motivate the encoding of conceptual categories. This cognitive process of concept formation is based on prototypes (1995:40-6). Taylor contrasts the classical view on categories with the prototype category in the following terms: the classical view "permits only two degrees of membership, i.e. member and non-member", but for the prototype category membership "is a matter of gradience" (Taylor, 1995:54) (cf. Labov, 1973).

Taylor further mentions that prototypes emerge on the basis of efficiency. He states that "new entities and new experiences can be readily associated, perhaps as peripheral members, to a prototype category, without necessarily causing any fundamental restructuring of the category system" (1995:53). Along the same lines, Geeraerts (1985:141) states that "[c]ognition should have a tendency towards structural stability" and that "the categorical system can only work efficiently if it does not change drastically any time new data crop up." On the other hand "it should [also] be flexible enough to adapt itself to changing circumstances" (1985:141).

Taylor applies the idea of prototypes not only to monocentric language categories (involving monosemy), but also to polycentric ones (involving polysemy), which are the more common categories in natural language. He proposes that membership in a category is "a function of similarity to one of several prototype representations", i.e. a kind of family structure (Taylor, 1995:99), as Aijmer (1985) also notes, but Taylor prefers the terms 'core' and 'peripheral' members. The modal auxiliaries, being polycentric, would then also adhere to a family structure of meaning prototypes. Although this thesis does not intend to interpret its data in terms of cognitive linguistic ideas, prototype theory is very useful in describing the semantics of modality, the categories of the system and the gradience within those categories.

Salkie (2009:79) asserts that previous studies that "have argued for a prototype approach to modality ...have not constructed such an account consistently". Salkie (2009:88; original numbering) therefore proposes four criteria for core members of the modal category, namely:

A. They express possibility or necessity.
B. They are epistemic or deontic.
C. They are subjective...
D. They are located at one of the extremes of a modal scale.

While criterion A links with the model of Van der Auwera and Plungian (1998), criterion B draws on the most traditional distinction in modal meanings, and criterion C involves the speaker's "commitment ...to the propositional content of the utterance" (Verstraete, 2001:1517) (see Salkie, 2009:81-7). All of the distinctions supposed by criteria A to C will receive more attention later in this section. Salkie's (2009:87) criterion D links with the scalar quality of modal expressions, i.e. gradience in meaning, where a core member would typically be located on the extreme end of the scale or continuum of meaning.

Palmer (2001 [original edition, 1986:4]) implicitly provides a formal view on gradience, which can be described as the 'symptom' of a particular stage in the grammaticalisation process, which is in itself a 'matter of degree'. He states:

> "Even at the formal grammatical level, grammaticalization is a matter of degree, of 'more or less' rather than 'yes or no'. Inflectional mood is a very clear example of grammatical marking, but the markers of modality may be modal verbs, clitics or

> particles. Whether these are grammatical or not can only be decided in terms of the degree to which they have syntactic restrictions and the extent to which they can be defined as a limited rather than open-ended system of items. … ...a modal system (or any other grammatical system) will develop gradually over time, and at any one point in time will have reached a particular stage of development and so show a particular degree of grammaticalization." Palmer (2001 [original edition, 1986:4]).

Palmer therefore argues that, as the language evolves, a certain stage of the grammaticalisation process will render a grammatical class like the modal verbs a 'limited system', which supports the notion that they are indeed still a set, even if their meanings are sometimes fuzzy, and that the fuzziness is often owing to an 'incomplete' stage in the process of grammaticalisation. Traugott (2006:112) also refers to the idea that "certain areas of modality are synchronically indeterminate or vague in certain contexts"; a claim in studies on grammaticalisation and semantic change that has been that "at periods of change, indeterminacy (rather than ambiguity, which presupposes discrete alternatives) is to be expected" (cf. Denison, 2001; Enfield, 2003). This period of indeterminacy can last a very long time – centuries and even millennia (Traugott, 2006:112).

With the modal system and its meanings being as complex and oftentimes 'fuzzy' as it is described above, the question might arise as to the origin of the study of modal meaning. The general apprehension of modality stems from a branch of philosophy of language, modal logic, which is basically concerned with certain modal expressions (Wright, 1951; Bybee and Fleishman, 1995:4; Garson 2009). Garson (2009:1) defines modal logic as follows:

> "Modal logic is, strictly speaking, the study of the deductive behavior of the expressions 'it is necessary that' and 'it is possible that'. However, the term 'modal logic' may be used more broadly for a family of related systems. These include logics for belief, for tense and other temporal expressions, for the deontic (moral) expressions such as 'it is obligatory that' and 'it is permitted that', and many others. An understanding of modal logic is particularly valuable in the formal analysis of philosophical argument, where expressions from the modal family are both common and confusing."

This allusion to the "common and confusing" nature of modal expressions seems to be an accurate characterisation of the views given above, and this 'nature' has inevitably led to terminology issues. De Haan notes that in typological studies of modality "[t]here is as yet no consensus on the proper terminology for modal meanings" (De Haan, 2006:28). The categories or types of modality are indeed

represented by various terms in the literature, and I offer a general exploration of those terms here, with the aim of consolidating the general macrosemantics of modality into specific terms to be used in this thesis.

In broad terms, Biber *et al.* (1999:491-497) distinguish between intrinsic and extrinsic modality: intrinsic modality expressing deontic or personal meanings and extrinsic modality expressing epistemic or logical meanings. Huddleston's (2002:177-179) classification is similar to this, drawing a distinction between epistemic modality, being logical and objective, and deontic modality, which is seen as subjective (cf. Krug, 2000:39-42). Palmer (2001:7-10) essentially characterises the same basic types, but by means of different terminology and with some added detail. He proposes propositional and event modality, where, in English, propositional modality signifies epistemic[68] meaning (as well as evidentiality[69]), and event (or root) modality entails both deontic (e.g. externally imposed obligation or permission) and dynamic (e.g. individual ability or volition) meanings, i.e. depending on whether the source of the modality is respectively external or internal to the relevant individual in his/her circumstances (this links with the objective/subjective distinction discussed later in this section) (cf. Nuyts, 2006:2-8). Coates (1983:13), favouring the terms root and epistemic modality, suggests that fuzziness is more typical of root modality (deontic) than of epistemic modality, but adds that epistemic meanings can also be fuzzy. Larreya (2009) also distinguishes between root and epistemic modality. He argues that root modality includes physical modality, e.g. *He had to stop – he was exhausted* (which is essentially equivalent to the dynamic meaning) and deontic modality, e.g. *You must stop* (Larreya, 2009:10). Epistemic modality on the other hand includes problematic modality, e.g. *He must be tired*, and implicative modality, e.g. *You have to be mad to do that* (2009:10).

---

[68] Alethic modality (which is a "term derived from modal logic") is sometimes adopted "as part of [the] theoretical framework for the analysis of modal verbs", where it "contrasts with epistemic and deontic modality" (Crystal, 2008:19). It is essentially "concerned with the necessary or contingent truth of propositions", as is for example found "in the sentence *A triangle must have three sides*, i.e. 'It is impossible for a triangle not to have three sides'" (Crystal, 2008:19). Salkie (2009:82) argues that alethic modality belongs to the periphery of the modal prototype.

[69] Evidentiality is essentially "a type of epistemic modality", in which "propositions are asserted that are open to challenge by the hearer, and thus require justification" (Crystal, 2008:176-7). The evidential construction, e.g. sentences like "'I saw it happen', 'I heard that it happened', 'I have seen evidence that it happened (though I wasn't there)', or 'I have obtained information that it happened from someone else'" expresses "a speaker's strength of commitment to a proposition in terms of the available evidence (rather than in terms of possibility or necessity)" (Crystal, 2008:176-7).

Van der Auwera and Plungian[70] (1998:80) present a universal semantic map proposing "possibility and necessity as paradigmatic variants" in four domains of modality (cf. De Haan & Hansen, 2009:1-2; Salkie, 2009:81). The first domain is "participant-internal modality", being "a kind of possibility or necessity internal to a participant engaged in the state of affairs" and the second is "participant-external modality", which is based on "circumstances that are external to the participant, if any, engaged in the state of affairs and that make this state of affairs either possible or necessary" (Van der Auwera & Plungian, 1998:80). The third domain is "deontic modality", which is essentially categorised as "a subdomain or special case of participant-external modality" that "identifies the enabling or compelling circumstances external to the participant as some person(s), ...and/or as some social or ethical norm(s) permitting or obliging the participant to engage in the state of affairs" (1998:81). The fourth domain is "epistemic modality", where "a proposition is judged to be uncertain or probable relative to some judgment(s)" (1998:81). In essence, this firstly entails that the possibility paradigm can involve either epistemic (uncertainty) or non-epistemic meanings, the latter of which is made up of both participant-internal possibility (including dynamic possibility, ability and capacity) and participant-external possibility (including non-deontic possibility and deontic possibility/permission) (1998:82). Secondly, the necessity paradigm can also involve either epistemic (probability) or non-epistemic meanings, the latter of which constitutes, again, both participant-internal necessity (a need) and participant-external necessity (including non-deontic necessity and deontic necessity/obligation) (1998:82). Therefore, the semantic map of Van der Auwera and Plungian (1998) in broad terms distinguishes between non-epistemic and epistemic modality, where the non-epistemic domain can include dynamic and deontic meanings (corresponding to the idea of a 'root' meaning). The layer of dynamic modality is however often adopted together with deontic and epistemic meanings as a "tripartite classification scheme", as e.g. in the analyses of e.g. Collins (2009a:159) (cf. Nuyts, 2006:2-8), but e.g. Salkie (2009:81-2) argues for the assignment of "dynamic uses of modals to the periphery" of the modal prototype (cf. Palmer, 1990:37).

[70] The model of Van der Auwera and Plungian (1998) has recently been reconsidered, developed and refined by e.g. Van der Auwera and Ammann (2005), Van der Auwera (2008) and Van der Auwera *et al*. (2009) (cf. De Haan & Hansen, 2009:2).

These above-mentioned classifications, in essence, overlap in their use of the terms *deontic* and *epistemic* as part of their various groupings, and these two terms will therefore be favoured in this thesis (but not with the exclusion of the dynamic layer). Moreover, this is the traditional major distinction in modality (Krug, 2000:41; cf. Nuyts, 2006:2) – "[a]t the heart of many analyses" (Salkie, 2009:82) – and is also adopted by e.g. Lyons (1977), Sweetser (1990) and Traugott (1997). Both these terms were indeed originally used in modal logic (Wright, 1951:1), and are etymologically derived from classical Greek. Krug (2000:41), for example, offers an explanation of this, as well as examples to aid understanding. The word *deontic* is "immediately derived from the deverbal noun δεον 'that which is needed', but goes back ultimately to the verb δεω 'bind', 'need'". This means that deontic modality is concerned with meanings of obligation and permission, for which, respectively, examples from Krug (2000:41) include:

(1) I **must** go home now. (Mum says so.)

(2) You **may** go home now. (School is over.)

These *deontic* meanings of obligation and permission are associated with social functions; moreover with the possibility or necessity of acts performed by modally responsible agents (Lyons 1977:823; cf. Bybee & Fleischman, 1995). On the other hand, the word *epistemic* derives from Greek επιοτημη 'knowledge', and "commits the speaker to statements about the truth of a proposition", which means that epistemic semantic domains include probability, (logical) necessity and possibility (Krug, 2000:41; cf. Lyons, 1977:793). Thus, the epistemic meaning is involved with expressions of knowledge and belief (Lyons, 1977:793). Examples of respectively probability and possibility are (Krug, 2000:41):

(3) He **should** be at home (The Super Bowl is on tonight and wouldn't miss it.)

(4) They **might** be in the cinema. (They talked about going this afternoon.)

A point of some disagreement is whether the meaning of volition, which is sometimes called 'desire', should be included in the modal semantic domain. Krug

(2000) for example follows the time-honoured tradition of e.g. Jespersen (1924) and Givón (1984), who categorise volition as part of deontic modality (cf. van der Auwera & Plungian, 1998). Krug (2000:41) offers some justification "inherent in the term *deontic* that warrants the inclusion of volitional modality: the middle form of Greek δεω i.e. δεομαι actually has a 'desire' reading." As mentioned earlier in this section, Palmer (2001) and Collins (2009a) subsume volition and ability meanings under the dynamic category of modal semantics, which adds an additional layer of meaning to the traditional deontic-epistemic distinction. In this thesis I will favour the inclusion of dynamic modality.

Bybee *et al.* (1991:23) offer yet another model for types of modality. Their three categories of modality include:

(a) *agent-orientated* (desire, obligation ability, root possibility, permission)
(b) *epistemic* (possibility, probability)
(c) *speaker-orientated* (imperative, hortative, optative)

Krug (2000:42) notes that this approach is not fundamentally different from the more traditional distinctions, and suggests that agent-orientated modality overlaps with deontic modality to a great extent. Permission and obligation are traditionally associated with deontic meaning, and the inclusion of the desire meaning among these deontic meanings follows the conventional approach of e.g. Jespersen (1924). Apart from the conventional view, the inclusion of desire and ability into this category makes even more sense if one considers the dynamic category of e.g. Palmer (2001) and Collins (2009a), which is indeed orientated toward the agent of the modal expression.

Within these macrosemantic categorisations certain microsemantic classifications are possible, for example distinctions between levels of deontic obligation. Such degrees of strength in obligation are categorised in terms of high (required), median (supposed) and low (allowed) degrees by Halliday and Matthiessen (2004:620), where the modal verb *must* is generally regarded to represent a higher degree of obligation, the quasi-modal verb *HAVE to* a median degree and the modal *should* a lower or weaker degree (cf. Bybee *et al.*, 1994; Huddleston, 2002:177; Leech *et al.*, 2009:86-88;114-116). Halliday (1970), in his above-noted

functional application to modality, groups *must* under his ideational function of compulsion, which is described as 'required', and in its negated form *must* also fulfils the function of permission together with *may* (Halliday, 1970:340). *Should,* on the other hand is grouped under the function of obligation, which is described as oblique (tentative/hypothetical) – expressing notions of being 'obliged', 'supposed', 'desired' or 'expected' (1970:340). Furthermore, when combined with his interpersonal function, Halliday links the functions of obligation (including *should* and *ought to*) and compulsion (including *must*) to respectively 'relative certainty' and 'absolute certainty' (1970:348), which are both part of the 'strong' committed function, whereas he groups e.g. *may* under the 'weak' committed function.

Furthermore, Alexander (1988:228) proposes that "*must* conveys more strongly than *HAVE to* the idea of inescapable obligation", and similarly Sweetser (1990:540) notes that "[*m*]*ust* has connotations of a directly applied and irresistible force, while *HAVE to*, *ought* and *need* are resistible forces". In contrast to this description, Collins (2009a:33) prefers the dichotomy of strong forms *must, HAVE to, (HAVE) got to, need and NEED to* and medium strength forms *should, ought and BE supposed to.* In essence, these approaches are not that different, as they characterise *must* and *HAVE to* as expressing a stronger degree of obligation (or compulsion in Halliday's [1970] terms) than *should*, and, indeed *must* remains the strongest option of the 'stronger' obligation modals. In this thesis I will therefore, in accordance with Halliday and Matthiessen (2004), use the terms 'high' for a stronger degree of obligation and 'median' for a weaker degree.

Verstraete (2005:1416) argues that an extra layer of meaning needs to be added on top of modal strength when describing deontic expressions, since such expressions "have both a modal source and a modal agent, and carry certain presuppositions about the modal agent's willingness to carry out the action described in the clause". The main distinction here is between the agent's willingness in expressions of permission and obligation, where permission "presuppose[s] the agent's willingness" and obligation "presuppose[s] unwillingness" (Verstraete, 2005:1416). Naturally this thesis is more interested in obligation, which implies a certain level of unwillingness from the agent, but the focus is on modal strength within obligation meanings rather than on its layered differences with permission meanings.

These degrees of modal strength (or illocutionary force) are sometimes linked to the idea of subjective and objective sources of modality, but the literature is not always clear as to the exact nature of the link. After an explanation of subjective and objective meanings regarding modal *must* and quasi-modal *HAVE to* in this section, the link with modal strength in SAfE will be evaluated in Chapter 4 in terms of the literature. In the simplest terms, objective meaning involves the orientation of the source of obligation as external to the speaker, whereas in the subjective[71] meaning the speaker is the deontic source (Collins, 2009a:60; cf. Salkie, 2009:82-7). There are various views and various labels applied as to the subjective and objective nature of *HAVE to* and *must*, which Depraetere and Verhulst (2008:3) briefly survey. They list Coates (1983), Quirk *et al*. (1985) and Huddleston and Pullum (2002) as proponents of the idea that *have to* can only be objective, Palmer (1990) that it is neutral and circumstantial, Tregidgo (1982) that it is not speaker-orientated and Larreya and Rivière (2005) that is has a neutral orientation. Essentially all these labels point to a more objective source. On the other hand, e.g. Tregidgo (1982) and Huddleston and Pullum (2002) label *must* to include both subjective and objective meanings, while Coates (1983), Quirk *et al.* (1985) and Larreya and Rivière (2005) prefer the subjective label, and Palmer (1990) the subjective and neutral label. Depraetere and Verhulst (2008:2) note that the common argument involves that "*have to* expresses 'external' necessity, originating in e.g. a rule or regulation, and that *must* typically expresses an obligation imposed by one of the discourse participants, i.e. the speaker in affirmative sentences and the addressee in questions". The following examples[72] illustrate this line of argument:

(5) You **have to** come in now.[73] ('I'm likely to be relaying someone else's instruction')

(6) You **must** be back by 10 o'clock.[74] ('You are obliged to be back…' 'I require you to be back…')

---

[71] Timotijevic (2009:114) argues that "subjectivity is a pragmatic notion", in that "[u]ses of modals that have been called 'subjective' crucially involve the speaker and thus involve more pragmatics than uses which are 'objective'" (cf. Recanati, 2004).

[72] Taken from Depraetere and Verhulst (2008:2).

[73] From Huddleston & Pullum (2002:205-6).

[74] From Quirk *et al*. (1985:225).

Along similar lines, Westney (1995:151) notes that deontic *HAVE to* relates to "an external, existent obligation that can be perceived or described independently of the speaker", and Collins (2009a:60) further notes (with reference to Westney [1995]), that deontic *must* is "a very general marker of obligation", and relates to "senses such as urgency, irresistibility and unconditionality". Collins (2009a) adds Perkins (1983) on top of Palmer (1990) and Coates (1983) (also mentioned by Depraetere and Verhulst [2008]) to the list of scholars who agree "that *HAVE to* contrasts with *must* in that its distribution is skewed towards the objective (Palmer's 'external') end of the subjective/objective scale" (Collins, 2009a:60).

Collins (2009a), on the basis that there is some disagreement in opinion as to the stronger force of *must* in relation to *HAVE to* (McCallum-Bayliss [1985] for example perceives *HAVE to* as stronger), proposes that the notion of strength is "a less useful concept for differentiating these items than subjectivity/objectivity" (Collins, 2009a:61). Yet Collins admits that in his analysis of native varieties of English (which will receive more attention later) "the subjective/objective distinction cannot be consistently applied to instances of deontic *HAVE to* insofar that it is commonly used" (cf. Leech, 1987:79), but that it is 'skewed' toward the objective, which "would appear to suggest that deontic *must* is stronger than *HAVE to*" (Collins, 2009a:61) . Here a link between these two areas is suggested.

Collins' interpretations of the subjective/objective distinction however suggest that the traditional view as noted above is not always as clearly reflected in his contemporary data for BrE, AmE and AusE. He for instance finds that the default interpretation of deontic *must* is that "the speaker is identified as the deontic source", but that there are cases in which "there [are] no necessary connections between subjectivity and the use of *must*" and that *must* can even be "objective, with the source of obligation ...external to the speaker" (2009a:35). Collins' findings further indicate a tendency "for the subjectivity and objectivity of deontic *must* to correlate with the person of the subject", for example, most of the instances of subjective deontic *must* have second-person subjects. This matches Coates' (1983:33-5) findings, which will be further discussed below. Due to the large number of indeterminate cases in his data, Collins (2009a:37) does not quantify the subjective/objective dimension, but, based on the idea that this dimension correlates with the subject, the majority of third-person over second-person subjects "suggest[s] that deontic *must* is more commonly used objectively than subjectively". This shows

that the traditional view might be in need of some re-evaluation.

Depreatere and Verhulst provide new insights into the topic concerning BrE in their article "Source of modality: a reassessment" (Depreatere and Verhulst, 2008), with a focus on what they prefer to call 'root modality' or 'non-epistemic necessity' (2008:2;4), which basically translates into the concept of deontic modality, but also includes dynamic modality. They find that the "[s]emantic and pragmatic features of the different sources, ...are ...more diverse than has previously been claimed" and that "the traditional distinction in meaning between the so-called 'objective' *have to* and 'subjective' *must* is not as outspoken as is assumed" (2008:1). Their argument is that a more sophisticated approach regarding the "the SOURCE of the necessity, i.e. the driving force behind the necessary state of affairs expressed by *have to* and *must*" is needed instead of the "simple binary division" (2008:1, emphasis original).

Depreatere and Verhulst (2008:5) propose different terms and subcategories for the various sources in which the non-epistemic necessity originates, viz. discourse-internal (from the speaker or hearer – correlates with subjective sources) or discourse-external (from a regulation, condition or circumstances – correlates with objective sources). They also offer an additional conceptual distinction between the source and the 'channel' of necessity, the latter of which refers to "the entity that communicates the necessity" (Depreatere and Verhulst, 2008:3), most explicitly in cases of reported speech or thought (2008:5). Their study of BrE shows that "no modal has 'exclusivity rights' on a particular source", and also that "no particular source is exclusively associated with one particular auxiliary" (2008:4). The only exception is the preference for *HAVE to* to be used in the circumstances category (2008:23-4). They furthermore observe that register and modal strength are among the main criteria for choosing a modal: "*have to* typically occurs in spoken English", whereas "*must* predominantly in written English", and "*must* conveys necessity with more insistence than *have to*" (2008:24). Here reference to modal strength is made, and Depreatere and Verhulst indeed also evaluate the relevance between source and strength.

They offer two factors that determine modal strength "with certainty": the possibility of noncompliance or the gravity of noncompliance (Depraetere and Verhulst, 2008:15). Depraetere and Verhulst interpreted an utterance as strong if the necessity is inescapable by means of e.g. rules, regulations and laws. They share some

of the methodological challenges of such microsemantic analyses and make some observations:

> "Assessing the outcome of noncompliance is ...a subjective matter and not ...simple (...). In the case of discourse-internal sources (...) the strength of the necessity depends on the social relationship between the discourse participants and the addressee's attitude towards authority. Comparing the strength of a particular 'discourse-external source' necessity with that of a particular 'discourse-internal source' necessity is even more delicate. ... ...utterances with generic *you* or *one* as subject are less likely to have an impact on the addressee than when the latter is straightforwardly addressed with a second person (nongeneric) *you*. ...it is difficult to achieve an accurate measurement. ... ...any measurement is bound to be relative and to a certain extent vague" (Depraetere and Verhulst, 2008:15-16)

Depraetere and Verhulst (2008:16) furthermore observe that "there is no strict correlation between source and strength such that one might say, for instance, that discourse-internal sources are strong and discourse-external sources less to." This issue will be investigated in the semantic analyses of this thesis.

The question now arises whether there is an actual link between the source and force/strength of modality in SAfE. Coates (1983) explores this link in her corpus-based study on BrE. In her discussion on root modality (which includes both deontic and dynamic meanings), She suggests that, in the terminology of fuzzy sets, the root meaning of *must* is "typically fuzzy", with meanings ranging from strong obligation (paraphrasable as 'it is imperative/ obligatory') at the core, to weak obligation (paraphrasable as 'it is important') at the periphery of a cline or continuum (Coates, 1983:32). This however accounts for only half of the picture. Coates adds that the "presence or absence of another feature", namely "the speaker's involvement in the utterance", or level of subjectivity, complicates the interpretation of *must* (1983:32). She notes that "[b]etween [the] two extremes [on the cline], there are two inter-related but independent clines: Subjective-Objective and strong-weak" (1983:34). From Coates' quote it can therefore be deduced that these two clines can interact or not. For the cline of subjectivity, she however implies that a core example or stereotypical extreme[75] on this cline (the 'performative' or imperative stereotype) has the following features where it interacts with the core of the strong obligation cline, (Coates, 1983:33):

[75] This links with Salkie's (2009:87-8) fourth criterion of a modal prototype, namely "[e]xtremeties of the ...scale".

(i) Subject is animate
(ii) Main verb is activity verb
(iii) Speaker is interested in getting subject to perform the action
(iv) Speaker has authority over subject.

Additionally, to illustrate the relative force of obligation, Coates (1983:36) plots her data against the following parameters, where parameter (a) represents the strongest end of the force and (h) the weakest (note how the stronger end of the scale corresponds to the features (i) to (iv) above):

(a) second person subject
(b) speaker involvement [links with subjectivity]
(c) speaker has authority over subject
(d) verb is agentive
(e) paraphrasable by 'it is obligatory/absolutely essential that'
(f) animate subject
(g) paraphrasable by 'it is important that'
(h) inanimate subject

Coates however finds that such core examples are infrequent. In her corpus-based study "performative examples account for roughly 1/14 of all [r]oot examples" (1983:33). She argues that this shows that root modals are more frequently objective; therefore, their meanings are more often situated in the periphery (1983:33), which corresponds with Collins' (2009a:35) suggestion that the traditional view does not always apply. Thus, Coates' conclusion in this case is that "strong examples are usually subjective and weak examples are usually objective, but this is not necessarily true [for all cases]" (1983:37). From this perspective, Coates asks the question of where the 'performative' stereotype or core use comes from (in e.g. Taylor's terms this will be the prototype), and explores the answer that "children use [r]oot [*must*] as a performative" and that this connects with strong obligation. This is however not the case for adult language, as Coates (1983:38) notes:

> "...early language seems to concentrate on the strong/subjective end of the continuum, i.e. the core and close to the core, while adult usage ...is far more fuzzy, covering the whole spectrum from most to least strong and most to least subjective, with the majority of examples being assigned to points on the continuum between the core and the periphery".

Apart from Coates' findings on the features that are present when core strong *must* overlaps with core subjective *must* (i-iv), and apart from theoretical considerations, no clear corpus evidence of *must* or *HAVE to* (or *should* for that matter) as a 'stronger' or 'weaker' carrier of force has been provided yet, which means that the 'usefulness' of the distinction (cf. Collins, 2009a:61) is not really explored.

However, Collins (2009a:37-8) reports the following for *must*:

> "...deontic *must* is semantically strong, but it is commonly prone to pragmatic weakening. Its strength – what Sweetser (1990:54) calls its 'resistibility', degrees thereof being determined by the severity of the consequences for non-fulfilment of the obligation – is closely associated with subjectivity/objectivity. The present data evidenced a tendency for subjective uses to be strong, objective uses to be weak".

Apart from this general tendency, Collins (2009a:35) also reports that "[s]ubjective deontic *must* is commonly used in cases where the speaker is not in a position to – or may not even wish to – require actualization, as in …advice …or …[a] request …or …exhortation." The words 'tendency' and 'commonly' in the above quotations suggests that both uses are established. The link or the relevance of the link between modal strength and modal source is therefore not clear from Collins (2009a). This thesis will assess this link by analysing both the subjective/objective source distinction and the distinction in the degree of strength for *must, should* and *HAVE to* diachronically for SAfE – the method of which is presented in Chapter 3.

As seen above, there are many different approaches to defining and categorising modality in terms of grammar and meaning. Van der Auwera & Plungian (1998:80), who take a pragmatic view on definitional issues, note that because "modality and its types can be defined and named in various ways", and because "[t]here is no one correct way", "[t]he only requirement is that one makes clear how one uses one's terms". Thus, to reiterate my preference: in this thesis I opt for the traditional distinction between deontic and epistemic modality, but with the added layer of dynamic modality (deontic and dynamic modality are still sometimes referred to under the umbrella term 'root', when applicable). I therefore support the position of polysemy within the modal system, which implicates (mostly) discrete categories and a degree of gradience within such categories. Traugott (2006:128) deduces:

> "It seems best to conclude that because modality is a gradient notion, semantically as well as morphosyntactically, it can be represented in a variety of morphosyntactic ways. Furthermore, because it is an enormously complex system that directly expresses speakers' beliefs and evaluative attitudes, it is not surprising that certain subparts may non-deterministically come to be focal or "gravitational" areas (Krug, 2000:242-247) "verb by verb, submeaning by submeaning, dialect by dialect" (Van Kemenade, 1992:306)."

The idea of 'focal' or 'gravitational' semantic areas links with the notion of discrete categories and leaves space for gradience; this summarises my view perfectly. I also prefer the terms 'high' and 'median' for degrees of obligation, rather than 'weak' and 'strong'. From the above exposition, it can be concluded that the modal system is prone to polysemy, gradience and fuzziness, but that a broad framework of categories can still be drawn within this polysemous system, even if the boundaries of this framework can sometimes vary in their degree of discreteness and modal meanings can sometimes be indeterminate, especially during ongoing processes of grammaticalisation. In this thesis I will opt to use the term 'cluster' to denote each grouping of a broad category of meanings, as in e.g. 'the obligation and necessity semantic cluster', in view of the gradient nature of modal semantics.

The next sub-sections will chronicle the gradual development of the English modal system through the process of grammaticalisation, with the goal of exploring the evolution of many polysemous relationships within and among the modal and quasi-modal auxiliaries and their 'focal' or 'gravitational' areas or categories.

2.4.1.2 The development of the modal class

As mentioned above, the English modal system has followed a gradual path of extended-period evolution. This section describes the evolution or emergence of the polysemantic modal auxiliaries via grammaticalisation.

The emergence of the English modal auxiliary system occurred largely due to the decline of the effectiveness of mood oppositions (especially indicative vs. subjunctive) and the internal dynamics of the developing modal group (Warner, 1993:192). Some of the functions of the indicative-subjunctive opposition (which was already losing its distinctiveness in OE) were taken over by modal group verbs, which were, at first, not simply equivalent to these mood expressions, but their growing

frequency, coherence and grammaticalisation accelerated the decline of the subjunctive (1993:192).

From the emergentist perspective (Hopper, 1998; Bybee, 2006), forms that evolve into a schematic grammatical unit may expand the occurrences of the instances of the schema in a language (the schema being a general concept structuring a perception of the world [Matthews, 2007; cf Lakoff, 1987]), because the schema itself achieves salience and increases the probability of use of the forms that instantiate the schema. Deutscher (2005) argues for an extensive contribution of the role of analogy in expanding the use of certain grammatical patterns in a language. Therefore, as modal verbs became more frequent and salient in ME, especially *must* and *should*, which had been around even before the 10th century (Biber *et al*., 1998:206), the process gained its own internal momentum and become established as grammatical class.

Krug (2000:44-5) recounts that the central modal verbs evolved from an inventory of OE preterite present verbs, which had a highly irregular morphology and semantics (even pre-OE), but in OE syntactically behaved as main verbs. The status of the lexical verbs therefore changed to that of auxiliaries, which means they underwent recategorisation, which is a prime example of syntactic change and a clear instance of grammaticalisation (Hopper & Traugott, 2003:55). Krug stresses the formal (e.g. *NICE*) properties mentioned in the previous section as the criteria these verbs gradually developed in order to become a class (Krug, 2000:45). Yet, the set of present-day modals does not wholly correlate with the OE inventory of preterite present verbs, owing to the fact that some preterite presents were lost, such as *UTON (let's)*, **THURFAN (need)* and *WITAN (know)*, but, on the other hand, the set was supplemented by one non-preterite present verb *WILLAN (will)*, which also had modal semantics, seeing that it originally already denoted volition (2000:45).

Just like the system itself, the semantics of the modals also continued to develop gradually, without ever attaining a state of a one-to-one relationship between form and meaning, but maintaining various many-to-many pairings. One example of this overlapping nature present in the development of a grammar system can be seen in the OE *wile* (later *will*) and *sceal* (later *shall*), where futurity was often difficult to isolate from volition or obligation. *Will* was frequently used as a future auxiliary, sometimes showing no trace of its original sense, but *shall* retained "some tinge of obligation or necessity until ME, although it was plainly used for a prophetic future or

future of appointment" (Warner, 1993:167).

Another area of semantic overlap in OE occurred in the possibility and permission domain, especially between the verbs *mot*/*mote* (*must*) and *mæg* (*may*). In OE *mot* was already centrally deontic, but examples do occur in which *mot* seems "to combine a notion of objective possibility with the sense that it is allotted by fate or deity" (Warner, 1993:163). This suggests that in OE there were already dynamic uses of *must* to some extent (1993:163) (see § 2.4.1.2.1 on the semantic development of *must*). Warner adds that *mot* was never a simple equivalent to *mæg*, and that *mæg* was often subject-orientated in OE (1993:163). It is therefore clear that the modals often shared certain meanings, but still remained distinct.

Abraham (2001:30) notes that the modals have undergone a special kind of semantic change since OE: a type of attrition where English modals "have lost their deontic/root meaning to a great extent" (except for *must*) (cf. Abraham, 1998, 2002; cf. § 2.4.1.2.1), which has also happened in German. This means that the originally deontic modals have also gradually moved into the domains of the epistemic and of futurity (Abraham, 1998). Linking with the notion of Abraham (2001), Conradie (1987:171) proposes that changes in the semantics of the English, Dutch and Afrikaans modals have systematically passed along the following unidirectional spectrum or axis (naming the salient fields) (cf. Conradie, 1980) since the Middle Ages:

knowledge ➔ ability ➔ permission ➔ obligation ➔ futurity

This axis of semantic change applies to the Dutch modals 'kunnen' *(can)*, 'mogen' (*may)*, 'moeten' (*must*) and 'zullen' (*shall*) (Afrikaans *kan, mag, moet,* and *sal,* respectively) and Conradie (1987:171) affirms that the changes in these modals "closely parallel changes in English" (cf. Traugott, 1972:198-9). Harris (1987:187) notes that this kind of semantic evolution agrees with a tendency for speakers to "attenuate their utterances", by means of e.g. politeness. This is done to "avoid recourse to the authority of either S[peaker] or H[earer] and [an] illocutionary instrument" (Conradie, 1987:179).

Semantic change has indeed also occurred in the area of temporality. Even pairs of present and preterite tense modals, which usually started out as semantically equivalent in OE, followed on the path of semantic distinction during ME, to

represent different nuances of meaning in different contexts. The original past tense form of *shall* is *should*[76], which in its OE form often signalled the "destiny, duty or obligation of the subject in the past" (Bybee, 1995:504), but today it does not have past readings anymore and needs *have* attached to it in order to achieve a past reading, which has been the case ever since the ME period (1995:514). Goossens (1987:130) stresses the near-synonymy that the older forms of *should* and *ought to* enjoyed even since OE, where in main clauses "as a rule, *sceold-* is paraphrasable as should/ought to" (cf. Visser, 1963-73). However, now it appears that *ought to* is on the path to elimination, and, as Harris (1987:184) states, "marked semantic overlap with a rival form and major morphosyntactic irregularity are not normally conducive to the survival of any linguistic item – and yet it is only now that *ought to* seems in serious danger of passing from the language" (cf. Perkins, 1983; Harris, 1986:356). This suggests that a semantic overlap between to modals can threaten the continued existence of one of the forms, even more so if the 'threatened modal' displays an irregularity in form to the other modal in the relationship.

Indeed, like in the case of *should* and *shall*, the modals *would, might* and *could* are historically the past tense forms of *will, may* and *can* respectively, but in contemporary English the meanings and uses of the past tense modals are not always the past equivalent of the present modals (1995:503). Bybee (1995:503) divides the uses of the past tense modals into three types:

(i) hypothetical uses
(ii) present tense uses
(iii) past tense uses.

The ME period indeed saw both *should* and *would* making their way into present contexts, especially in contexts where the synthetic subject is in the first or second person (1995:505). Bybee (1995:505-6) states:

> "To explain the present uses of Past forms, we must refer to the special properties of modal verbs. Modal verbs, whether they express desire, obligation, necessity, intention or ability, have in common the semantic property that they do not imply the completion of the action or event expressed by the infinitive with which they occur. (This property

[76] The cognate forms of *shall* and *should* in Afrikaans and Dutch are 'sal'/'sou' and 'zullen'/'zouden' respectively, but the cognates will receive further attention in § 2.4.3.1.

> is in contrast with other complement-taking verbs such as *finish* and *manage*, which do imply completion [Givón 1973].) In the present tense, however, these modal verbs do imply a close relation between the agent and the main predicate, which may imply or predict a future completion of the event or activity of the main predicate."

This suggests that the present tense forms of modals more readily imply a sense of future completion in Contemporary English, than their past tense counterparts. The past forms of modals, however, imply that there existed a certain state before the moment of speech, but it is not clear "whether that state still exists in the present or not" (Bybee, 1995:506). Past modals therefore render two vague areas: firstly, "whether or not the predicate event was completed", and secondly, "whether or not the modality remains in effect" (1995:506). Indeed many uses of *shulde (should)* that refer to what was to take place, but with no implication that it did in fact take place (as shown in [7]), are already found in ME (1995:509).

(7) And went on his way with his wyȝen one,
þat **shulde** teche hym to tourne to þat tene place.[77]

'And (he) went on his way with the man
That was to show him how to get to that perilous place.'

This notion of incompleteness adds a layer of hypothetical meaning (expressing the *irrealis*) to the past tense modals (Larreya, 2003:21).

Bybee (1995:511) mentions that the trend toward the use of modals like *should* and *would* in present and hypothetical contexts "has continued to the point where both modals appear in these contexts almost to the exclusion of past contexts". Coates (1983), in her analysis of spoken and written texts, recognises that the hypothetical uses are indeed the more common ones in the past modals of Contemporary English (cf. Bybee, 1995:503). Where obligation meanings were present for *should*, the effect of the hypothetical layer was the weakening of the force of obligation (i.e. a level of politeness) via the implication of outstanding conditions, which has lead to the modern usage of *should* to express weak obligation (1995:513). This, for example, contrasts with a modal like *must*, which, being in the present tense (despite its earlier

[77] Source: *Sir Gawain and the Green Knight* (translation from Bybee, 1995:509).

past tense reading [cf. e.g. § 2.4.1.2.1]), does not express weaker obligation in CE like *should*. Indeed the earlier hypothetical use of *must*, which arose in ME, was already obsolete by the 19th century ((OED Online, 2012b); cf. § 2.4.1.2.1). This discussion of the separate semantic evolution of present and past tense modals shows that even a grammatical set such as *shall* and *should* has moved in the direction of polysemy over time.

Within the contemporary modal system as a whole, the modals *can, could, may* and *might* either denote a permission, possibility or ability meaning, whereas *must, should, need* and *ought* either denote an obligation or necessity meaning. Furthermore *will, would* and *shall* can either have a prediction or volition meaning (Biber *et al*., 1999; Huddleston, 2002; Palmer, 2001). The modal system has therefore become a polysemous system, since its development through grammaticalisation, but the polysemy does not end here. The semantics of the system further expanded with the grammaticalisation of a new subclass, the quasi-modals, which will be discussed in § 2.4.1.3, after some observations on the semantics on *must*.

2.4.1.2.1 The semantic development of *must*

The general semantic development of *must* from OE to present is explored below, based on dictionary entries[78], since *must* is the focus of much of the semantic analyses of this thesis and yields exceptionally interesting results (see e.g. § 4.3.2). The OED Online (2012b) is an excellent resource to trace the semantic development of *must*, since the various senses are accompanied by dated textual examples, casting light on the periods of English in which the senses arose. As also noted in § 2.4.1.2 and § 2.4.1.3, it appears from the OED Online entries that *must* denoted root (deontic and dynamic) meanings in OE, but also expressed permission and possibility meanings. As further explained in § 2.4.3.1, the ancestor of *must* in OE (*moste*) was the past tense form of the OE present form *mot/mote*[79] (OED Online, 2012b). Therefore, the deontic meaning in OE that has remained in use until present is that of "necessity or obligation in the past", but the two OE uses that have become obsolete after ME

[78] The continued development of *must*, as studied via corpora (mainly over the 20th century), will receive attention in § 2.4.2.

[79] See examples (24), (25) and (26) in § 2.4.3.1.

include expressions of "permission or possibility in the past" (see example [23] in § 2.4.3.1) (OED Online, 2012b).

Many new root meanings arose for *must* (*moste*) in ME, mainly because its temporal domain expanded to include the present tense[80]. These meanings have remained in use in CE. New meanings in ME included dynamic and deontic uses, e.g. "[e]xpressing necessity ... in the present tense" and "[e]xpessing an insistent demand or a firm resolve on the part of the speaker or imputed to another person", the latter of which indicates an associated strong or high degree of obligation (OED Online, 2012b). Other root uses that arose in ME are the impersonal (formulaic) expression equivalent to the French *il me faut*, which in recent use only occurs as 'needs must' (expressing dynamic necessity), as well as the use "[w]ith *have* and [the] past participle", which expresses "the present necessity or obligation of a past action or state" (basically equivalent to 'should have done') (OED Online, 2012b). *Must* has also been used as part of a formulaic phrase "[i]n explanatory clauses" from the ME onward (with a root meaning), namely 'you must know/understand' (OED Online, 2012b). One root sense (including both deontic and dynamic uses) which arose in ME, and which did not survive, is the use of *must* "[i]n the main clause of a conditional sentence, or with a condition implied but not stated, expressing hypothetical necessity or obligation" (OED Online, 2012b). Such meanings equivalent to "would be obliged to" in the present have become obsolete since the 19th century (cf. § 2.4.1.2), and meanings equivalent to "would have had to do" in the past, where *must* is used with "*have* and [the] past participle", have become rare since the mid-20th century (OED Online, 2012b). The ME period also saw the expansion of the semantic domain of *must* into epistemic meanings that remain in use in CE. These include the expression of "a fixed or certain futurity ... in the present tense" (equivalent to "shall certainly or inevitably"), and the expression of "the (present) inferred or presumed certainty of a fact" (equivalent to "it is to be inferred that..."), which can also be used "[w]ith reference to the past", moreover "[w]ith *have* and [the] past participle", where the meaning is equivalent to "it is to be concluded that..." (OED Online, 2012b).

The EME period saw the development of new root/deontic uses of *must*, but no new epistemic uses. Among these are two new deontic uses that are still in use,

[80] See § 2.4.3.1 for a further discussion of *mot/mote* and *moste*.

viz. the past time equivalent of the previously noted use "[e]xpressing an insistent demand or a firm resolve on the part of the speaker or imputed to another person", which had developed in its present form during ME (suggesting a high degree of strength of obligation), as well as the use "[i]n the negative, expressing prohibition ...[with] reference to the present", which "also sometimes [expresses] denial of permission to oneself" (OED Online, 2012b). After developing in the EME period, the use of *must* "[a]s a past or historic present tense, used satirically or indignantly with reference to some foolish or annoying action or some untoward event", has however become rare since the early 20th century (OED Online, 2012b). Two EME formulaic phrases that have remained in use in CE are 'I must say', meaning "I cannot help saying", and 'if you must know', which is "used to introduce information provided against the judgement or inclination of the speaker" (OED Online, 2012b).

During the LME period new uses only developed in the epistemic domain[81]. The construction 'must be', which expresses "past inferred certainty" in "oblique and virtual oblique narration" developed in the 18th century and remains in use (OED Online, 2012b). The equivalent use where this construction expresses "the (present) inferred or presumed certainty of a fact" has however become obsolete after the 18th century (OED Online, 2012b). During the 19th century, the construction *must* "[w]ith *have* and past participle" arose in AmE, where it was used "[i]n the negative, expressing presumed certainty of something not being the case" in the past (OED Online, 2012b). The equivalent of this use, with reference to the present, came in use during the mid-20th century, i.e. CE, and remains in use (OED Online, 2012b). According to the OED Online (2012b), the only other new use that developed during the CE period is deontic, viz. the use "[w]ith *have* and past participle... [w]ith reference to the past", that expresses prohibition in the negative (equivalent to "am [is, are] not allowed to have").

By examining the dated examples of the entries for *must* in the OED Online (2012b) we can gain a sense of the diachronic semantic development of *must* from the OE to the CE period. However, other dictionary entries often provide insight into the standard use of a word like *must* at a certain, synchronic, point in history (we might refer to this as the prototype in a certain sense). The next few paragraphs will explore

[81] The OED Online also lists a regional use of *must* that was only used in Derby from the 18th to 19th centuries, namely "requesting permission ...[i]n questions", where *must* is equivalent to *may* – in a sense reverting back to one of the original OE uses (OED Online, 2012b).

such dictionary definitions of modal *must* from points in time that span the earlier periods of the historical corpus (see § 3.2.2.1) – from the early 19th century (and somewhat further back to the mid-18th century) to the early-20th century. The dictionaries quoted here are mostly representative of the British standard[82], except for *Webster's Complete Dictionary of the English Language*, 1886:870), which represents the American standard.

Earlier dictionaries of course did not usually rely on usage-based evidence, as the OED Online (2012b) does, and were naturally rigid and prescriptive in their approach – sometimes even biased. A famous example of this is Samuel Johnson's *A Dictionary of the English Language* (Besalke, 2012:1342) (an example of BrE), published in 1755, in which the definition for modal *must* simply appears as "[t]o be obliged". Hence, only the deontic meaning is given in this mid-18th century entry, with the implication of a high degree of modal strength and a sense of social impact. The same definition regarding obligation is provided by Noah Webster (an example of AmE), but with the addition of "to be necessitated; expressing both physical and moral necessity; ...[t]o be morally required; to be necessary or essential to the character or end proposed" (*Webster's Complete Dictionary of the English Language*, 1886:870). Webster therefore essentially adds the meaning of dynamic necessity to his definition by adding a reference to 'physical necessity'. His definition also refers to a moral meaning, which is more inclined towards the deontic sense. Words such as 'required' and 'essential' suggest a high degree of modal strength. Funk (1898) provides a late 19th-century definition for modal *must* in the *Standard dictionary of the English language* (once again an example of BrE), in which two general senses are indicated. The dictionary entry appears as follows (Funk, 1989; original numbering and emphasis):

> **1.** To be necessitated or obliged; be compelled either by physical compulsion, by constraining authority, or by overwhelming influence; as, one <u>must</u> work or starve.
> **2.** To be essential or requisite; be logically or morally necessary; be necessary as a point or conclusion since a given antecedent is a fact; as a judge <u>must</u> be upright. If the men had been captured, it <u>must</u> have been since the captain's departure.

[82] According to the implications of Bekker's (2012) view on the extended formation of SAfE, Schneider's (2003; 2007) Phase Two, 'exonormative stabilization', could be said to have only really started around the 1930s, which means that British norms reigned well into the 20th century (see § 2.3.3.3).

The first sense is clearly deontic and dynamic (root) and the second epistemic. From words such as, 'constraining authority' and 'overwhelming influence', the standard deontic meaning of *must* during this era can be regarded as conveying a high degree of modal strength. The dynamic meaning is suggested by the description of a 'physical compulsion' as the force of the obligation (linking with Webster's [1886] entry), which implies an internal or circumstantial need. With the choice of the verbs 'work' and 'starve', something of a negative tone underlies the examples for this sense, i.e. *must* may already have a negative connotation in the Standard English of the late 19th century (see § 2.4.2.2 for 20th-century connotations). In the second sense, the epistemic meaning is described as being 'logically' derived from 'facts', which lends it a more neutral tone than in the first sense, even if words like 'essential', 'requisite', 'morally' and 'conclusion' indicate that a judgement of some sort is made. Especially the use of 'morally' invokes a sense of social pressure. Hence, neither of these definitions of *must* appear to have very positive connotations. In the early 20th century the *Oxford Dictionary of Current English* (1919:535) again simply defines *must* as "[b]e obliged to (do)", which basically reverts back to the earlier definition by Johnson, implying the stronger sense of obligation, i.e. a high degree of social impact. In all these earlier dictionary entries, a high degree of modal strength is suggested (i.e. negative connotations and social impact) with regards to the deontic (root) meaning of *must*.

This subsection has shown that *must* originally denoted only deontic and dynamic (root) meanings in OE (together with permission and possibility senses), and continued to develop new deontic/dynamic uses in ME and EME, as well as in CE today. Epistemic meanings of *must* only arose during the ME period, and new uses in this semantic domain developed in LME in CE. This confirms that *must* has not lost its (root) deontic meaning since OE, as many other modals have (cf. Abraham, 2001:30; § 2.4.1.2), despite the fact that it has gained the epistemic meaning, i.e. it has remained polysemous. The entries in the OED Online (2012b) and the various other dictionaries from certain points in history (from the mid-18th to early 20th centuries) have revealed that deontic *must* (expressing obligation) is generally considered as conveying a high degree of modal strength. The connotations of *must* from the 20th century onwards will be further explored in § 2.4.2, but first, the next section will discuss the rise of the quasi-modals, in keeping with the idea of a

polysemous modal system.

2.4.1.3 The development of the quasi-modal class

Towards the end of the 14th century and the start of the 15th century, a new kind of modal class began to develop, which is yet another case of grammaticalisation (Biber *et al*., 1998:206). This modal class goes by a plethora of different terms in the literature, but the most commonly used terms are 'quasi-modals', as preferred by Hopper and Traugott (2003), Collins (2009a) and this thesis, as well as 'semi-modals', as used by Biber *et al.* (1999) and Leech *et al*. (2009). Other terms include the 'emergent modals' (Krug, 2000) and 'marginal modals' or 'semi-auxiliaries' as used by Quirk *et al.* (1985), the classifications of which they employ "in a more syntactically restricted way" (Leech *et al*., 2009:91). Leech *et al.* (2009:91) note that 'semi-modal' (or 'quasi-modal' for that matter) is not a precise term syntactically speaking, seeing that it "refers to a loose constellation of verb constructions", but they do not enter into a debate on terminology, seeing that their focus is primarily on modal meaning, as is indeed the focus of this thesis. The emergence of this class is briefly discussed below.

NEED to was the first quasi-modal to emerge during the period mentioned above, around the turn of the 15th century (alongside *ought to*, which is generally treated as a modal rather than quasi-modal), with *HAVE to* emerging some time after 1500 and other modals only appearing more recently, e.g. *BE supposed to* after 1650 and *(HAVE) got to*, as well as *WANT to* in the 1800s (Biber *et al*., 1998; Krug, 2000:119). Despite their relatively short history, quasi-modals have become increasingly productive in British English, especially during the 20th century (Biber *et al*., 1998:208) and indeed also in the present (Mair & Leech, 2006). Their increasing productivity has however been largely register-specific, having higher frequencies in drama across all periods (as a simulation of spoken data) than in written texts (Biber *et al.*, 1998:208-209). Collins (2009:159) reports that while quasi-modals occur more frequently in spoken AmE, AusE and BrE, the difference is much greater in the case of AmE than the other two varieties. This indeed suggests that modal evolution may

be linked strongly with register – an important consideration for this thesis, to which I will return in the next section of this chapter.

The traditional view of *NEED to* is that *NEED*, originally a lexical verb in OE and ME (Warner, 1993:203), was impersonal in OE, but moved into various personal constructions in ME (Visser, 1969:1345) and "in the sixteenth century it starts to show modal characteristics" (Warner, 1993:203). Example (8) shows the use of *NEED* with the *to*-infinitive form dating from 1551:

(8) What **needeth to** alledge one sentence of him?[83]

Krug (2000:202) however found that, in Shakespeare, during which time e.g. Warner identifies *NEED* as "showing incipient modal behaviour", "modal constructions by far outnumber main verb constructions", but this trend has reversed in contemporary English.

The traditional account of how deontic *HAVE to* has evolved (see e.g. van der Gaaf, 1931; Visser, 1969; Fleischman, 1982; Brinton, 1991 and Heine, 1993) assumes five stages of grammatical re-analysis from a construction denoting a possession meaning to a modal meaning (as summarised in Krug, 2000:55; cf. Heine, 1993:41). In stage 1 the construction *I have a letter* represented a possession schema[84] (as in OE *HABBAN*), whereas in stage 2 the construction *I have a letter to mail* involved the added layer of a purpose/goal adjunct to the possession schema. Stage 3 involved that the possessive meaning of *HAVE* has been bleached out in the construction *I have a letter to write*, but it was only during stage 4 that *HAVE to* started to function as a unit lexeme which expressed the modal notion of obligation, e.g. *I have to write a letter.* In a construction like *I have to write*, in stage 5, the object complement could now be deleted and the construction is fully re-analysed and grammaticalised. The first attestations of the contiguous construction arose in ME, and in EME it was still rare, but a slow, gradual increase in its use ensued in the 19th century with rapid growth toward the end of the 19th and into the 20th century (Krug, 2000:74).

---

[83] Source: R. Record's *Pathway to Knowl.* (OED Online, 2013).

[84] Crystal defines 'schema', a psycholinguistic term, as "a mental structure in which knowledge is organized. SCHEMA THEORY has been developed to explain how people use background knowledge to shape their expectations about what a text (spoken or written) will contain. Readers create mental models (**schemata**) which they actively use to make sense of a text" (Crystal, 2008:424; emphasis original).

Krug (2000:61) proposes that the conceptual development of deontic *HAVE to* followed the tendency that "meanings based in the external described situation" evolves into "meanings based in the internal (evaluative/perceptual/cognitive) described situation" (Traugott, 1989:34). The cognitive processes in the evolution of obligative *HAVE to* therefore involve the notion of subjectification, which can be seen as a specific kind of metonymy (Traugott, 1996:5). She states the following:

> If the meaning of a lexical item or construction is grounded in the world of reference, it is likely that over time speakers will develop polysemies grounded in their world, whether reasoning, belief, or metatextual attitude to the discourse. In other words, subjectification is a semasiological development of meanings associated with a meaning-form pair such that the latter comes to mark subjectivity explicitly. (Traugott, 1996:5)

After the deontic *HAVE to* construction had evolved, it also developed an epistemic meaning. Traugott (1988:411) states that in the development of epistemic meaning from volitional or deontic meaning, "there is strengthening of the subjective element, and of focus of belief and knowledge". She provides the following example:

> "...if I say *You had to go* in the obligation sense, I invite the inference that I believe you did go. Therefore, in *You had to have gone*, derived from *You had to go*, the inference of the speaker's belief in the truth of the complement is strengthened." (Traugott, 1988:411)

Krug (2000:61;75) notes that possessive *HAVE got* surfaced during the later 16th to early 17th century, but that the obligation reading thereof only came into use in the 19th century (as mentioned above (cf. Biber *et al*., 1998)), and rapidly increased during the 20th centaury. Krug (2000:63) however argues that *HAVE got to* did not undergo the same kind of development as *HAVE to*, which is described above, in the sense that it started out with less clearly deontic semantics. In its frequent collocation with personal pronouns in the 19th century, all forms of *HAVE* in the construction were consistently used in enclitic form (pronounced with so little emphasis that it is shortened[85]), as in the contracted construction *you've got to*. Krug (2000:64) suggests that because of these contractions, "it is not unlikely that by semantic analogy (i.e. on the model of) the older have to construction have got extended its range of complements from nominal complements (as in *I've got a house*) to infinitival ones

[85] Concise Oxford English Dictionary (2006:469).

(as in *I've got to go*)". Krug further notes that "deontic readings of *HAVE to* carry much more semantic weight than auxiliary uses" and that deontic and possession meanings are very distinct, which for reasons of information processing leads to phonological reduction (2000:64-5). However, the resulting "high currency" of personal pronoun plus *HAVE*, "automatically triggers their coalescence" and may therefore explain why *got* has been inserted to reinforce the cases where *HAVE* has been reduced (2000:65).

The development of *WANT to* is a remarkable case, seeing that a verb or noun *WANT* did not exist in OE, when the volitional (pre-)modal verb was *WILLAN* (Krug, 2000:119). Krug (2000:119;165) reports that it first entered the language during the early Middle English period as both a transitive verb and a noun meaning 'lack', which is supported by the fact that *WANT* was borrowed from Old Norse (cf. Allen, 1995:224). The earliest uses of *WANT* in ME are commonly impersonal, as seen in (9), which was the "typical construction of its Old Norse ancestor *vanta*" (Krug, 2000:121).

(9) For ther nys no creature so good that hym ne **wanteth** somewhat of the perfeccioun of God, that is his makere.[86]

'For there is no creature so good that it does not **want** [lack] some of the perfection of God, who is his maker'.[87]

This is paralleled to some extent by the situation of many other English modals, e.g. *may, shall* and *will*, which also occurred with impersonal constructions in OE before they acquired all the morphosyntactic criteria to make them 'proper' modals (Warner, 1993:122; Krug, 2000:121). One point on which *WANT to* (and *HAVE to*) differ from the central modals is that they need *do*-support in e.g. negation (Krug, 2000:165). It is only toward the second half of the 18th century that "the modal construction – that is volitional semantics followed by the infinitive – became substantially more common" (2000:165).

Considering semantics, the parallel development of modals and quasi-modals had the effect that a modal such as *must* marked a strong sense of personal obligation

[86] From Krug (2000:122), taken from the Helsinki corpus.

[87] My translation.

or logical necessity, while *should* and *ought to* marked a weaker sense of this. Similarly the quasi-modals *HAVE to, (HAVE) got to* and *NEED to* mark a stronger sense of personal obligation and logical necessity, whereas *BE supposed to* and *had better* mark a weaker sense (cf. Collins, 2009:78). Therefore, *must* together with *HAVE to, (HAVE) got to* and *NEED to* are related semantically, just like *should*, *ought to, BE supposed to* and *had better* (cf. Biber *et al*., 1998:205). Tagliamonte (2004:34) summarises the development of deontic modality in English in terms of the grammaticalisation of *must* relative to *HAVE to* and *(HAVE) got to* (cf. Jespersen 1961:51-54, Visser 1963-73:1478-9, Warner 1982:163). Tagliamonte (2004:34) writes, as was also noted previously in § 2.4.1.2.1, that in OE *mot* (*must*) primarily denotes deontic (obligation) meanings along with permission or possibility readings, but by the ME period it develops the epistemic meaning and loses its permission reading. In ME *HAVE to* is rare, but already present. During the 19th century *HAVE to* is attested and *(HAVE) got to* develops, and by the 20th century *must* is already seen as 'old' or established, as the contracted form *'ve got to*, which is considered as 'American', becomes present, and truncated *got to* also develops (2004:34). From this summary it is clear that much grammatical development has occurred in the expression of deontic modality, where quasi-modals *HAVE to* and *(HAVE) got to* have entered the domain of *must* and became increasingly frequent in the 20th century. This parallel development of modals and quasi-modals also applies to other senses such as volition/prediction and permission/ability[88]/possibility. The quasi-modals therefore function on the same semantic framework as the central modals due to generally shared meanings, even though, on a microsemantic level, they denote varying nuances of meaning. This often leads to the modals and quasi-modals being in competition with each other. The next section explores the more recent changes among the modal and quasi-modal verbs in the large native varieties of English.

[88] Aijmer (2004), in her typological exploration of the ability meaning, regards *BE able to* as having an extra layer of modal meaning on top of ability, namely actuality, which links it to aspect and therefore renders it different in nuance to *can*.

2.4.2 Modality and English varieties

The system of modality has been analysed comprehensively within major inner-circle English varieties in recent research (e.g. Collins, 2009a and 2009b; Leech, 2003; 2011; Mair & Leech, 2006; Leech & Smith, 2006, 2009, Leech *et al.*, 2009; Myhill, 1995; Smith, 2003). The first part of the discussion will be on the general 20th-century trends reported by Leech and his associates (e.g. Mair & Leech, 2006) for British English and American English (BrE and AmE respectively), including the semantic trends related to monosemy and polysemy, and the second part of the discussion will focus on the synchronic trends reported by Collins (2009a) for BrE and AmE, as well as Australian English (AusE).

2.4.2.1 General trends in native Englishes

Short-term diachronic studies in American English (AmE) and British English (BrE) using the LOB/F-LOB and Brown/Frown sets of written corpora (hereafter referred to as the Brown corpora) report a general decline in the frequency of the modal verbs or modals and a general increase in the use of the quasi-modals or semi-modals over the course of the 20th century, from 1931 and 1961 to 1991 as chronological data points (Leech, 2003; Mair & Leech, 2006; Leech & Smith, 2009). The decline of modal verbs is also confirmed by Leech (2011) for BrE and AmE across a longer time span, from earlier in the 20th century (1901/1910s) and into the 21st century (2006/2000s). With no solid evidence of interdependency between the rising and falling frequencies of these kinds of specialised auxiliaries (Leech, 2003:235; Mair & Leech, 2006:327), the reasons behind these changes were linked to tendencies towards colloquialisation (the stylistic movement towards more oral patterns of language) and democratisation (the social trend towards avoiding claims to power or authority) (e.g. Leech, 2003:236-7).

Christian Mair in his book *Twentieth-century English: History, Variation and Standardization* (2006:183) describes colloquialisation as the "linguistic correlate" of "a trend towards informality" that has characterised "the social history of the past century", as manifested in the "narrowing of the stylistic gap between speech and

writing". Indeed Biber and Finegan (1989) found that over the past two centuries the written genres of essays, letters and fiction have all undergone a 'drift' toward a more oral style, as also noted in § 2.2.2.1. Biber and Finegan suggest that the motivation for this includes the following, quoted from the summary by Mair (2006:184): "an aesthetic preference for colloquial or plain styles shared by such unlikely allies as the seventeenth-century experimental scientists who founded the Royal Society[89], the late eighteenth-century Romantic poets, and the American nationalist Noah Webster", as well as "a changing demography, in which schooling became available to wider sections of the population and writing ceased to be the elite pursuit that it had been" (cf. Biber & Finegan, 1989:512-4). Biber (2003:169) describes these changes since around 1600 as follows:

> "Written registers in English have undergone extensive stylistic change over the past four centuries. Written prose registers in the seventeenth century were already quite different from conversational registers, and those registers evolved to become even more distinct from speech over the course of the eighteenth century. … However, over the course of the nineteenth and twentieth centuries, popular written registers like letters, fiction, and essays have reversed their direction of change and evolved to become more similar to spoken registers, often becoming even more oral in the modern period than in the seventeenth century."

This drift toward oral styles has accelerated toward the second half of the 20th-century, which Mair (2006:185) suggests is a reflection of "an unprecedented mobilization of formerly relatively stable class-based hierarchies in large parts of the English-speaking world". He further notes that the findings of linguists regarding "the modern preference for colloquial over elaborated[90] style, for informal over formal conduct, and for spontaneity of expression over ritual and custom" "converge with the assessments of cultural historians, social theorists and critical discourse analysts" (Mair, 2006:185). Terms such as 'technologisation', 'democratisation' and 'informalisation' also characterise this general tendency, with the widespread availability of technology for communication in some ways giving rise to the "eroding of the power of traditional elites", motivating the projection of

---

[89] According to The Royal Society's website *www.royalsociety.org* it is a London-based "[f]ellowship of the world's most eminent scientists" founded in 1660 by King Charles II.

[90] Here writing is perceived as an elaboration on the natural linguistic 'baseline', which is informal face-to-face conversation (Mair, 2006:183).

conversational discourse from the private into the public spheres (Mair, 2006:186-7; cf. Fairclough, 1992:98). These terms and the processes they denote are complementary to the degree that the term colloquialisation integrates the concerns of "historical corpus-linguists, cultural historians, and critical discourse analysts" (Mair, 2006:187). The term 'colloquialisation' is indeed first introduced by Mair (1997), where in linguistic terms regarding twentieth-century English it refers to the stylistic shift firstly "away from a written norm which is elaborated to maximal distance from speech and towards a written norm that is closer to spoken language", and secondly, "away from a written norm which cultivates formality towards a norm which is tolerant of informality and even allows for anti-formality as a rhetorical strategy" (quoted from Mair, 2006:187). Regarding chronology and the process of colloquialisation, Mair (2006:187) further notes the following:

> "...the colloquialization of the written language is a development which is in evidence to a greater or lesser extent in different communities synchronically, and at different times diachronically. One thing, however, is certain, too: it will never work itself out to the limit, as a written norm which is identical to colloquial speech would be highly dysfunctional."

This is because written language, as an elaboration on informal conversation, is motivated functionally, seeing that, for example "the fact that writing usually involves communication across space and time will make it necessary to spell out references to the context of situation which may be left implicit in speech" (Mair, 2006:183).

This thesis therefore does not consider colloquialisation to be an endpoint of informal language or simply a result of merging registers, but rather a process toward more informal language, sometimes embodied in a narrowing gap between stylistic variation in the registers (but with the different functions of the registers motivating this gap to remain present, even though its extent is reduced). This is of course providing that the direction in which the registers are 'narrowing' in terms of their stylistic gap is speech-like and more informal. If the latter is not true, the move toward register convergence may be an independent trend from that of colloquialisation.

The suggestion of colloquialisation enjoys some implicit support from Collins' (2009a) findings related to contemporary modality within the large inner-circle varieties used in the studies of e.g. Leech and his associates, plus Australian English (AusE). Collins (2009a:9) generally reports that the quasi-modals are more frequent

in the spoken than in the written register, which is true for all the relevant native varieties. This quantitative finding, together with the stable, widespread rise of these auxiliaries across both written and spoken registers mentioned above, might point toward the increasing preference or popularity of a speech-like style of language (as also noted by e.g. Mair (2006)), which involves a usage preference shift from modals to quasi-modals. The general pattern reported by Collins (2009a) is however not as simple as this, as the overall ratios between the modals and the quasi-modals were found to vary within varieties; AmE enjoyed a more marked "stylistic gulf between semi-modals and modals" (2009a:9) and BrE less so, while the differential of AusE emerged between that of the two other varieties. AmE can therefore be said to enjoy the highest degree of colloquialisation among these varieties.

The patterns reported by Leech and his associates are of course more complex than the mere rise of quasi-modals and the decline of modals, seeing that different tendencies emerge within the three semantic groups or clusters of modality. The more general trend corresponds with the synchronic quantitative patterns of the obligation/necessity cluster, where the quasi-modals *HAVE to* and *NEED to* have superseded *must* and *should* (Mair & Leech, 2006; Collins, 2009:34). However, it is not compatible with the other two semantic clusters, volition/prediction and permission/ability/possibility, where modals are still very prevalent in contemporary data, e.g. *will* and *would* are more than double the frequency of *BE going to* and *WANT to* respectively, and e.g. *can* almost enjoys ten times the usage frequency of *BE able to* (Collins, 2009a:92;126). On top of this, these quantitative semantic patterns are not equivalent to the behaviour of all individual modals or quasi-modals.

It was found that the high-frequency modals, e.g. *will, would* and *can,* suffered the least in terms of a frequency drop, whereas the low-frequency modals, e.g. *shall, need* and *ought,* have diminished the most, and the mid-frequency modals have experienced a moderate degree of frequency decline (Leech, 2003; Mair & Leech, 2006; Leech & Smith, 2009; Leech *et al*., 2009). Within the obligation/necessity cluster, and even beyond, the most salient modal in terms of frequency drop is *must.* A mid-frequency modal by the first half of the 20th century (with higher social impact than *should*), *must* decreases by 29% in BrE and 34,4% in AmE from 1961 to 1991/2 (Mair & Leech, 2006:327), which is a steeper decline than for *should* (having decreased by 11,8% in BrE and 13,5% in AmE), plus *could* and *might* (Smith 2003).

The decline has occurred to such an extent that today *must* is even frequently counted among the low-frequency modals in AmE, BrE and AusE (Collins, 2009a:43-44; Mair & Leech, 2006). Leech (2011:551) indeed confirms the continued decline of *must* in BrE into the 21st century: it dropped from 809 tokens per million words (henceforth [pmw]) in 1991 to 533 tokens (pmw) in 2006. This (533 tokens) is less than half of its original frequency in 1901, which was 1185 instances (pmw).

It is apparent that the decline of *must* has been the greatest for AmE, according to e.g. Mair & Leech (2006). In his study Leech (2011) indeed found evidence of a statistically significant decline of *must* in AmE across a larger time span than the previously-mentioned studies: this modal shows a steady decline across the 20th century and into the 21st century, e.g. from 1165 tokens (pmw) in the 1910s to 405 tokens (pmw) in the 2000s. Myhill (1995:162) also confirms the strong decline (almost by half) of this modal in AmE – even since earlier periods than Leech and associates, as his study covers the 19th to 20th centuries – but moreover explores the reasons behind the trend. This retreat of a socially 'charged' modal traditionally denoting strong obligation (and often negative commands in AmE) can be seen as an effect of democratisation (Myhill, 1995:166-167). This effect, as touched on above, tends to avoid face-threatening[91] language and hence the creation of social distance – in this case notably by means of the more "objective" quasi-modal *HAVE to* (Collins, 2009a:60), which has increased by 9% in BrE and 1,1% in AmE, as well as the most salient quasi-modal in terms of frequency escalation, *NEED to,* which rises by 249,1% in BrE and 123,2 % in AmE (Mair & Leech, 2006).

As a comparison with Leech (2003), Neil Millar (2009) undertakes an investigation into modal frequency by means of the TIME Magazine Corpus ranging from 1923 to 2006, which reaches further back, as well as further into the future, than the earlier studies of Leech and associates (e.g. Mair & Leech, 2006). The TIME Corpus consists of 100 million words, which is clearly much larger than the corpora used by other scholars in this regard. His results are "largely congruent with Leech's (2003) exploration of the Brown corpora" (2009:251), in that semi-modals or quasi-modals increase (2009:204), but he does not find an overall decline in the modals and rather an increase of 22,9%, which contrasts with Mair and Leech's (2006) findings that modals diminished overall by 9,5% in BrE and 12,2% in AmE, and that most

[91] See § 2.2.1.2.2 (cf. Brown & Levinson, 1987).

modals displayed some degree of decline. Millar, by contrast, reports that the modals *can, could* and *may* show a definite increase. *Can* steadily increases by 113,4% over the span of the 20th century, whereas *could* increases by 103% and *may* by 59,7%. On the other hand, Millar (2009:199; 251) only detects a clear decline in the three modals *shall, must* and *ought. Must* declined by 48,4% from 1923 to 2006, with the decline accelerating toward the end of the 20th and beginning of the 21st century (2009:199). *Shall* has gradually declined by a massive 95,6% from the 1920s to 2000s and *ought* has gradually declined by 44,9% (2009:199).

Millar suggests the reasons behind these declines "to be shifts towards forms previously associated with spoken language, shifts from formal towards informal, and shifts towards forms that are more inclusive" (2009:251) and he therefore concurs with Leech and associates on the matters of colloquialisation and democratisation. However, Millar suggests a new (tentative) hypothesis, viz. "that socio-political or socio-historical factors may play a role in the types of modality being expressed and this may be reflected in fluctuations in the frequencies of particular modals" (2009:251) – an important consideration for this thesis. Millar (2009:216) however refers to Leech's notion that generalising corpus findings to represent a larger population "must be regarded largely as an act of faith" (Leech, 1991:27) and adds that both the TIME corpus and the Brown family[92] of corpora (as used by Leech and associates for AmE) have many weaknesses along with their strengths, so the results form either of these sets of studies should rather not be seen as absolute.

However, Leech (2011), in his response to Millar's paper, showed that the modals "have indeed been declining in frequency", by means of an enlarged set of BrE and AmE corpora to cover a larger time span, and consequently found that Millar's conclusion is "unjustified" (Leech, 2011:547). Leech (2011:549) criticises Millar for his assumption that "*if usage of modal verbs is declining in the language as a whole, it is expected that this will be reflected in the TIME magazine corpus*" (Millar, 2009:194; original italics). Leech however notes that Millar's results are based on "only one diachronic corpus consisting entirely of TIME Magazine material,

[92] The Brown family of corpora includes LOB (the Lancaster-Oslo/Bergen Corpus) that represents the written BrE of 1961 and its Freiburg update, F-LOB, representing 1991/2 of the same variety, along with the Brown corpus (Standard Corpus of Present-Day American English) that represents the written AmE of 1961 and the Freiburg-Brown corpus of American English (Frown), representing 1991/2 of the same variety. More information on these corpora are available from: http://www.helsinki.fi/varieng/CoRD/corpora/index.html.

which cannot represent the (written) language overall, but only one magazine which is well known, incidentally, to have a distinctive style" (Leech, 2011:548). The Brown family of corpora used by Leech and associates however include various registers. Millar implies that the limitation of Leech's (2003) study, which showed the modals to decline, is its use of only two data points (1931 and 1991) (the TIME Corpus has 83 data points) and that this could "present an inaccurate picture of the overall trend (Millar, 2009:191), but Leech argues that the difference between these two studies is "much more likely to be due to the limitation of Millar's study to one very particular genre" (Leech, 2011:549). Leech evaluates what he calls one of the 'flaws' in Millar's argument:

> "Where I believe Millar goes wrong is in attributing to the TIME Corpus a representativeness of the language that it just does not have. He rightly commends the TIME Corpus for its size (over 100 million words), internal consistency, representativeness (by which he means representativeness of TIME magazine, not of the language as a whole!), temporal span (covering the period 1923-2006) and the short chronological intervals between data points it allows (in his case one year). However, all these virtues, although they may provide excellently robust results for those wishing to study the evolving language of TIME Magazine, do not justify any assumption that if the English of TIME Magazine is changing in a certain direction, the same change will apply to English in general." (Leech, 2011:549)

Leech's (2011:550-554) results for BrE are however based on the inclusion of, firstly, a third corpus on top of LOB and FLOB, namely BLOB, covering 1931, which was also previously included by Leech & Smith (2009) and Leech *et al.* (2009), secondly, the BE06 corpus which includes data from 2006, and thirdly the BLOB-1909 Corpus. This means that Leech's study now covers five data points for BrE (1901, 1931, 1961, 1991 and 2006). For AmE Leech used two recent online corpora: COHA (Corpus of Historical American English) and COCA (Corpus of Contemporary American English), which together provide 10 data points: each decade from the 1910s to 2000s. The results obtained through these corpora confirm the overall decline of modals from the second half of the 20th century and show that this trend extends at an accelerated rate into the 21st century. The modals of BrE decline by 16,3% from 1931-2006, and by 13,14% in AmE from 1960-2009 (Leech, 2011:550-554).

Nevertheless, Leech (2003) and Millar (2009) do agree on the matter of frequency group shifts, as also mentioned earlier, which involve that the most frequent modals (including *can, could, will* and *would*) are becoming more frequent

and the least frequent ones (e.g. *shall, ought, need*) are becoming even less frequent. But Leech (2011:555) notes that the difference in their findings relates to where the boundaries of these frequency groups are. Leech's results show that, for both varieties, the high-frequency modals *can, could, will* and *would* remain generally stable across the 20th century, but that, "from the seven remaining modals", especially *must, may, shall, ought* and *need* have all declined quite steeply (2011:556).

There is therefore evidence from various corpus-based studies (Leech, 2003; 2011; Mair & Leech, 2006; Leech & Smith, 2009; Leech *et al*., 2009; Millar, 2009 and Myhill, 1995) that *must* is in serious decline since the 20th century in the large native varieties of British and American English. Myhill's (1995) proposal that *must* is declining due to its connotation with face-threatening meanings (an effect of democratisation), suggests that semantics are important to consider when explaining frequency change. Indeed Leech *et al*. (2009:83) suggest that certain modals and quasi-modals may be on a path of semantic 'restriction' within their broad semantic groups or clusters. Such a suggestion may in turn point toward the fact that modals and quasi-modals are or were semantically interchangeable to a degree at a point in time before moving toward semantic refinement (cf. Leech and Coates, 1980). The question now turns to whether these kinds of suggestions indicate an ongoing shift from modal polysemy at one stage, toward modal monosemy at another.

2.4.2.2 Trends related to polysemy and monosemy

Apart from these more recent events, the development of the modal system throughout the history of English, long before those periods covered by Leech and associates, is a well-known, intricate tale of interactions between the forces of change and stability. The system's intricacy additionally lies in the interaction between its grammar and semantics (Warner, 1993:13), which is most notably visible in the rise of the quasi-modal grammatical class during the late Middle English period to the early Renaissance, lasting even until the early 19th century (Biber *et al*., 1998:208), which progressively moved into some of the semantic domains of the central modals that have already been there since Old English (Biber *et al*., 1998:206; Warner, 1993:192) (as also mentioned earlier).

Two especially frequent and salient modal verbs in Middle English were *must* and *should*, (Biber *et al*. 1998:206) – the broad meanings of which have over time come be expressed by e.g. the quasi-modals *HAVE to* and *NEED to*. These and other semantic overlaps within and among the central modal and quasi-modal classes naturally lead toward a modal polysemy over time, in that many different levels of meanings within relative semantic confines or semantic clusters can be conveyed by a single modal or quasi-modal, depending on the context of use. As mentioned earlier, Leech *et al.* (2009:70-90;107-117) examine the reasons behind the drop of particular modals and the rise of particular quasi-modals, in order to account for the broad tendency between the modal and quasi-modal usage differential. These examinations all denote tendencies toward some degree of interplay between levels of polysemy and monosemy, which are also noted by e.g. Smith (2003) and Myhill (1995).

It was found that the lowest-frequency modals, *shall*, *ought* and *need* may be undergoing so-called "paradigmatic atrophy" (Leech *et al*., 2009:80-1), in that they are restricted to very specific uses and linguistic-internal patterns, i.e. a strong degree of monosemy, e.g. the rarity of *shall* is attributed to its virtual restriction to either first-person subjects, archaic uses in literary quotations or administrative language (i.e. "distributional fragmentation"), whereas the increasing rarity of the negative constructions *shan't* and *oughtn't* contribute to the scarcity of *shall* and *ought*, and, contrastingly, *need* is restricted only to be used in its negative construction form, *needn't*, within non-assertive contexts (Leech *et al*., 2009:80-82; cf. Van der Auwera *et al*., 2012).

For the quasi-modals or semi-modals, signs of abstraction and generalisation or semantic weakening are becoming more prevalent as part of the process of grammaticalisation (2009:107-177). *HAVE to* and *(HAVE) got to* both tend toward expressing more abstract meanings (2009:109), whereas various other quasi-modals are all in the process of acquiring generalised meanings, which can produce monosemy, and less frequently even polysemy and as side-effects. For example, *BE going to* often enjoys a future-referring function, *NEED to* often loosens the link with the subject referent via the passive voice and expands into the domain of *HAVE to, BE able to* often expresses ability meanings and expands into the deontic domain of *NEED to*, and *WANT to* extends into the deontic domain of *NEED to* and *ought to* (2009:107-114).

The most relevant and interesting observations regarding a trend toward

monosemy entail the changes in the 'ecology' and thus the function of individual modals and quasi-modals of the permission/possibility/ability cluster and, more importantly, the obligation/necessity cluster (2009:83-89;117). Leech *et al.* (2009:114) define the concept of ecology as "the idea that each form evolves its own niche in the expression of modality, expanding, contracting or maintaining its 'habitat' in relation to other, partially competing, forms"[93].

The use of the term 'ecology' in this respect should not be confused with Schneider's (2003) use of the same term to denote his different strands of communities in the evolution of a New English. In its various discussions, this thesis will maintain the distinction between the uses of these two terms by referring to either 'social ecology' (in Schneider's [2003] terms) or 'semantic ecology' (in Leech *et al.*'s [2009] terms). The changes in semantic ecology in Leech *et al.*'s (2009) terms can firstly be seen within and among the declining mid-frequency modals, *may, must* and *should*, where *may* and *should* are found to portray noticeable trends toward monosemy, and where *must* is becoming somewhat more restricted to context and therefore also moves toward occupying its own niche (2009:83-89).

May moves toward predominantly expressing epistemic meanings, with the permission meaning still present, but the root possibility meaning yielding ground to *can*. Similarly, Millar (2009:203) finds evidence of an increase in epistemic possibility meanings in *may* from 1923 to 2006, and a sharp decline in root possibility meanings, as well smaller declines in permission, quasi-subjunctive and other unclear meanings. The epistemic possibility meaning of *may* is illustrated in (10), and the root possibility meaning of *can* in (11) (taken from Leech *et al*. [2009:83;85]):

(10) We **may** be able to hide you.

(11) Nomia melanderi **can** be found in tremendous numbers in certain parts of the United States.

---

[93] Leech *et al.* (2009:114) choose to analyse four modals (*must, should*, *ought* and *need*) and four quasi-modals (*HAVE to*, *(HAVE) got to*, *NEED to* and *(had) better*) from the obligation and necessity cluster in terms of semantic ecology, which contrasts with the traditional view of S-curve progression, where one of two competing forms "finally yields the field to the other". They note that this kind of progression is "far from the truth in this case" (2009:114).

Must shows a clear decline in the frequency of expressing deontic meaning as a by-product of its sharp decline in frequency to avoid such meaning, i.e. collocating with first and second-person subjects (where the source of subjective imposition in the text or discourse is clearly the speaker him/herself) (Smith, 2003:242; Leech *et al.*, 2009:114-116), which does not mean that it now expresses less deontic meaning, but that the restriction to deontic or subjective meaning has lessened its frequency. Deontic *must*, as in (12), has therefore become highly discourse-oriented (Palmer, 1990:69-70) and hence increasingly monosemous, which links with the notion of democratisation (Leech *et al*., 2009:88; Collins, 2005). *Should* is indeed found to be a beneficiary of the decline of *must*, in that *should* progressively expresses 'weaker' deontic meanings (13) than *must* (2009:86-7).

Must has furthermore been found to incline towards expressions portraying epistemic necessity (14), whereas *should* experiences a decline in its epistemic meaning as a result (Leech *et al.*, 2009:115-116; Collins, 2005). The movement toward epistemic expressions in *must* has been partially influenced by its deontic restriction (and hence decline) and partially because there is no other obvious alternative expression with the same propensity for expressions of epistemic necessity, e.g. *HAVE to* (2009:87-9). These examples are taken from Leech *et al*. (2009:86-7)

(12) That woman **must** go!

(13) They **should** be kept out of reach of children.

(14) I thought I saw the cook, then I decided I **must** be mistaken.

Similar to the trends mentioned above, Millar (2009:203) most notably reports that the weak obligation (root/deontic) meaning of *should* has increased steadily toward the 2000s in the TIME Corpus, where it comprises 80% of meanings. On the other hand, Millar (2009:203) also finds evidence of *must* shifting very gradually toward epistemic meanings across the 20th century, but *must* only starts to move away from deontic meanings after the 1960s. By the 2000s the root deontic meaning of *must is* still the highest (67%) in the period, but has become less so since the 1920s.

It is interesting to note that deontic *must* has become so discourse-oriented (cf. Palmer, 1990:69-70) as to lose its place among the mid-frequency modals. It was

noted earlier in this main section that the modals have generally "lost their deontic/root meaning to a great extent" since the Anglo-Saxon era, except for *must* (Abraham, 2001:30; cf. Abraham, 1998, 2002), to avoid recourse to authority (Conradie, 1987:179). It appears that, because *must* retained its deontic meaning to a larger extent than the other modals, it is now falling out of use and, when it is used, more often denotes epistemic meaning. In some ways there is a suggestion of a trend toward monosemy here, but not at all to the same extent as for e.g. *should*. Leech *et al.* (2009:87-8) confirm that *must* has "not been moving appreciably towards monosemy", because of its restriction to certain contexts due to democratisation.

Similarly, Close and Aarts' (2010:179) results from the Diachronic Corpus of Present-Day Spoken English (DCPSE) "show that even in the 1960s data, the frequency of root *have to* has surpassed that of root *must*", and that "[t]he continued decline of *must* and the rise in *have to* in the 1990s data suggests that *have to* is being used in contexts where previously we would have found *must*". They further report that "although both epistemic and root senses of *must* have declined, as proportions of the total number of instances of *must* they have remained fairly constant" (2010:179). In the DCPSE "*have to* is the most frequent marker of root modality, which is remarkable because we might expect the newer form *have got to* to be more frequent", and "the use of epistemic *must* has decreased, but the use of epistemic *have to* has not increased" (2010:179). This indeed "supports Leech's (2003) claim that *must* is not becoming monosemous" (2010:179).

The quasi-modals *HAVE to* and *NEED to* however display a trend towards monosemy, as both appear to have found their ecological niche – where *HAVE to* has a tendency to express general root modality (as also noted by Close and Aarts [2010] above) and is hence becoming less face-threatening (15), and *NEED to* is becoming even more face-saving (16), insofar that it is never replaceable by *must* (Leech *et al.*, 2009:115; Collins, 2005). Smith (2003:242) adds that *HAVE to* tends to characteristically associate with objective or external meanings (as also noted earlier [cf. e.g. Mair & Leech, 2006; Collins, 2009a]), where the source of imposition is clearly not the speaker – contributing to this modal becoming more face-saving. The following examples are from Leech *et al.* (2009:115).

(15) Only the electric bill. It's up again. We'll **have to** go easy on the immersion heater next quarter.

(16) ...hard questions **need to** be asked about the Government's record since 1981, and the Socialist Party's leadership's current lack of direction.

When examining the internal dynamics of the obligation/necessity cluster in BrE and AmE across the 20th century more closely, Smith (2003:242) draws on the previously-noted assumption that some modals and quasi-modals are 'stronger' markers of the obligation/necessity cluster in terms of deontic meaning, which include *must, HAVE to, (HAVE) got to, need* and *NEED to*, whereas the 'weaker' markers are *should, ought to,* and *BE supposed to* (as also mentioned by Leech *et al*., 2009:86-87 in terms of *must* and *should*). This is similar to the description of the parallel semantics of these auxiliaries given in § 2.4.1.3. The basic distinction is contextual: the stronger the obligation, the more severe the consequences will be if it is not fulfilled (Smith, 2003:242), which links with the previously-mentioned criteria of Coates (1983) and with that of Depraetere and Verhulst (2008:15) regarding the gravity of noncompliance. Smith also indicates that there may be internal degrees of strong and weak obligation, in that the stronger uses of the weak forms (e.g. *You should get a move on*) may *occasionally* convey more force than the weaker uses of strong forms (e.g. *You must come and visit some time*) (2003:242). This is however not the general picture. Generally, this distinction between strong and weak obligation is proposed by Smith (2003:242) to be clearer than between subjective and objective use, which contrasts with Collins' (2009a) above-mentioned view. This reveals a relative monosemous nature of the modals and quasi-modals in the obligation/necessity cluster – even in its micro-semantics.

Along similar lines, but in terms of general trends toward monosemy in AmE across both the 19th and 20th centuries, Myhill (1995) asserts that for each modal that decreased in frequency, others (usually two) within its semantic cluster increased in frequency in order to 'replace' it during these periods (1995:163). More intricately, the 'old' and the 'new' modal or quasi-modal expressing e.g. strong obligation are increasingly used in dissimilar ways, e.g. strong obligation *must* tends to be related to social norms, while e.g. strong obligation *HAVE to* is typically connected with habitual obligations (1995:163). A trend toward micro-level monosemy can therefore be said to be evident in both BrE and AmE, across the last century and even beyond.

Based on the findings of Leech *et al*. (2009), Smith (2003) and Myhill (1995) (which do portray differing levels of interplay between trends toward polysemy and monosemy in other semantic clusters), the obligation/necessity cluster can be said to be most evidently moving toward monosemy in BrE and AmE, where each modal or quasi-modal tends to enjoy its own specific set of meanings or even micro-meanings, instead of sharing them amongst one another in a polysemous relationship. This may in turn influence individual frequencies of modals or quasi-modals and be influenced by the process of democratisation, as in the case of *must*. This kind of semantic shift toward monosemy suggests that the modal and quasi-modal verbs are on a path towards less fuzzy semantics, where meanings become more clearly distinguishable. The question now arises whether the same changes are taking place within the modal system of SAfE, which will be addressed in Chapter four.

As noted previously, diachronic change is manifested synchronically as variation, and it is therefore useful to examine the findings of Collins (2009a) next.

2.4.2.3 Synchronic variation in native Englishes

Collins undertakes a study of modals and quasi-modals in three major English varieties around the world, namely BrE (British English), AmE (American English) and AusE (Australian English) in his book *Modals and Quasi-modals in English* (2009a). Even though his exploratory study of dialectal and stylistic variation is mostly synchronic in nature, his investigations also touch on diachronic variation, resulting in, what Collins calls "the most comprehensive study yet attempted in the area" (2009a:1). In his book, Collins employs a million-word corpus per variety – the standard number of the ICE corpora. The International Corpus of English for Great Britain (ICE-GB) is used to represent BrE, whereas its counterpart for Australia, ICE-AUS, is used to represent AusE. A corpus resembling the ICE corpora in size and composition was compiled for English in the USA, which Collins calls C-US (2009a:1). ICE corpora typically consist of a spoken component, which includes dialogues (private or public) and monologues (scripted or unscripted), as well as a written component, including printed (e.g. news) or non-printed (e.g. letters) material (cf. Aarts, *et al.*, 2002:307-8). This section will only briefly give an account of the

main frequency patterns that Collins provides, as well as the very broad semantic patterns. Collins' semantic accounts will be explored in more detail in Chapter 4, when they are compared with those in SAfE.

Within what he names the necessity and obligation category[94], Collins (2009a:34) reports that *HAVE to* is overall the most frequent of the modals and quasi-modals (3 940 tokens [pmw]), with its highest uses in firstly AmE and secondly AusE. *Should* is the second most frequent modal in this cluster (3 115 tokens [pmw]), ranking the highest in AusE and second highest in BrE. Next on the frequency list is *must* (1 690 tokens [pmw]), which in all three varieties is basically two times less frequent than *should*. BrE displays the highest use of *must*, whereas AmE uses it less than *NEED to*, which is fourth on the overall frequency list (1 096 tokens [pmw]). The highest frequency of *NEED to* consequently occurs in AmE, and is followed by AusE, although *must* is still more frequent than *NEED to* in AusE. *(HAVE) got to* is next (844 tokens [pmw]), with its highest frequency detected in BrE, followed by AusE. This is followed by *BE to*, with 432 tokens (pmw), which is used the most in BrE (221 tokens [pmw]) – almost three times its usage frequency of AmE (76 tokens [pmw]). The fourth least frequent choice is *BE supposed to*, which is more widely used in AmE and BrE, but this quasi-modal only comprises almost a third (273 tokens [pmw]) of the frequency of *(HAVE) got to* when the frequencies of all three varieties are totaled. *Ought*, which is still used more in BrE than the other varieties, is next, (with a total of 167 tokens [pmw]), and has the smallest frequency in AusE. The second last auxiliary on the list is *(had) better*, with a total of 112 tokens (pmw). This quasi-modal has its highest usage frequency in AusE (48 tokens [pmw]). The least frequent modal in this semantic cluster is *need* (only 68 [pmw]), which is also mostly used in BrE. Hence, in order to express meanings of necessity or obligation, BrE and AusE both use *HAVE to* and *should* with almost equal preference, whereas AmE prefers *HAVE to* above *should*. This pattern corresponds with the findings of Leech and associates described in the above section that *must* has lost ground to the quasi-modal *HAVE to*.

Must mostly carries deontic meaning across all three varieties (57,3%), with the highest count of deontic uses in BrE. Epistemic uses were found in 32% of the overall count and the remaining 9,9% is divided among dynamic (6,3%) and

[94] Collins (2009a) adds the low-frequency quasi-modal *BE bound to* to the list of the obligation/necessity category, but it is not analysed in this thesis (cf. Biber *et al*., 1999).

indeterminate uses (3,6%). This contrasts with the semantic pattern found in *should*, where 69% of overall uses signified deontic meanings (with the highest count in AusE) and only 11,8% epistemic meanings. The remaining 19,3% of *should* is covered by the subjunctive meaning (6,6%), the meaning equivalent to *would* (2,7%) and indeterminate cases (10%). It is interesting to note the frequencies of deontic *must* (969 tokens [pmw]) and *should* (2 148 tokens [pmw]); deontic *should* is a little more than double the frequency of deontic *must* (Collins, 2009:35;45). Furthermore, *must* has higher productivity in epistemic uses than *should*, all of which corresponds to the semantic tendency found by Smith (2003:242) and Leech *et al.* (2009:114-116) that *must* is increasingly used in epistemic contexts. This case is strengthened by the even higher frequency of deontic meanings in *HAVE to* than *should*, which number at 2 999 tokens (pmw) and covers 76,1% of the semantics of *HAVE to* (Collins, 2009a:60). This particular quasi-modal expresses dynamic meanings in 22,4% of cases, as well as epistemic meanings in only 0,7% of cases (0,8% is indeterminate), which supports the trend that the epistemic 'weight' is shifting toward *must* and, consequently, deontic meanings have become more likely in *HAVE to* (Leech *et al.*, 2009:87-9).

Within the semantic cluster denoting meanings of possibility, permission and ability, *can* is by far the modal with the highest overall frequency (10 608 tokens [pmw]) and has similar frequencies in BrE, AmE and AusE (Collins, 2009a,92). *Could* is overall the second most frequent modal in this cluster (4 969 tokens [pmw]), with the three varieties once again displaying near equal frequencies. The third most common modal in this category is *may* (2 924 tokens [pmw]), but this auxiliary is used most frequently in BrE (1 218 tokens [pmw]) and bears near equal weight in AusE and AmE. *Might* is the second least frequent modal with 1 917 tokens (pmw) and enjoys almost equal distribution in AusE and BrE with a lower distribution in AmE. Last on the list is *BE able to* with the least overall uses (1 167 tokens [pmw]) and with a higher preference in BrE (434 tokens [pmw]).

Collins (2009a:92) reports a very high percentage of epistemic meanings in *may* (79%), whereas the remaining 21,0% is divided between deontic (7,0%), dynamic (8,1%) and indeterminate cases (5,9). This contrasts with the meaning distribution of *can*, which yielded 81% to dynamic meanings and the remaining 19% is divided between deontic (9,9%), epistemic (1,1%) and indeterminate meanings (8,1%). This pattern agrees with the report of Leech *et al.* (2009:83-89) that *may* is moving toward chiefly expressing epistemic meanings and that the (root) possibility

(deontic/dynamic) meaning of *may* is shifting to *can*. *Could* was found to mostly convey dynamic meanings (77,4%), with epistemic meanings expressed in 14,1% of cases, deontic in 2,9% and indeterminate cases in 5,6%. On the other hand, *might* most frequently denotes epistemic meanings (77,2%), which is followed by dynamic meanings (16,6%), indeterminate meanings (5,5%) and lastly deontic meanings (0,6%). The meanings of *BE able to* are divided as follows: in 42,9% of cases it expresses ability, in 47,5% of cases it conveys theoretical possibility, indeterminate cases occur in 5,2% of uses and deontic meanings in 4,4%.

The third semantic cluster Collins analyses, is the prediction and volition cluster. This cluster is "dominated by *will* and *would*, which together account for 75,2% of all tokens across the three corpora" (Collins, 2009a:125). This agrees with the patterns reported by Leech and associates that these two modals, along with *can* (as seen above), have suffered the least in terms of frequency drop across the 20th century. *Will* yielded 8 505 tokens (pmw) and has near equal frequencies across BrE, AmE and AusE, whereas *would* amounted to 7 775 tokens (pmw) and is more popular in AmE, with 4 001 compared to the 3 404 and 3 585 tokens (pmw) in BrE and AusE respectively (2009a:126). The third most frequent item in this semantic category is the quasi-modal *BE going to* with overall 2 721 tokens (pmw). There is clear popularity of use for *BE going to* in AmE (2 413 tokens [pmw]). *WANT to* is next with 1 897 tokens (pmw) for the three varieties collectively and this quasi-modal is, once again, popular in AmE (1 425 tokens [pmw]). Last on the list is *shall*,[95] overall totaling at only 343 tokens (pmw), with the highest frequency of use found in BrE (223 tokens [pmw]), which is not surprising given the more 'conservative' nature of this variety when compared to e.g. AmE (in terms of retaining a diminishing, archaic modal). The low frequency of *shall* found in Collins (2009a) corresponds to the pattern of decline found in the studies of Leech and associates, as mentioned in the previous section.

When considering the semantics of this cluster, *will* mostly denotes epistemic meanings (in 59,2% of cases), with near equal distribution among AusE, AmE and BrE, which is followed by dynamic meanings (30,4%), deontic readings (1,7%) and indeterminate cases (5,2%) (Collins, 2009a:126). AmE has a higher frequency of dynamic meanings (1 563 tokens [pmw]) than the other two varieties. *Shall*, on the other hand, is divided between deontic (46,1%) and dynamic meanings (42,1%), and

[95] Seeing that *BE about to* is also excluded from analysis in this thesis.

epistemic meanings only account for 10,8% of its use, with the remaining 1,2% being indeterminate cases (2009a:135). The meanings of *would* correspond to that of *will* in the sense that *would* also mostly denotes epistemic meanings (64,3%), like *will*, and is in the second place characterised by dynamic meanings (22,9). The remaining 12,8% of *would* tokens were indeterminate (2009a:140). *BE going to* also follows this pattern in terms of its two most frequent meanings; in 52,3% of instances it denotes epistemic meanings and in 38,9% of cases it expresses the dynamic meaning. The remaining 8,9% is divided among deontic (0,9%) and indeterminate cases (8%) (2009a:144). Lastly, *want to* was almost exclusively found to express dynamic meanings across all three varieties (overall 98%), with the residual 2% split between deontic (1,1%), epistemic (0,8%) and indeterminate meanings (0,2%) (2009a:151).

From Collins (2009a) the following list of modals and quasi-modal from most to least frequent can be drawn: *can, will, would, could, HAVE to, should, may, BE going to, might, WANT to, must, BE able to, NEED to, (HAVE) got to, BE to, shall, BE supposed to, ought, (had) better* and, finally, *need.* These frequencies of modals and quasi-modals in the large native varieties of BrE, AmE and AusE, with the addition of NZE data for the obligation and necessity cluster from Collins (2005), will be used as the basis for synchronic comparison with SAfE in the results chapter (§ 4.2.2). The semantic findings of Collins for these native varieties will be compared with SAfE in the diachronic results section as well (§ 4.2.1). The next section will explore the path of modality in South Africa.

2.4.3 Modality in SAfE

English reached the shores of South Africa well after the modal system had been grammaticalised into the language and during the periods when some of the quasi-modals were becoming more usual, fresh input form Britain through the various waves of settlers, as well affinity to England, was still strong (until the late 19th century). Yet the coming of English did not instantiate the first transplantation of an Indo-European, or more specifically, a Germanic modal system to the country, as noted earlier. This section first provides a concise reflection on studies on SAfE

grammar involving Afrikaans influence (as considered in the introductory chapter of this thesis), before briefly discussing the correspondences and differences between the English and Afrikaans modal system, because, as seen in previous sections, the contact between these two languages of Indo-European descent is a critical key in the development of SAfE as a native, Southern-Hemisphere variety.

As shown in the previous chapter, Rossouw and Van Rooy's (2012) pilot study of modality in SAfE did not yet consider sociohistorical factors such as language contact in the first exploration of the topic. However, those previous studies on the grammatical influence of Afrikaans on SAfE that have been appraised in Chapter one (§ 1.1.1) included Jeffery and Van Rooy (2004), Schneider (2007:184), Lass and Wright (1986) and Mesthrie (2002). Although Jeffery and Van Rooy (2004) detected nativisation of a construction originating from Afrikaans, and Schneider (2007) recognised loan translations in complementation patterns and preposition uses from Afrikaans, both Lass and Wright (1986) and Mesthrie (2002) did not perceive the '*busy* + progressive' construction as a direct feature of Afrikaans contact. But these peculiar features of SAfE and the influence Afrikaans may or may not have had on them do not sketch a complete picture of Afrikaans influence on SAfE grammar or even the variety in general, and should of course not be viewed as such. This is indeed an important consideration for this thesis. Even if this study yields results that suggest an Afrikaans influence on SAfE modality, it should be kept in mind that that influence applies only to the specific area of modality described and would not necessarily be 'proof' of an Afrikaans influence on the grammar of the variety. The same applies to the reverse: if no Afrikaans influence is found in SAfE modality, then this does not refute a general influence of Afrikaans in other areas of grammar. Section 2.3.3.4 has indeed already proven the existence of such a general and widespread influence in the social history of South Africa, and it perhaps only remains to inquire into the possible areas of language, and more importantly, grammar, in which this influence has manifested.

2.4.3.1 English and Afrikaans modal cognates

When considering the contact of two languages of European descent on a territorially transplanted basis, the fact that English and Afrikaans are both West-Germanic languages means that they share many features (especially on a lexicogrammatical level), e.g. cognates – the shared idea of a modal system encapsulating many cognate modal verbs at the forefront (Mortelmans *et al*., 2009:11). In general, the Afrikaans modals share the following grammatical similarities with English: they do not inflect for person, but do inflect for tense to a limited degree, seeing that only some of the preterite forms are left (cf. Ponelis, 1979:246). This section will explore these cognates in terms of structure and meaning. Meaning, however, is regarded here as the central aspect of comparison, i.e. I will compare those modals which share parallel meanings in the first place, and secondly mention grammar and phonology where applicable. The Afrikaans examples are from Ponelis (1979:246-253), unless stipulated otherwise, and the English translations are my own. I also include more idiomatic English equivalents of some examples were applicable.

The modal verb *can* is cognate with the Afrikaans *kan*, which agrees strongly with *can* regarding all three criteria stipulated above, i.e. meaning, structure and pronunciation. *Kan* can be both deontic and epistemic in its use, denoting ability, possibility (and futurity) and permission (Ponelis, 1979: 249-253), which corresponds with its uses in English (see Table 3 in Chapter 3). The past tense form of *can* in English, *could*, is cognate with the Afrikaans past form *kon*, with strong correspondence as to the meaning, but weaker correspondence as to form and therefore phonology. The ability meanings of *kan* and *kon* are exemplified in (17) and (18) respectively:

(17) Sarel Seemonster[96] **kan** rook deur sy neusgate blaas.

‘Sarel Seemonster can blow smoke through his nostrils.’

(18) Toe ek jonger was, **kon** ek die myl onder vyf minute draf.

‘When I was younger, I could run the mile in under five minutes.’

[96] A character in a popular Afrikaans children’s television programme of the 1970s-90s, *Wielie-Walie*.

The Afrikaans cognate of *may* is *mag,* seeing that these two verbs concur semantically, grammatically and to some extent phonologically. The archaic past tense of *mag*, i.e. *mog*, which is equivalent to English *might,* has fallen out of use in Afrikaans (Ponelis, 1979:246) and therefore the meaning of *might* is conveyed by the present *mag*. It is interesting to note that the preterite forms of Afrikaans modal verbs also have a hypothetical value, as mentioned above in terms of the English past tense modals, which suggests that, because *mog* has fallen out of use, present tense *mag* can be expected to have taken over the hypothetical meaning (Ponelis, 1979: 246). An example of the permission meaning of *mag* is given in (19):

(19) Jy **mag** maar gaan.

'You may go.'

The quasi-modal *BE able to* has no cognate form in Afrikaans, as *kan* is mainly used to denote ability (1979:251), but similar meanings are conveyed by the periphrastic constructions *in staat wees om te* ('to be in a state to') and *die vermoë hê om te* ('to have the ability to'). These circumlocutions are however not grammaticalised in the language – as seen in the use of the nouns *staat* ('state') and *vermoë* ('ability') respectively.

The cognates perhaps most relevant to this thesis are the English *must* and the Afrikaans *moet*, that share the same Proto-Germanic root, **mot* (compare Old Saxon *motan*/*mōt*/*muot* ['to be obliged to, have to'], Old Frisian *mota*/*mōt*, Old High German *muaʒ*/*muoʒa* ['find room or opportunity', 'leisure'], Middle Low German *moten*/*mōte*, Dutch *moeten*, German *müssen* ['to be obliged to'], Gothic *gamotan*/*gamōt* ['to have room to, to be able to']) (Onions, 1966:598; Online Etymology Dictionary, 2013). The Proto-Germanic root is believed to stem from the Proto-Indo-European (PIE) root **med*, meaning 'to measure' or 'to take appropriate measures' (Online Etymology Dictionary, 2013).

Both contemporary English *must* and Afrikaans *moet* do not inflect for person anymore, share high degree of similarity of meaning and enjoy phonological resemblance (cf. Abraham, 2001:23-30; Conradie, 1976; Traugott, 2006:120). *Moet* is

the central and unmarked[97]/neutral modal verb in Afrikaans regarding commands or requests (Ponelis, 1979:251) and can used in the negative form, as in (20), as well as in the entrenched negative contracted form, as in (21), and in the positive form, as in (22):

(20) Jy **moet** nie hier sit nie.

'You must not sit here / Don't sit here'.

(21) **Moenie** die goed hier wegvat nie.

'You mustn't take the things away from here / Don't take the things away from here.'

(22) Ons vra dat jy dit asseblief tog **moet** doen

'We ask that you must please just do it / (We ask you to) please just do it.'

Afrikaans *moet* is therefore possibly used more widely to communicate requests than English *must*, seeing that the English translations in (18), (19) and (20) are interchangeable with a construction without a modal, and the latter is perhaps the more natural option. This suggests that, *moet*, as the neutral form for requests, does therefore not necessarily expresses a high degree of strength, but can often convey a median degree of obligation.

Afrikaans *moet*, together with its past form *moes*, can also denote epistemic meaning, in the same way that English *must* can. As there is no contemporary grammatical past form for present *must* in English, both *moet* and *moes* express the meaning of *must*. However, despite its present reading today, *must* is actually the contemporary remnant of the original past form of the OE present form(s) *mote/mot* etc., namely *moste* (OED Online, 2012b) (cf. § 2.4.1.2.1). Historically, therefore, the real cognate pairs are OE *mote/mot* and *moet* (Dutch *moeten*) (also compare German *müssen*), as in contrast to OE *moste* (the ancestor of contemporary *must)* and *moes* (Dutch *moesten*) (also compare German *müssten*) (cf. Mortelmans *et al*., 2009). The

[97] Very simply speaking, markedness is described by Trask (2007:163) as "the property which distinguishes less neutral linguistic forms [marked forms] from competing ones which are more neutral [unmarked forms]." Trask also mentions that although "the concept is older, the term markedness was introduced by the European linguists of the Prague School in the 1920s" (2007:163).

use of past tense *moste* in OE is illustrated in (23), which is more easily translatable as *could* or *had to*:

(23) þenden he burhwelan brucan **moste**[98].

'While he the prosperity of a city could enjoy'[99].

The OED Online (2012b) notes that *moste* denoted both past and present senses since OE, even though it is the original past form, whereas *mote/mot* only denoted present senses. *Moste* was therefore the more productive form, and this is perhaps the reason why *mot* only survived until the late 17th century (OED Online, 2012a). The Online Etymology Dictionary (2013) notes that the *moste* was more commonly "used as [the] present tense from c.1300, from the custom of using [the] past subjunctive as a moderate or polite form of the present".

Examples (24), (25) and (26) respectively illustrate the uses of *mote/mot* in OE and ME, as well as the latest recorded example (1895) from the OED Online (2012a):

(24) Londrihtes **mot** þære mægburge monna æghwylc idel hweorfan[100].

' Of land-rights must of your clan every man become deprived'[101].

(25) He **moot** reherce as neigh as evere he kan euerich a word[102].

'He must report, as nearly as he can, every least word'.[103]

(26) Saint Francis, loving bee thee till, An' thankit **mote** thou be for thy gudewill[104].

Today, however, in mostly epistemic cases, the construction *must have* can indicate past time in English (cf. § 2.4.1.2.1). In deontic *moes* in Afrikaans, however, the force

---

[98] Source: *Beowulf*, line 3100. From the OED Online (2012b).

[99] Translation from: http://www.heorot.dk/beowulf-rede-text.html

[100] Source: *Beowulf*, line 2886 (OED Online, 2012a).

[101] Translation from: http://www.heorot.dk/beowulf-rede-text.html

[102] Source: Geoffrey Chaucer's *Canterbury Tales, Prologue,* line 732-3. (OED Online, 2012a).

[103] Translation from: http://pages.towson.edu/duncan/chaucer/duallang8.htm

[104] Source: J. L. Robertson's *Dunbar*, line 84. (OED Online, 2012a).

of obligation is more frequently somewhat weaker than *moet*, because of its hypothetical reading in the preterite form (Ponelis, 1979:246). Examples (27) and (28) illustrate past readings of epistemic uses of *moet* and *moes* respectively, which both translate into the same English expression:

(27) [Hy] **moet** oneindige energie gehad het.

'He must have had inexhaustible energy.'

(28) [Hy] **moes** oneindige energie gehad het.

'He must have had inexhaustible energy.'

Conradie (1987:177-8) describes the expressions of commands and requests via the Dutch equivalents of *moet* (Dutch *moeten*) and *sal* (Dutch *zullen*) (English *must* and *shall* respectively) as follows: "In the use of *zullen,* the source of modality is the will of S [the speaker], projected onto the general order of things whence it derives absolute necessity". On the other hand, in the case of "the more predominant *moeten*", "the source of modality and the will of S [the speaker] may be obliged to motivate his request" (1987:178). Conradie (1987:178) illustrates this use of *moet* through the second sentence of the following example (from a novel):

(29) Jy **moet** my terugneem. Ek **moet** in die Kaap kom.

'You must take me back. I must get to Cape Town.'

The second sentence suggests that a "*moet*-command – unlike a *sal*-command – may be embedded in an *Ek wil ... hê* ('I want')-matrix[105], indicating that the choosing of *moet* does not make it redundant for S to add a reference to his will as motivation" (Conradie, 1987:178). This kind of usage corresponds to that of dynamic *must*, where, according to (Collins, 2009a:40-1), its use may not be sharply distinct from deontic *must* in that "it is understood that an obligation is imposed by a deontic source", but that it is more salient in its expression of "a circumstantially derived need".

[105] A 'matrix' clause or sentence, as a linguistic term that is derived from mathematics, and especially used in generative grammar, refers to "the superordinate sentence within which another sentence is embedded, e.g. *The student who shouted left*, where *The student left* is the **matrix sentence**, and *The student shouted* is the embedded sentence" (Crystal, 2008:297-8; emphasis original).

It is interesting to note that the *Pharos Afrikaans-Engels/English-Afrikaans Dictionary* (2005:356) lists multiple translation options for Afrikaans *moet* to other (mostly) modal and non-modal constructions in English, viz. "must, have to, be compelled/forced/obliged to; should, ought to", "[have] got to", "need be", "is supposed to", "had better" and "need to". This list basically covers the entire continuum of modal strength: from the traditionally strong English modal *must* to the much weaker *ought*. The negative Afrikaans construction *moenie* is consistently translated as "don't" (2005:355-6) – matching my translations in (20) and (21). Equivalents of formulae containing the negative forms *moet nie* or *moenie* together with the polite Afrikaans adverb for requests 'asseblief', meaning 'please' (as in "[moet] asseblief nie") are given as "please do not", "kindly abstain from" and "please refrain from" (2005:356). This indicates that *moet* in its negated form is often used as part of politeness formulae in Afrikaans. The past form *moes*, apart from being translated as "must have" and "should have", is even translated to "could" and "would" in some formulaic expressions (2005:356). *Moet* is therefore a verb that can be used in a wide range of contexts, many of which are traditionally associated with various other modals or constructions in English.

Should, unlike *must*, does not have semantic correspondence with its Afrikaans cognate, *sou*, which, being the past tense of *sal* (cognate with English *shall*), expresses a meaning closer to that of *would*. The meaning of *should* is expressed to a greater extent via the Afrikaans construction *behoort te*, which essentially acts as a quasi-modal and expresses both deontic and epistemic meanings (Ponelis, 1979:247;249). Epistemic meaning is denoted in (30):

(30) As hulle geen teëspoed kry nie, **behoort** hulle so teen vieruur hier **te** wees.

'If they don't have bad luck, they should be here by about four o'clock.'

The Afrikaans construction, *hoef te*, has a similar meaning to both *need* (as a modal) and *NEED to*, but is mostly used in the negative form *hoef nie te*, which is semantically

equivalent to *need not* and *do not need to*[106] (cf. Ponelis, 1979:247). Example (31) shows *hoef te* in its negative construction:

(31) [Jy] **hoef** niks **te** doen nie.

'You don't need to do anything.'

Other Afrikaans constructions that can express the meanings of *NEED to* and *need* include *is nodig dat* and *het nodig om te* (both not grammaticalised constructions containing the adjective *nodig* [English *necessary*]), but generally deontic and dynamic *NEED to* (as in [16]) translates as *moet* (cf. Collins, 2009a:75; Pharos, 2005:356). *Ought* shares no cognate with Afrikaans, but its meaning is also close to that of *behoort te* (and also to some extent to *moet*), which is similar to the case of *should* (cf. Pharos, 2005:356). Both *HAVE to* and *(HAVE) got to* have no cognate forms in Afrikaans and their meanings are both conveyed by *moet* (cf. Pharos, 2005:356), as for example seen in the dynamic senses in (32). Here *moet* can also be translated as *NEED to*.

(32) Ek **moet** met jou praat.

'I need to talk to you. / I have to talk to you. / I've got to talk to you.'

The meaning of English *BE supposed to*, on the other hand, is similar to the Afrikaans expression *veronderstel wees om te*, but also corresponds to the meaning of *moet* (cf. Pharos, 2005:356).

The meanings of both *shall* and *will* correlate with Afrikaans *sal*, which is originally the cognate of *shall* (as noted above), sharing much of its structure and phonology as well. However, as seen in sections above, *shall* is falling out of use in everyday contexts and has become restricted to special environments, and is more readily replaced by *will. Will* also shares the meaning of the Afrikaans verb *gaan*, which is more closely related in structure (compare *go* and *gaan)* and semantics to the

[106] This corresponds to the Dutch verb *hoeven*, which is also "a negative polarity item the meaning of which can be compared to that of *need* in English or *brauchen* in German" (Mortelmans *et al*., 2009:17).

English quasi-modal *BE going to*. Example (33) illustrates the equivalent meanings of *sal* and *gaan*, together with *shall/will and BE going to* in a future prediction.

(33) Riaan[107] **sal** nou die nuus lees. / Riaan **gaan** nou die nuus lees*.

'Riaan will now read the news. / Riaan is now going to read the news.'

*My sentence

The actual cognate form of *will* in Afrikaans is *wil*, but today it shares its meaning with the English quasi-modal *want to*, indicating volition (as in [34]), and similarly, the past form of *wil*, which is *wou*, corresponds semantically with *wanted to* (as in [35]). The past form of *will*, *would*, is semantically equivalent to the past form of Afrikaans *sal*, which is *sou* (as in [36]), as also noted above.

(34) **Wil** jy my gewete wees?

'Do you want to be my conscience?'

(35) Niemand **wou** luister nie.

'Nobody wanted to listen.'

(36) Hiervolgens **sou** Nederlandse bedrywe dieselfde reëls in ag moet neem vir swart werknemers en hul gesinnne as vir blankes[108]

'According to this, Dutch companies would have to apply the same rules for black employees and their families as for whites.'

Note that the modal combination of *sou moet* in Afrikaans cannot be translated as *would 'must'* in English, but either as *would have to* or *would need to*. The Afrikaans auxiliary *gaan* is cognate with quasi-modal *BE going to*, but technically *gaan* only has structural correspondence with *going*. *Gaan* in Afrikaans indeed translates as *go* or *going* in English, since Afrikaans does not mark the progressive aspect morphologically.

---

[107] Refers to Riaan Cruywagen – a famous and popular Afrikaans newsreader from the 1970s to 2012.

[108] From *Beeld:* Ponelis (1979:250).

This section has briefly shown that English and Afrikaans share many cognate modal verbs, and that, some cognates still correspond in terms of semantics today, while others only correspond in structure, but not semantics. Those modals that have retained both a strong grammatical and semantic resemblance over time are *can* and *kan*, *may* and *mag*, *must* and *moet* and *shall* and *sal*. Table 1 provides a summary of the most prominent Afrikaans and English cognate modals, which is based on meaning rather than structure.

Table 1

Afrikaans and English modal cognates/semantic equivalents

Afrikaans	English
kan	*can*
kon	*could*
mag	*may/might*
moet/moes	*must/NEED to/HAVE to/(HAVE) got to*
behoort te	*should/ought*
hoef te	*need/NEED to*
sal/gaan	*will/shall*
sou	*would*
wil/wou	*want to/wanted to*
gaan	*BE going to*

From Table 1 it is clear that *moet*/*moes* is by far the modal pair in Afrikaans with the widest range of equivalent expressions in English, including modals and quasi-modals. The modal options in Afrikaans can therefore be said to be more limited than in English, especially regarding the expression of obligation and necessity.

In this section, the field of modality was explored by describing the modal system in terms of its birth through grammaticalisation, its semantic properties and its development in the more recent past within varieties of English. The correspondence of the English and Afrikaans modal systems was also reviewed and exemplified. Owing to the fact that it remains a relatively unexplored issue whether Afrikaans contact has had a profound influence on the grammar of SAfE, the results of this thesis could contribute to answering this question by comparing the frequencies of English and Afrikaans cognates and exploring the role of semantic relationships within those frequencies where relevant.

2.5 CONCLUSION

This chapter has reviewed the literature relevant to this thesis with the aim to provide a descriptive synthesis between theory and context in terms of the diachronic and synchronic study of modality in SAfE. Three main ideas have been considered: firstly I have examined the general theoretical and empirical frameworks that support the research in this thesis, secondly I have undertaken an exploration of social history and context, and thirdly, I have described the grammatical features central to this study – i.e. modal and quasi-modal verbs. These three ideas have been divided into three levels of specificity – from the most general notions of language change to the most prominent frameworks for varieties of English, and finally to the specific variety of (native) South African English.

The composite field of sociohistorical linguistics has been described in terms of its two subfields: historical linguistics (including grammaticalisation theory) and sociolinguistics (including frameworks for English varieties and politeness theory). The empirical framework of corpus linguistics, which includes register studies, has been elucidated, in order to illustrate how theory and method interact in this thesis. Two main conclusions can be drawn from this first main section: the corpus-based study of historical texts is greatly valuable in analysing language change, and the context surrounding those texts, whether social, historical or linguistic, is of pivotal importance to provide reasons for the changes perceived in those texts.

The social history of language change has been explored in terms of language evolution, English evolution and the expansion of English around the globe and, specifically, to South Africa. The history of English in South Africa has provided a basis for evaluating frameworks for English varieties in terms of their application to South Africa itself, and it has been pointed out that the country defies classification in many respects due to its unique context. It can be concluded that neither Trudgill (2004) nor Schneider's (2003) models are wholly applicable to South African English in terms of the factors they propose, even if they are adequate in many respects. Bekker's (2012; 2013) model has provided a new and perhaps more comprehensive perspective into this matter, as it considers more evidence related to immigration and migration in South Africa than previous studies. From this second main section it can

be concluded that migration, language contact and identity are the three most prominent factors of language change in general, in English and in SAfE.

The third main section has explored the system of modality in terms of terminology and semantics, as well as the development of the modal system through grammaticalisation. It has also provided a critical look into the many changes that the system has underwent, and is still undergoing, in the native varieties of English around the world. The correspondences of the Afrikaans and English modal system have also been explored. From this section the conclusion can be drawn that modality, and specifically the modal and quasi-modal verbs, is part of a vastly complex system involving intricate semantics, notions of gradience, fuzziness and oftentimes indeterminacy, but also that, despite this, it is not impossible to analyse. Furthermore, changes are occurring in the native Englishes of the world regarding e.g. the rise of quasi-modals and decline of modals, which can be attributed to social factors.

The question now arises as to the state and development of modality in South African English, as well as the motivations behind such changes and variations. Another question is whether the unique context of the country has rendered the system of modality in SAfE unique among the World Englishes. The study can be anticipated to be a complex one, due to firstly, the complexity of the country's context and secondly, the complexity of the grammatical and semantic workings of the modal system. Complexity of this kind may well hinder a certain level of detailed investigation, as Zadeh (1972) suggests:

> "In general, complexity and precision bear an inverse relation to one another in the sense that as the complexity of a problem increases, the possibility of analyzing it in precise terms diminishes."

However, it is the ideal of this thesis to analyse, from texts, SAfE modality in terms as precise as possible, regardless of the linguistic and sociohistorical complexity found in both the relevant texts and contexts. The next chapter will describe the many methodological aspects of the analysis of modality in SAfE, by linking with the focus on text and context provided in this chapter.

CHAPTER 3

RESEARCH METHODOLOGY

"What is a scientist after all? It is a curious man looking through a keyhole, the keyhole of nature, trying to know what's going on." (Jacques Yves Cousteau)

"It is the mark of the educated man to look for precision in each class of things just as far as the nature of the subject matter admits." (Aristotle)

3.1 INTRODUCTION

This chapter will shed light on the research methodology applied in the data collection, analyses and interpretations related to this dissertation. General, as well as specific methodological issues receive attention, with the emphasis on research design, general analytic and quantitative methods, as well as interpretative, qualitative methods used for examining the modals and quasi-modals. Firstly, the section on research design will describe the study population of this dissertation, as well as the different corpora utilised for analysis. These corpora, in representing something as vast and untold as a language, is indeed only the 'keyhole' through which this thesis looks in an attempt to unravel a much bigger picture, as is the nature of any scientific description, according to the quote by Jacques Yves Cousteau above. Corpus compilation and preparation, as well as the content of the historical corpus of SAfE will receive special attention here. Secondly, the general analysis tools and quantitative methods will be presented by describing the analytic processes of e.g. using concordances and normalising raw figures. Lastly, the qualitative methods for semantic interpretations will be explained, including macro- and microsemantic classification, as well as the specific parameters selected to conduct the latter. With these specific parameters, along with the more general semantic interpretational

methods described in this chapter, I seek to establish a degree of precision in the analysis of such gradient phenomena as modal meanings, as Chapter 2 (§ 2.4.1.1) describes. This can indeed only be achieved "just as far as the nature of the subject matter admits", as Aristotle observes in the second quote above.

3.2 RESEARCH DESIGN

Seeing that the general research design of this study relies on the corpus-based approach, the study population and the corpora representative of that population will receive special attention here. These descriptions will link with the historical circumstances, as well as the empirical framework described in the previous chapter.
The study makes use of both diachronic and synchronic corpora as a collective linguistic database, in order to retain the balance between, and exploit the extended scope across, these two chronological approaches, as described in Chapter two. Before an exposition of the corpora, however, we need to become acquainted with the users of SAfE – the study population of this dissertation.

3.2.1 Study population

As mentioned in Chapters 1 and 2, the native speaker community of SAfE is the targeted study population of the current study. This includes the historical settler community[109] from the 1820s and their descendants across the rest of the 19th century, and into the 20th century, up until the contemporary community of the 1990s and onward (as described in subsequent paragraphs). More specifically, the authors or writers of the texts from the past, as well as the writers and speakers of the contemporary texts, were counted as the linguistic community of study.

During the historical corpus compilation process, each author of a text (dating from the 1870s to 1950s) that was considered for use in the corpus was assessed

[109] The circumstances of the Settlers and their descendants are clarified in detail by e.g. Butler (1974) and Hockley (1966).

according to birthplace and date of birth. To ensure the credibility of the study, it is essential that the authors were clearly born and raised (and spent most of their lifetimes) in South Africa – sometime between the 1820s and the present. However, between 1820 and around 1850 to 1870 (30 years representing one generation and Trudgill's (2004) construct of a 50-year colonial lag, described in § 2.3.2.3, taken into consideration) authors of texts are considered as representatives of the input type of English (19th century dialects of BrE) that evolved into SAfE over time, according to the traditional view of SAfE formation (see § 2.3.3.3). These authors were therefore born and raised in Britain.

When attempting to establish the heritage of each author, the literary texts used in some cases provided such biographical information, but in other cases it was necessary to conduct a search in various encyclopaedic sources to obtain these details. By the end of the compilation process, the corpus included texts written at various South African localities to cover the largest-possible scope of local dialects, including the Eastern (e.g. Grahamstown and Port Elizabeth) and Western Cape (e.g. Cape Town), Natal (e.g. Durban), Johannesburg (and surrounding area), Kimberley and occasionally England itself (however by authors still considered South African, while e.g. on vacation). This therefore includes both users of the Natal and Eastern Cape dialects (which emerged from the formation of SAfE, as described in the previous chapter), as well as the communities in and around the dialect mixing hot-pots in the north, arising from the various strands of immigration and migration (due to e.g. the gold rush described in Chapter two).

The authors and recipients of the letters include, among others, some well-known writers of South African English literature, such as Olive Schreiner, Pauline Smith, Sir Percy Fitzpatrick and Alan Paton. Some extracts of these writers' works of shorter or longer prose were also included in the corpus, and some newspaper texts also included discussions of their work, life and times. Not all texts are connected with literature or famous writers, but the fact remains that texts of such a nature (especially letters) are more easily accessible and more likely to be preserved in archives or museums, as I have found to be true. Other texts that do not follow this pattern usually report pivotal events in South African History, such as the discovery of gold and diamonds on the Witwatersrand (Johannesburg-district) and Kimberley respectively, the Anglo-Boer War and the road to a democratic South Africa – which is especially true for journalists' reports in newspapers, and sometimes even letters

written by persons involved in or affected by such events, e.g. political figures, would-be political analysts, military officers, owners of companies, etc.

Because of the nature of social history in earlier decades, most of the older social letters in the corpus are written by women, whereas business letters and newspaper articles are, by definition, written by men, but this pattern evens out towards the beginning (for business letters) and the end (for news) of the 20th century. An overall even distribution of male and female authors occurs in the fiction genre, with a general male dominance in non-fictional writings across the entire time span.

As for the contemporary linguistic community of this study, Jeffery (2003:343) describes the relevant study population as the prototypical WESSA, but not excluding the fluently bilingual or multilingual (not necessarily English mother tongue) "educated in English to matriculation[110] level or beyond" (as also mentioned in Chapter one). These contemporary, educated users of a native or near-native variety of English in South Africa are described as being "geographically mobile" and therefore regional variation in the corpus is not such a pivotal factor. Nevertheless, users from the Eastern and Western Cape, as well as KwaZulu-Natal (previously Natal) and Gauteng (the area including Johannesburg and Pretoria) were targeted (2003:343). In order to draw synchronic comparisons between SAfE and Afrikaans, first-language users of contemporary Afrikaans, the indigenous language derived from colonial Dutch (as described in the previous chapter), are included as another study population altogether.

3.2.2 Corpora

This section will illuminate the corpora used for this dissertation. These include, for diachronic data, the historical corpus of SAfE, and, for synchronic data, both ICE-SA and an Afrikaans corpus with which to draw comparisons with the former.

[110] The South African Grade 12 or final secondary school year.

3.2.2.1 The historical corpus of SAfE

In order to acquire a historical corpus that covers the range from the early 19th to the late 20th century, an add-on approach contributes to the convenience and reliability of the entire collection. One existing corpus was used to represent the current state of SAfE, while a new corpus was specifically compiled for this thesis to cover the various historical stages beforehand. The registers chosen for this corpus are letters (business [formal] and social [private]), news, fiction and other narrative texts, and non-fiction. The texts were then organised into four periods, namely Period 0, 1, 2 and 3, which will be fully discussed below. This new corpus of SAfE is largely compiled from data gathered[111] from the following libraries, archives, museums and collections:

(i) **National Library of South Africa** (NLSA): Cape Town Campus
(ii) **National Archives and Record Service of South Africa** (NARS) – Western Cape Provincial Archives and Records Service: Cape Town
(iii) **National English Literary Museum** (NELM): Grahamstown
(iv) **Cory Library for Historical Research**: Grahamstown
(v) **William Cullen Library** of the University of the Witwatersrand (WITS) (Historical Papers collection): Johannesburg
(vi) **Afrikana collections** of the North-West University (NWU) campuses: Potchefstroom and Vaal-Triangle.
(vii) Website for **British 1820 Settlers to South Africa**[112]

Other sources were, smaller personal collections of the author and friendly donations[113].

Although it was intended to keep proportions regarding register representation in each period as equal as possible, some practical restrictions arose, relating to

[111] I personally undertook the data collection the locations listed from (i) to (vi), with the assistance of Wimpie Wasserman in (i), (ii) and (iii), Prof. A.J. van Rooy and Mariet Steinmann in (iv and vii), as well as Christa Rossouw in (vi). The travel costs and fees for photocopies, microfilm-prints, text processing, etc. were generously funded by research grants from the NWU (North-West University) and NRF (National Research Fund).

[112] http://www.1820settlers.com

[113] Most of these donations can be attributed to Prof. A. J. van Rooy, Prof. L. K. van der Walt, previously of the UPE (now the NMMU) and Mrs. M. van der Walt.

firstly, the amount of available data (especially regarding the availability of texts in earlier periods), a restricted time scale for myself in which to compile the data, and the restriction on available manual labour by means of which to process the texts. Copyright issues are a further restriction on corpus size, as great care is taken in ensuring that all the requirements of archives, libraries, special collections and museums are met, according to each institution's policy.

Where possible, texts were photocopied and/or scanned before being processed electronically. In the case of many older texts (which were not allowed to be photocopied or scanned), the originals were retyped on site. Newspapers on microfilm were either printed and then typed, or, where possible, saved in electronic format. The electronic versions of all the texts were finally converted to plain text (txt) format and manual checking and editing ensured the accuracy of each converted text. General markup was done in accordance with the ICE Markup Manual for Written Texts (Nelson, 2002). The composition of the corpus is outlined in Table 2[114].

Table 2

Total word counts for registers in the historical corpus of SAfE per period

	Letters	**News**	**Fiction/ Narrative**	**Non-fiction**	**TOTAL**
Period 0 1820s – 1860s	48 313	72 633	58 565	14 560	194 071
Period 1 1870s-1900s	41 090	42 111	61 626	26 359	171 186
Period 2 1910s-1950s	41 633	31 873	30 568	37 091	141 165
Period 3 1990s	14 415	39 455	35 346	64 132	153 348
TOTAL	145 451	186 072	186 105	142 142	659 770

Table 2 shows that the collective diachronic corpus amounts to 659 770 words. The texts are organized into the periods that Schneider (2007:175-185) surmises to be representative of different phases in the development of SAfE. By ordering the periods according to these phases, some insight might be gained into the various stages of the development of SAfE, in view of Bekker's (2012) emphasis on its extended formation period (which implies an extended Phase 1) (see § 2.3.3.3).

[114] For a more detailed description of the corpus word counts per decade, refer to Table 19 in Appendix 1.

Period 0 in the diachronic corpus covers the first era of settlement and therefore the time of input, the 1820s to 1860s, representing so-called 'Proto-SAfE' – the baseline for the formation of SAfE (cf. Mesthrie & West 1995). The words in this selection amount to 194 071. The label Period 0 was chosen to represent this phase in SAfE history, because the language used in these texts cannot yet truly be called South African English. This period more or less corresponds with the start of Schneider's (2007:175-178) Phase 1. Period 0 would therefore still be in Trudgill's (2004) mixing (of transplanted dialects) phase and Schneider's (2003) first phase of foundation. It is only around 1870 (according to Schneider 2007), that what can be called the first-generation users of SAfE emerge. Therefore, the remainder of the historical corpus was compiled to cover the periods described below.

Period 1 (171 186 words) ranges from the 1870s to the 1900s (along the lines of Schneider's 2007:178-181 Phase 2)[115], whereas **Period 2** (141 165 words) is the 1910s to the 1950s, and **Period 3** (a selection from written ICE-SA[116] to match the registers of the historical corpus [also described in § 3.2.2.2]) is the 1990s (together representing Schneider's [2007:181-185] Phase 3), which contains 153 348 words. Period 0 to Period 3 hence covers about 170 years in the development of SAfE. When taking into account Leech's statement (2003) mentioned in Chapter 2, relating to emerging grammaticalization which is already being discernable within one generation, the anticipated developments in the modal system of SAfE over the course of nearly two centuries in this case should provide a much more comprehensive picture of such a process.

A short description of the texts used in each period will follow, together with some examples of authors and their locations. A basic city map of South Africa is provided below to illustrate some of the places where the texts were written.

---

[115] If one considers Bekker's (2012) view (cf. § 2.3.3.3), the implication is that Schneider's (2003; 2007) Phase 1 (foundation) could only really commence by the late 19th century. The general era surrounding Period 0 and 1 in the historical corpus could therefore broadly represent Phase 1, and Period 2 could broadly represent Phase 2 (exonormative stabilization), with the exception of the 1910s-20s possibly still forming part of Phase 1 (see § 5.3.2).

[116] See Jeffery (2003) on ICE-SA.

Figure 3

A contemporary map of South Africa[117]

The letters in Period 0 (48 313 words) are divided between 16 091 words for social or private letters and 32 222 words for business or formal letters. The predominance of formal letters in this period exists due to the free availability of e.g. complaint letters addressed to the authorities on the *British 1820 Settlers to South Africa* website. The formal, business letters for Period 0 include the following authors: George and Mary Gatehouse, George and Richard Pigot, John Campbell, John Ingram, Anne Oliver, Mary Campbell, J. W. Weir, etc. In fact, letters from 19 different authors were included in the 1820s alone, with 2 authors each for the 1840s, 50s and 60s, amounting to 25 authors of business letters in Period 0. These letters

[117] Map available from: http://www.toursa.com/travel_south_africa/southafrica_map.htm

were written from e.g. Cape Town, Grahamstown, Lovedale (an Eastern Cape mission station), Durban, and a few business letters were written in England.

On the other hand, the social letters for the 1820s-60s include selected correspondence of Sir Thomas Pringle, letters from distressed settlers written to family members or friends in England or South Africa, letters addressed to Sir Alexander Duff Gordon, selected correspondence of Jane Elizabeth Waterson, as well as letters from various other authors. There are social letters from 10 different authors in the 1820s, 2 authors in the 1830s and 1840s respectively, and 3 in the 1860s, amounting to a total of 17 authors of social letters in Period 0. The locations whence these letters were written include Genadendal (the oldest mission station in Africa, Western Cape), Grahamstown, Somerset West and Lovedale. Apart from the few letters written in England, which are also representative of early 19th century input, the social and business letters of Period 0 that are written in South Africa therefore cover the areas of the first and second waves of settlement, as well as the areas of governance, namely, the Cape Colony (including contemporary Western and Eastern Cape), as well as Natal (contemporary KwaZulu-Natal). The business and social letters of Period 0 amount to 48 313 words, and were written by both male and female authors.

The newspapers used for Period 0 (72 633 words) and their places of publication, include the following: for the 1820s, *The South African Commercial Advertiser* and *The South African Chronicle and Mercantile Advertiser* (both Cape Town); for the 1830s, *The Graham's Town Journal* (Grahamstown); for the 1840s, *The Natal Witness* (Pietermaritzburg) and *The Port Elizabeth Telegraph* (Port Elizabeth); for the 1850s, *The Friend of the Free State* (Bloemfontein); and for the 1860s, *The Transvaal Argus* (Potchefstroom/Johannesburg) and *The Cape Standard* (Cape Town). The news register of Period 0 therefore comprises 8 different newspapers written all over South Africa.

As far as fiction and other narrative writing are concerned (58 565 words), extracts from *The Chronicle of Jeremiah Goldswain* represents the 1820s, whereas excerpts from *The Wrongs of the Caffre Nation* (by Robert Mackenzie) and a *Narrative of an Expedition into Southern Africa* (by William C. Harris) constitute the 1830s. Extracts from *A Narrative of a Residence in South Africa* (by Thomas Pringle) and *A Missionary Narrative of the Triumphs of Grace* (by Samuel Young) represent the 1840s. The 1850s consist of selected tales from *A South African Miscellany of*

Historic, Narrative, and Descriptive sketches, tales and poetry (by George Duff), as well as parts of *The Reminiscences of Thomas Stubbs*, and the 1860s are represented by excerpts from *South African gold fields: a narrative of discovery* (by Alexander Wilmot) and *The Reminiscences of John Montgomery*.

The non-fiction component of Period 0 (14 560 words) comprises the *Report of the Committee of the Society for the Relief of Distressed Settlers in South Africa* held in Cape Town (1823)[118], as well as a *Biographical Sketch* of Thomas Pringle written by Josiah Conder (1842). Non-fiction is therefore the register with the smallest word count in Period 0, with news as the register with the highest word count. To conclude: the word count for this period totals at 194 071 words. The next few paragraphs will shed light on the corpus makeup for Period 1.

The letters in Period 1 are divided between 8 464 business and 32 626 social letters. The business letters from the 1870s were written by 7 different authors, including the Secretary to the Transvaal Government (M. Osborn) at Lydenburg (in contemporary Mpumalanga), as well as at Pretoria. Selections from Sir James Rose Innes' letters are included in the formal letters of the 1880s, together with those of 2 other authors; these were written at Cape Town, Alice (a Scottish mission station in the Eastern Cape) and Kimberley. The 8 authors of business letters in the 1890s, writing from Cape Town, Pretoria and Kimberley, are e.g. W. H. Hockly and Jane Elizabeth Waterson, and the three authors of business letters in the 1900s include the major of Kimberley, H. A. Oliver, writing from e.g. Kimberley, Cape Town, Grahamstown and Natal. Hence, there are altogether 21 various authors of business letters for Period 1.

The larger portion of letters for Period 1 however consists of social letters, which, in the 1870s, were written from e.g. Dordrecht and Kraai River (both Eastern Cape), Cape Town, Lovedale, Durban and Kimberley, by, among others, John X. Merriman and Olive Schreiner. Three authors wrote the private letters used for the 1880s, namely the two mentioned above, as well as Jane Elizabeth Waterson – who wrote from e.g. London (while on holiday), Cape Town, Lovedale and Grahamstown. The letters of the 1890s hail from Cape Town, Middleburg (Eastern Cape – not to be confused with the Middleburg in contemporary Mpumalanga), London, Karree Kloof (near Carnarvon, Northern Cape), Pretoria, Kimberley, King William's Town

[118] From: http://www.1820settlers.com

(contemporary KwaZulu-Natal), and are from the pens of 9 authors, including Chas Levinson, J. A. Stevens, Sir James Rose Innes, Olive Schreiner and P. Macowan. The social letters of the 1900s (12 authors) were written in Johannesburg, Cape Town, East London, Lovedale and London, by e.g. Samuel Evans, H. L. Hall, F. E. Garret and Eunice M. Nestie. The letters of Period 1 (collectively 41 090 words) were therefore written by 27 different male and female authors, and their places of writing cover a vast area of the country, including the Western, Eastern and Northern Cape, Natal and the Transvaal (e.g. contemporary Gauteng and Mpumalanga).

The newspapers included in the corpus to represent Period 1 (42 111 words) include *Crocott's Penny Mail* and the *Graham's Town Journal* (both Grahamstown) for the 1870s. The texts from the 1880s also include these two newspapers, along with the *Eastern Star* (Grahamstown) and the *Diamond Fields Advertiser* (Kimberley). The latter is also present in the texts covering the 1890s, together with the *Cape Times*, the *Cape Argus* and the *South African Telegraph*, (all Cape Town), as well as the *Graham's Town Journal*. The *Wynberg Times* and *The South African News* (both Cape Town), as well as the *Transvaal Leader* (Johannesburg) and the *Graaff Reinet Advertiser* (Graaff Reinet) represent the newspapers of the 1900s. Hence, the texts for the news register of Period 1, amounting to 42 111 words, are from 12 different newspapers, which were published in the Eastern, Western and Northern Cape, as well as the Transvaal.

The register of fiction and other narrative writings for Period 1 (61 626 words) include the following for the 1870s: extracts from *Sport and war or recollections of fighting and hunting in South Africa* by John Jarvis Bisset, as well as selected narrative articles from the *Cape Monthly Magazine*, including 'My First Journey' by the Rev. Richard Ridgill and 'From Graham's Town to the Gouph' by Dr. William Guybon Atherstone. Extracts from *Fighting and farming in South Africa: a narrative of personal experiences* by Fred G. Browning, a *Narrative of an explorer in tropical South Africa* by Francis Galton, *Alone among the Zulus: the narrative of a journey through the Zulu country*, *by a plain woman* by Charlotte Barter, *The history of the battles and adventures and the Zulus, &c., in Southern Africa* by Duncan Campbell Francis Moodie, as well as *The far interior: a narrative of travel and adventure* by Walter Montague Kerr, represent this register for the 1880s. The texts for the 1890s include the short story, *The Black Mamba*, by Ernest Glanville and an extract from *Eighteen-Ninety-Nine* by Olive Schreiner. Excerpts from *The Cloud Child* by Stephen

Black and *Jock of the Bushveld* by Percy Fitzpatrick represent the first decade of the 20th century, the 1900s.

The non-fiction component of Period 1 comprises 26 359 words. Selected articles from the *Cape Monthly Magazine* represent the 1870s, the titles of which include e.g. 'Cape English' (an extract of this article was discussed in § 1.1.3) and 'Extraordinary Hairbreadth Escapes: A Fragment of Frontier History' (by the Hon. Robert Godlonton); also included is an article from *Outlook on a century: South Africa 1870-1970*[119], namely 'Are You True Britons Still?'. For the 1880s, an information pamphlet on the *Lovedale College Scheme* by the chairman in 1888, J. C. Macintosh, was used, as well as articles from *Outlook on a century: South Africa 1870-1970* – these include 'A Settler View' by F. King and 'Conscience is a Power' by John Davidson Don. As far as the 1890s are concerned, selected articles from *Outlook on a century: South Africa 1870-1970* were once again used, the titles of which include 'Christianity and Conquest', 'Atrocities in the Congo' and 'Black and white' (by J. Moffat). Other texts from the same decade comprise a *Selection of Short biographies of the Galla rescued slaves, now at Lovedale: with an account of their country and their capture,* as well as extracts from *The Schröder Art Momento* by Chas Cowan. Selected articles from *Outlook on a century: South Africa 1870-1970* were once again used to represent the 1900s, including 'Among the Matabele' by a Mrs. Turner and 'African Type of Christianity?' by Brownlee J. Ross. To conclude: the word count for Period 1 (across all registers) is 171 186, with fiction and other narrative writing as the register with the most words, and business letters as the register with the least. The next part of this section will discuss the content of the corpus for Period 2.

The letters of Period 2 (collectively 41 633 words) are divided between 17 567 words in the business letters component and 24 066 words in the social letters component. The business letters of the 1910s, together written by 10 authors, include selected correspondence of the Royal Yacht Club, as well as the Claremont Congregational Church (both based in Cape Town). This period also includes selected correspondence of the Mission Council at Lovedale (mostly letters addressed to the Rev. J. Henderson), as well as the Alice Municipality and the Juvenile Musical

[119] *Outlook on a century: South Africa 1870-1970* is a collection of articles from the journal *South African Outlook*, which was founded in 1870 at Lovedale.

Society (based in Braamfontein, Johannesburg). Apart from these locations, the places of writing also include Pacaltsdorp (an old missionary station of the London Missionary Society – today a township of George), East London, Umtata and Blythewood in the former Transkei (all Eastern Cape), as well as Montagu (close to Worchester in the Western Cape). However, in the 1920s, which includes 20 different authors, the letters comprise selected correspondence of the Wesleyan Methodist Church of South Africa (Johannesburg), the Union Congregational Church (Grahamstown), the Brownlee Mission (King Williams Town), the Native Affairs Commission (Pretoria), the Juvenile Musical Society, the Witwatersrand General Council of the South African Party (addressed to the mayor of Johannesburg, John Wesley O'Hara), the East London Hospital Board and the Lovedale Institution (once again addressed to the Rev. J. Henderson). Many female authors were included as part of the correspondence of the Juvenile Music society, including Ivy. E. O'Hara (wife of the mayor of Johannesburg), Grete Pollak, Winifred Pomfrit, Ashley Hartley and Sarah Cohen. The business letters of the 1930s are written by 15 authors and include selected correspondence of the Lovedale Press, Alan Paton (a well-known literary and political figure), and Dr. P. W. Laidler, the erstwhile Medical Officer of Health for the Public Health Department (East London). The letters to Dr. Laidler were from, e.g. the South African Institute of Race Relations (Johannesburg), the Resident Commissioner's Office (Mafikeng), The Medical Association of South Africa (Cape Town), and W. T. Jackson, headmaster of the Mafikeng High School. Other areas in the country were the business letters of the 30's were written include East London, Pretoria, Durban and Amanzimtoti (near Durban). The 1940s, which is represented by the letters of 10 authors, include selected correspondence of the Royal Yacht Club, the Lovedale Press and James Gray, an analytical and consulting chemist in Johannesburg. Other places of writing include towns in the Natal area. The data for the 1950s hail from e.g. Heidelberg (south-east of Johannesburg), Johannesburg itself and Alice, and comprises 3 authors, including the secretary of The Christian Council of South Africa (Heidelberg), as well as the manager of Monotype Machinery (S.A.) Limited (Johannesburg). As a result, there are altogether 58 different authors (male and female) of business letters for Period 2.

As far as the social or private letters of Period 2 are concerned, the 1910s include 7 authors, such as John X. Merriman and his acquaintances Jane R. and Henriette Cochet, as well as a lady named 'Bella', the writer of a friendly post card.

The places whence these letters came include Port Elizabeth and a nearby town, Addo, as well as East London (all Eastern Cape). The private letters of the 1920s comprise selected correspondence of two literary figures, Alan Paton and Pauline Smith; the locations from which these letters were written include Illovo Beach (on the South Coast of contemporary KwaZulu-Natal) and Cape Town. Selected correspondence of Alan Paton and Pauline Smith also constitute the texts for the 1930s, together with selected correspondence from two other authors from the Royal Yacht Club. The places of writing for this decade are Johannesburg, Pretoria, Boksburg (east of Johannesburg) and Broadstone (Dorset – while Pauline Smith was in England). Paton and Smith's letters can also be found in the data for the 1940s (which has 10 different authors), not only written in Johannesburg and Dorset, but also in New York, USA, while Paton attended the Conference of Christians and Jews in 1946, as well as East London, Queenstown, Umtata and Durban. Some of the letters in this decade are addressed to Dr. and Mrs. Sheppard of Lovedale. Paton and Smith's sets of correspondence are also included in the 1950s private letters data (which have 6 authors), along with other letters between members of the Lovedale Press and a set of letters from a father, Jack Penn, addressed to his son, John, in which he offers him advice on life, love, moral matters, school, etc. (written in Johannesburg). The social letters of Period 2 therefore has 29 different authors.

The newspapers that represent the news register for Period 2 (31 873 words) include, for the 1910s, the *Diamond Fields Advertiser* (Kimberley), and for the 1920s, *The Star*, *The Sjambok* (both Johannesburg), the *Natal Mercury* and the *Natal Advertiser* (both Durban). The 1930s comprises news articles from *The Outspan* (a satirical Johannesburg publication exposing local corruption), as well as the *Cape Times* and the *Diamond Fields Advertiser*. On the other hand, the *Cape Times* and *The Star* comprise the data set of the 1940s, and *The Cape Argus* covers the 1950s. There are therefore altogether 8 newspapers represented in the news register of Period 2, covering the Northern and Western Cape, the Transvaal and Natal.

The total word count for fiction and other narrative writing in Period 2 is 30 568 words. The 1920s comprise Pauline Smith's short story *The Sisters*, together with Sarah Gertrude Millin's short story *A Sack for Ninepence*. The 1930s consists of extracts of Mary Byron's *Neighbours*, William Plomer's *When the Sardines Came* and Frank Brownlee's short story *Badeni's Bank Note*. Furthermore, an extract from *Cry, the Beloved Country* by Alan Paton, as well as *The Rooinek* (a short story) by

Herman Charles Bosman represent the 1940s. Dan Jacobson's short story *A Day in the Country* and two short stories by Nadine Gordimer, namely *Is There Nowhere Else Where We Can Meet?* and *Ah, Woe Is Me*, typifies the 1950s.

The non-fiction register of Period 2 contains 37 091 words. The data for the 1910s include selected non-fiction articles from *Outlook on a century: South Africa 1870-1970*, such as 'Recruiting in the Transkei' (by Godfrey Callaway), 'Dutch Reformed Church' (by E. Jacottet) and 'The Mendi Disaster' (by 'An Officer of the S.A. Native Labour Contingent'). For the same decade, extracts from J. S. Marwick's *Natives in the larger Towns* and S. G. Rich's *Notes on Natal* were used in the corpus. The 1920s also include selected articles from *Outlook on a century: South Africa 1870-1970*, e.g. 'Gratitude' (by A. W. Baker) and 'Missionaries and Colour Prejudice' (by an 'Observer'). As for the 1930s, the selections from the same publication include 'Night Schools' (by Edward W. Grant) and 'Powerless Workers' (by Tom Atkinson). 'Coal Mining', also by Tom Atkinson, together with 'African Hodge' (by K. A. Hobart Houghton) (from *Outlook on a century: South Africa 1870-1970)*, as well as an extract from the book, *The South African Opposition 1939-1945* (by Michael Roberts and A.E.G. Trollip), were used for the 1940s. Lastly, a biographical piece on 'Neil Macvicar' (by R. H. W. Shepherd) was also taken from *Outlook on a century: South Africa 1870-1970*, which, together with extracts from C. S. McLean's *Gold – South Africa's Major National Asset*, George Young's *The Shipping Links Binding South Africa and the United Kingdom* and Herman Charles Bosman's non-fiction work *Home Town*, comprise the data for the 1950s. Non-fiction is the register with the highest word count in Period 2, with news as one with the lowest word count. This concludes the discussion of registers for the historical corpus of SAfE from 1820 up to 1959.

The historical corpus has its limitations. It is clear form Table 2 that the amount of data (in terms of word counts) for each of the four textual genres is not perfectly balanced across the periods described above. However, the word counts for the news and fiction/narrative registers are similar, and likewise the word counts for letters and non-fiction are comparable. The total size of the corpus might seem like a relatively small number in comparison with the one million word corpora of British and American English (LOB/FLOB; Brown/Frown) used by e.g. Mair and Leech (2006). Seeing that the contemporary white L1 SAfE speech community was found to consist of a mere 3,1% of the entire South African population, as mentioned in

Chapter 2, it can be seen as a comparatively small community when contrasted against the much larger language populations of BrE and AmE, which could justify the representational credibility of a smaller corpus in the case of this study. However, the data for this dissertation cover a much longer time period (about 110 years longer) than e.g. Mair and Leech's (2006) study, and it reaches 90 years further back in time than that of Leech (2011).

3.2.2.2 A corpus of contemporary SAfE

For synchronic analyses of contemporary SAfE, the most recent version of the South African component of the International Corpus of English, ICE-SA, is used (see Jeffery, 2003). Those registers not comparable with the ones from the historical corpus were excluded from this study, namely non-printed student essays and exam scripts. This selection from written ICE-SA consists of 153 348 words in its written component, which is already noted in Table 2 and the spoken part of ICE-SA consists of 406 810 words. ICE-SA therefore does not yet contain the ICE standard word counts, which are 400 000 words for written and 600 000 words for spoken corpora (Aarts *et al*., 2002). The total word count of ICE-SA stands at 560 158, which is just over half the size of the other ICE corpora. The use of this corpus aids in conducting comparisons with the ICE corpora used in Collins (2009a,b), e.g. ICE-GB (Great Britain), ICE-AUS (Australia) and C-US (an ICE-equivalent corpus of the United States of America).

The selection from written ICE-SA was made from the following textcodes (Aarts *et al*., 2002:308)[120]:

(i) **Letters**
W1B-001 to W1B-015 for social letters and W1B-016 to W1B-030 for business letters

(ii) **Non-fiction**
W2A-001 to W2A-040 for academic writing; W2B-001 to W2B-040 for popular writing; W2D-001 to W2D-010 for administrative/regulatory writing; and W2D-011 to W2D-020 for writing on skills & hobbies

[120] Refer to Table 2 for the word counts of the registers.

(iii) **News**
W2C-001 to W2C-020 for press reports and W2E-001 to W2E-010 for press editorials

(iv) **Fiction/narrative**
W2F-001 to W2F-020 for novels & stories

The makeup of the spoken component of ICE-SA includes the following textual genres, which match the makeup for all ICE-corpora as stipulated in Aarts *et al.* (2002:308). Firstly, the dialogue component includes private (direct conversations and telephone calls) and public conversations (classroom lessons, broadcast discussions and interviews, parliamentary debates, legal cross-examinations and business transactions), and secondly, the monologue component includes unscripted (spontaneous commentaries, unscripted speeches, demonstrations and legal presentations) and scripted texts (broadcast news and talks, as well as non-broadcast talks).

3.2.2.3 Corpora of contemporary Afrikaans

Seeing that Afrikaans influence on SAfE needs empirical evidence, as mentioned in the Chapter 1, the following Afrikaans corpora are used to this end. A synchronic corpus of contemporary written Afrikaans, *Die Taalkommissie-korpus* (De Wet *et al.*, 2011), compiled by the standardizing body of Afrikaans, 'The Language Commission', is exploited to draw comparisons with the findings from written ICE-SA. This corpus consists of 57 million words and is balanced for a wide range of registers. For the comparison with spoken ICE-SA, the unpublished North-West University corpus of contemporary spoken Afrikaans is used. It comprises 72 000 words, evenly divided between informal dialogue and unscripted formal dialogue.

3.3 GENERAL ANALYTIC AND QUANTITATIVE METHODS

3.3.1 Tools

The corpora described above were analysed according to the corpus-based approach and its techniques, as described by e.g. Biber *et al* (1998). Both the synchronic and diachronic analyses were conducted electronically using *Oxford Wordsmith Tools* (version 5), by means of which concordance lists for all modal and quasi-modal verbs were complied. The concordance lists allow the analyst to view linguistic phenomena within their textual environment and conduct further search and sorting operations in order to eliminate irrelevant material and emphasise relevant data. By these means the lists of modals and quasi-modals were manually checked for accuracy. For diachronic frequency comparison of modals and quasi-modals, the log likelihood score indicating the statistical significance of the decrease or increase of an item were calculated by using the online log likelihood calculator by Paul Rayson[121] (see e.g. Rayson & Garside, 2000 and Rayson *et al.*, 2004). The interpretations of statistical significance rely on the following critical values, as stipulated in the calculator website[12]:

(i) 95th percentile; 5% level; $p < 0.05$; critical value = 3.84
(ii) 99th percentile; 1% level; $p < 0.01$; critical value = 6.63
(iii) 99.9th percentile; 0.1% level; $p < 0.001$; critical value = 10.83
(iv) 99.99th percentile; 0.01% level; $p < 0.0001$; critical value = 15.13

3.3.2 Methods

Normalisation, which is a corpus linguistic technique described in § 2.2.2, was a necessary method in this study. All raw synchronic and diachronic frequencies of modals and quasi-modals were normalised per 100 000 words in order to provide comprehensive comparative uniformity, not just SAfE-internally, but also with other varieties of English. Within the obligation and necessity cluster, the macro- and

[121] Available from: http://ucrel.lancs.ac.uk/llwizard.html

microsemantic analyses described later in this chapter were conducted for all the instances of the modals *must* and *should*, and the quasi-modal *HAVE to*, the results of which were normalised per 100 instances to aid comparison. For comparison between synchronic microsemantic results for contemporary spoken SAfE and contemporary spoken Afrikaans in § 4.3.2.2.3, all instances of deontic *must* and *moet/moes* with a second-person subject were analysed in both cases for modal strength and normalised per 100 instances to aid comparison with the normalised semantic results described above. The central investigations in this study, although orientated and triggered by the quantitative analyses, lies in the qualitative and, more specifically, the semantic interpretations.

3.4 INTERPRETATIVE, QUALITATIVE METHODS

3.4.1 Macro- and microsemantic analyses

As mentioned in Chapter 2, overlapping meanings exist within the modal and quasi-modal verb classes. These are illustrated in Table 3, wherein the modals and quasi-modals are classified into three broad or macrosemantic clusters. The selection of auxiliaries for this dissertation is based on the selections and groupings of Collins (2009a:7)[122], Biber *et al*. (1999:485) and Mair and Leech (2006:327-8)[123]. The fact that Table 3 groups the modals and quasi-modals into modal categories does not necessarily mean that these categories are discrete or that the auxiliaries are 'static' in their semantics. As explained in Chapter 2, gradience is a characteristic of the modal system, and the semantic boundaries in Table 3 should therefore be seen as indications of gradual, rather than abrupt transition into another broad semantic cluster.

[122] The quasi-modals *be bound to* and *be about to*, which are included in Collins' (2009a) analysis, were not included in this dissertation.

[123] *Used to* is not included in this dissertation, even though Mair and Leech (2006:328) included it in their list.

Table 3

Modal categories based on semantic clusters

Cluster	Modals	Quasi-modals
permission / possibility / ability	*can, could, may, might*	*BE able to*
obligation / necessity	*must, should, need, ought*	*HAVE to, (HAVE) got to, NEED to, BE supposed to, BE to, (had) better*
prediction / volition	*will, would, shall*	*BE going to, WANT to*

Apart from the general categorisations presented in Table 3, which relate to the sets of meanings inherently linked with each lexical item, each of the modals and quasi-modals can further be described as being systematically ambiguous (or polysemous) in terms of their type of modality, as mentioned in the previous chapter. These types of modality, represented by various terms in the literature, but referred to as either deontic, epistemic or dynamic in this dissertation (as explored in Chapter 2), apply to each of the three semantic clusters in the above table in the following way, according to Collins (2009a).

Within the permission/ability/possibility cluster, various meaning overlaps can be found. Permission meanings can be deontic (either objective or subjective) or dynamic (from an internal force), and ability meanings are prototypically dynamic (similar to dynamic possibility), whereas possibility meanings can be either epistemic, dynamic or deontic (equivalent to permission from an objective source) (Collins, 2009a:91-123). Within the volition/prediction cluster, prediction can be either epistemic, dynamic or deontic, whereas volition is described as being mostly dynamic (2009a:125-157). Within the obligation/necessity cluster, necessity expresses an epistemic meaning and obligation can be dynamic or deontic, which may also have a high or median degree of strength, as well as a subjective or an objective source of obligation (Collins, 2009a:33-89; Coates, 1983).

Selective macrosemantic analyses were done where it was necessary to explain noteworthy register-internal frequency patterns for some of the modals and quasi-modals, and all instances of *won't* were analysed to evaluate Bowerman's (2004:477) claim that it is used as a directive softener (as mentioned in § 1.1.2), but mainly, however, the systematic manual macrosemantic analyses were conducted only on the three most frequent auxiliaries of the obligation/necessity cluster in SAfE. This involved the classification of *must, should* and *HAVE to* into either epistemic,

deontic or dynamic modality, as discussed above. The semantic analyses proved to be complex, and included a degree of residual ambiguity, so the ambiguous instances were grouped separately under the label 'Indeterminate' (cf. Collins, 2009a). Furthermore, the microsemantic analyses in this thesis are specifically focused on the deontic (obligation) meanings of these three above-mentioned auxiliaries and include analyses for modal strength (the method of which is fully described later) and subjective/objective source. I analysed the auxiliaries separately for subjective/ objective source, in order to evaluate its significance to modal strength (cf. Collins, 2009:38) (as mentioned in Chapter two).

Methodologically, after I conducted the broad semantic classifications of the three above-mentioned modals, I checked it against the analyses and examples of Collins (2009a) to ensure accuracy and hence comparability. Only after these initial analyses were done, could I embark on the more intricate and micro-level semantic analysis of the deontic meaning in the obligation/necessity cluster.

The significance of the microsemantic analysis lies in the investigation of the trend toward monosemy in other native Englishes, as mentioned in Chapter 2, and in the claim of Bowerman (2004b:477) that *must* has less illocutionary force in SAfE than in other varieties of English, and is often a substitute for polite *should*. To recapitulate from the previous chapter, the degrees of modal strength in obligation are categorised in terms of high (required), median (supposed) and low (allowed) degrees by Halliday and Matthiessen (2004:620), where *must* is generally regarded to represent a high degree of obligation, *HAVE to* a median degree and *should* a lower or 'weak' degree (Collins, 2009a:26; Huddleston, 2002:177; Leech et al., 2009:86-88; 114-116). As also noted in Chapter 2, I prefer the terms 'high' and 'median' degree of strength to indicate respectively a stronger and a weaker force of obligation.

The above-mentioned macro- and microsemantic classifications required manual semantic analyses for all tokens of *must*, *should* and *HAVE to*, usually by interpreting a large portion of the context surrounding each of these tokens within the corpus via the 'source text' function of *Oxford Wordsmith Tools*. I concede that these semantic judgements involve gradient phenomena dependant on a context that is facilitated by corpora, which is valuable in this regard, but not always allows for conclusive answers, and are therefore open to many challenges. For this reason I set out to compile a set of parameters as basis for a more accurate interpretation involving high and median degrees obligation, which is presented in the next section,

seeing that no clear or comprehensive parameters were found in literature, except for some considerations of e.g. Coates (1983:33;36) and Depraetere and Verhulst (2008), as explained in § 2.4.1.1.

3.4.1.1 Microsemantic parameters

Table 4 lists the six parameters applied to the interpretation of either a high or median degree of obligation as part of the analysis of deontic *must*, *should* and *HAVE to*.

The first parameter I chose in this regard was the subject and main verb of the clause – which is suggested by Collins (2009a:37) and Coates (1983:33-6) to link with the idea of subjective/objective source of obligation. However, as mentioned earlier, the source of obligation constituted a separate analysis, in order to evaluate this link. This parameter was therefore not applied in isolation from the other parameters. Coates' (1983:33;36) features of a stereotypical core use of *must* with stronger force (often through the combination of subjective source), are the bases of this parameter. These features can therefore also be present for an auxiliary with a high degree of obligation, independent of its source.

If a subject was animate, apparently enjoyed a higher degree of status or authority than the addressee (cf. Depraetere & Verhulst, 2008:15-16), if the clause contained personal pronouns like *you* and *I* (and especially if the subject was in the second person), and if the main verb was an activity verb and was in the active voice, the instance was marked for a possible high degree of obligation. On the other hand, if the subject was inanimate (as depicted by e.g. the impersonal pronoun, *it*) or, in the instance of an animate subject (often in the third person), portrayed no status difference to that of the addressee, and if the main verb was stative and in the passive voice, the case was marked as a median degree of obligation.

In the close proximity to the auxiliaries, certain collocating adjuncts or expressions, as described by Halliday and Matthiessen (2004), served as the second parameter for interpreting the force or degree of the obligation in *must*, *should* and *HAVE to*. These include adjuncts with an interpersonal metafunction, which, as part of the mood or comment structure of the clause, express meanings of modality and intensity (2004:125;129).

Table 4

Parameters for high and median degrees of obligation

Parameter	High degree of obligation	Median degree of obligation
1. Subject and main verb	• Animate • Personal pronouns (second person subject) • Indication of status/authority • Activity/agentive verbs • Active voice	• Inanimate • Impersonal pronoun *it* (plus other third person subjects) • No difference in status • Stative verbs • Passive voice
2. Adjuncts/ Expressions	• High probability • Counterexpectancy • High usuality • Total or high intensity • Stronger comment/mood adverbials • Polarity	• Low or median probability • Low or median usuality • Low intensity • Weaker comment/mood adverbials
3. Mood of nearby clauses	• Imperative	• Indicative/declarative
4. Polarity/ Temporality	• Negative polarity/negation • Present or future time	• Positive polarity • Past time
5. Pragmatic functions	• Motivated by social norms • Motivated by emotions • Habitual obligation • Face-threatening situation • High degree of gravity or no possibility of non-compliance/non-actualisation	• Unmotivated by social norms • Unmotivated by emotions • Intermittent obligation • Polite or friendly situation • Hedging • Irony • Little or no gravity or a possibility of non-compliance/non-actualisation
6. Broad textual analysis	• Interpersonal content • Topic/intention, e.g. legal/regulatory/instructive/ warning	• Hypothetical/philosophical content • Topic/intention, e.g. opinion/advice/suggestions/ wishes/ socialising

Halliday and Matthiessen (2004:125;128) refer to certain adverbs serving as mood adjuncts of modality, all of which either denote a low, median or strong degree of probability or usuality (the latter of which ties in with Collins' notions on temporality [2009a:24-25] and Myhill's subfunction of habituality [1995:163]). I interpreted those adjuncts classified as low or median as indicative of a median degree of obligation and naturally high-degree adjuncts were flagged for a higher degree of obligation. Probability adjuncts of a high degree are e.g. *certainly, definitely* and *no way*, whereas the low or median degree is expressed by e.g. *probably, possibly, perhaps* and *maybe*

(Halliday and Matthiessen, 2004:128). Usuality of a high degree can be expressed by e.g. *always* and *never*, and the low or median degree by e.g. *usually, sometimes, occasionally, seldom* and *rarely* (2004:129). Halliday and Matthiessen (2004:129) further describe mood adjuncts of intensity, which can either denote a total (e.g. *totally* and *utterly*), high (e.g. *quite*) and low (e.g. *scarcely* and *hardly*) degree of intensity or an exceeding (e.g. *even, actually, really* and *in fact*) or limiting (e.g. *just, simply, merely* and *only*) level of counterexpectancy. Both exceeding and limiting levels of counterexpectancy were interpreted to denote a high degree of obligation.

The concordance lines were also searched for modal adjuncts that serve as comment or mood adverbials. Comment adverbials denoting a median degree of obligation may include e.g. *please* and *kindly* (linking with politeness in parameter number 5) and mood adverbials in this category are e.g. *probably, perhaps* and *maybe*. On the other hand, comment adjuncts such as *frankly* and mood adjuncts such as *certainly* and *really* usually denote a higher degree of obligation (Halliday and Matthiessen, 2004:354-355). Some degree of overlap can be seen to exist within differing types of mood adjuncts, which is largely due to the fact that some adverbs can also serve as adverbials, e.g. *probably* and *really* and are therefore mentioned in more than one list. Modal adjuncts of polarity can further aid in interpreting for degree of obligation, where e.g. *yes* and *no* may point toward an answer denoting a high degree of obligation to a matching question, and where e.g. *not* and *never* (the latter of which is also an adjunct of usuality) may point toward restrictiveness and therefore a high degree of obligation. Such black-and-white interpretations are of course not always watertight, which is why such cases where adjuncts were found were always checked against the other parameters before a decision could be reached.

Expressions equivalent to the conveyed meanings of many of these adjuncts were also helpful for interpretations. For example, there are some expressions equivalent to adjuncts denoting degrees of probability, e.g. *It is certain* (high), *probable or possible that...* (both lower), and polarity, *I am certain/sure that...*, *It is true that...* (high) (2004:148-149). Halliday and Matthiessen further emphasise the role of the metaphorical in expressing modality, which can be done by means of expressions of personal opinion, conviction or stance (which may aid in interpreting high and median degrees of obligation). Examples of such expressions are: *It is obvious that...* (equivalent to the adverb *obviously*), *Nobody tries to deny that...*, *There can be no doubt that...*, *Anyone will agree that...*, *Everyone knows that...*, *No sane*

person would pretend that... not..., Common sense determines that..., etc. (2004:616-617). Adverbs conveying a sense of urgency, such as *immediately*, *now* or equivalent fixed expressions like *at once* were also interpreted as indicators of high-degree obligation. The whole of parameter two proved to be a particularly useful one in interpreting the strength of obligation, as many of these adjuncts or expressions acted as reinforcement for other parameters. Various examples of such adjuncts and expressions mentioned above were found in the diachronic corpus, including: *perhaps* (median), *it is suggested that* (median), *seriously* (high), *no doubt* (high), *it is very important that* (high), *be compelled to* (high), etc.

The third parameter is that of the mood of a collocating clause in close proximity to those clauses in which the modals *must* and *should* and the quasi-modal *HAVE to* occur. If such a clause was in imperative mood, the instance was marked for a possible high degree obligation, whereas in cases of regular indicative or declarative mood being present in surrounding clauses, the instance was seen as a candidate for median obligation (see e.g. Halliday and Matthiessen, 2004:78; Verstraete, 2007:51-53). This of course does not mean to imply that face-threatening acts cannot occur by means of declarative or interrogative clauses, so this parameter in particular was not relied upon in isolation; a case was not marked for either degree of obligation after all possibilities were not considered.

The fourth parameter is related to Collins' suggestions involving negation and temporality (2009a:24-25). Through his examples and analyses when discussing theoretical preliminaries to, and results from, his study, Collins (2009a:25-27;41) implies that the interpretation of semantic strength is purely contextual, but does not offer clear criteria by which means to conduct such interpretations. Collins however hints that the source of negation in negated modals or quasi-modals, whether intrinsic or extrinsic (which is analysed separately in this dissertation as objectivity and subjectivity), could aid in analysing the propositional 'value' and hence the temporality of a modal utterance (2009a:24). This gives the impression that (1) negated modals or quasi-modals may be regarded as carrying a stronger degree of expressions of permission or obligation (negative polarity) and (2) modal expressions grounded in the present and future time (posterior proposition) may indicate stronger degrees of obligation than those situated in the past time (anterior) (2009a:24-25). Some modal adjuncts of polarity and mood adjuncts of usuality, as mentioned above, may contribute to the interpretation of such cases. Although these statements are

never overtly made by Collins (2009a), they offer useful suggestions on possible criteria for modal strength, and especially strength of obligation. This of course does not mean that positive statements (positive polarity) cannot denote a high degree of obligation (cf. Coates, 1983:39), which is why this parameter cannot be relied on in isolation from other parameters.

The following example, which was interpreted as an instance of a high degree of obligation, illustrates the usefulness of parameters 1, 2, and 4 in such an interpretation. Note that in the discussion of examples, the label 'P' (for 'parameter') plus the relevant number will be used to show which parameter applied in each case, e.g. 'P1' for 'parameter 1'.

(37) That we believe, but we fear he is comparatively impotent in restraining the various Kraals under him; and we are forced to repeat that we **must** not in considering this question, lose sight, for a moment, of the maxim, that a Chief is responsible for the conduct of his subjects. (1850s; News)

In example (37) the following factors indicated a high degree of obligation: animate subject, *we* as a personal pronoun, active voice (P1), expressions resembling adjuncts that express an exceeding level of counterexpectancy (*for a moment)* (P2), negative polarity and future time (P4), which overall is reinforced by both a verb (*forced*) and its complement verb phrase (*to repeat*) indicating a high level of intensity.

The fifth parameter relies on Myhill's (1995) reference to pragmatic functions in subcategorising the force of the modals and quasi-modals of obligation, as well as Brown and Levinson's (1987) ideas of politeness and face threatening acts. Myhill (1995:163) suggests that there are three subfunctions of a high degree of obligation, which include (1) obligation motivated by social norms and (2) emotion, as well as (3) habitual obligation. He exemplifies these subfunctions via various modals and quasi-modals as follows: (1) *Since you broke it, you must pay for it*, (2) *You've just got to help me!* and (3) *I have to take the bus to work every day*[124]. If the reverse of

[124] Myhill (1995:166) writes that "*got to* and *have to* are particularly associated with, respectively, obligations motivated by emotions and habitual obligations, and early *must* rarely had either the 'emotional' or the 'habitual' function". Therefore "as 'emotional' and 'habitual' obligations have come to be expressed more frequently, the language has switched to using particular markers associated with these functions, rather than expanding the usage *of* the old marker *must,* which was not typically associated with these functions" Myhill (1995:166; original emphasis).

these three subfunctions were noticed, e.g. if the obligation was found to be unmotivated by social norms or emotion and if it appeared to be intermittent or incidental in any way, the instance was marked for a median degree of obligation. As mentioned earlier, certain adjuncts assisted in identifying such subfunctions. I also interpreted a high degree of obligation in cases where the intention of the source of obligation could be seen as face threatening in any way. The term 'face threatening act' (Brown and Levinson, 1987), as mentioned in § 2.2.1.2.2, points to an act (speech act) in verbal or non-verbal communication that by its nature "run contrary to the face wants of the addressee" in the sense of 'keeping one's face' or sense of social security (1987:65). In contrast to this, if the social setting could be interpreted to be obviously friendly, polite or non-obstructive (not face threatening) in nature, median obligation was identified. Instances of polite or formulaic hedging aided in this process. Where a degree of irony could be identified, as e.g. in jokes between friends, the obligation was interpreted as functioning on the weakest end of the strength continuum, i.e. the median obligation label was given. Further considerations are whether non-compliance with the obligation (also termed non-actualisation or non-fulfilment of the obligation) will have a degree of gravity of consequence or not, or whether there is indeed a possibility of non-compliance or not, which is based on the criteria of Depraetere and Verhulst (2008:15) to measure the strength of obligation (§ 2.4.1.1; see also Sweetser, 1990:54; Smith, 2003:242; Collins, 2009a:45-6). An utterance was considered a high degree of obligation if the obligation is inescapable, i.e. impossible not to comply with (as in regulatory or legal contexts) or if there would be serious consequences if the obligation was not complied with (cf. Depraetere and Verhulst, 2008:15). Cases where non-compliance would have little or no gravity, or where non-compliance is possible, were naturally considered indicative of median obligation. The following example illustrates the application of this parameter to indicate a median degree of obligation, where the addressee is directly acquitted of any social pressure to perform the request of the speaker (P5).

(38) We had a great w/e away with the SY's. More on a social basis – so no pressure to **have to** do this, that or the other. (1990s; Fiction/narrative [W2F-005])

The sixth and final parameter is that of broad textual analysis and is the most generalised parameter in the case, which is why it is last on the list. It was found that, especially in the cases where parameters 1, 2, 3, 4 and 5 were not conclusive, the force of the obligation was often recoverable from the broad thematic analysis of the text. Those cases identifiable as median obligation were usually found to contain philosophical or hypothetical content (which links with Bybee's [1995:513] observation that the effect of the hypothetical meaning is often the weakening of the obligative force [§ 2.4.1.2]), as well as featuring personal opinions, friendly or neutral advice and suggestions, or merely projecting future wishes as part of the act of socialising (which links with the notion of Brown and Levinson [1987:65] that advice and suggestions are at the lowest end of the scale of face-threatening acts [cf. § 2.2.1.2.2]). If upon close examination the context surrounding the instances of *must*, *should* and *HAVE to*, whether clausal or broad textual, revealed interpersonal contact or interaction, and if the topic or intention of the text centred on legal, regulatory or instructive matters, or contained a warning, the modal was marked for a possible high degree of obligation.

Unsurprisingly, sometimes these interpersonal or interactive kinds of behaviour as captured by corpora are not clear-cut to analyse in written texts. Methodologically, if one or more of the parameters proved to apply to the individual case, the instance was marked for the relevant force of obligation, but, naturally, the more parameters applied, the higher the likelihood of either a high or median degree of obligation prevailed, so the instance was only finally marked after all the parameters were properly considered. Consider the following example from the historical corpus used in this study:

(39) My darling old Mary

We had such a splendid day up the mountain yesterday. You were the only person I wished was up there with me. ... Some day you **must** come up with me. You can have a horse. (1890s; Private letter)

Example (39) illustrates *must* in its use of a median degree of obligation, despite the fact that there are some of the parameters present which would indicate a high degree of obligation, such as: an animate subject, the personal pronouns *you* and *me*, the activity verb *come*, active voice (P1), a projection into future time (P4), and the

interpersonal content of the text (P6). The parameters that match the criteria for median obligation essentially overrule these factors – they are: no difference in status (P1), the presence of the adverb *some day* (contemporary 'someday') which lowers the degree of obligation (P2), positive polarity (P4), unmotivated by social norms or emotions, intermittent obligation, a polite or friendly situation (as this is a social letter between two friends of equal social rank and gender – indicated e.g. by the opening of the letter) (P5), and the fact that this is a suggestion referring to an indefinite or hypothetical time and situation (P6), strengthened by *some day*. In this case the expression *you must* is more one of a friendly wish (P6) or invitation equivalent to 'I would like you to', 'I wish for you to' or 'I invite you to'. Therefore, an example such as this one would additionally count as having a subjective source of obligation (the author), which indicates a non-prototypical use where a weaker/median degree of obligative force coincides with a subjective source.

Such analyses were conducted within each token of *must*, *should* and *HAVE to* for each period in the historical corpus. Apart from the more certain cases, it was found that some grey areas still seemed to arise, which is not unusual where gradient phenomena are concerned (see § 2.4.1.1). The six-parameter system aided in these difficulties to a large extent, but in a marginal number of cases, the uncertain occurrences were grouped under the tag 'Indeterminate', as also mentioned earlier.

As for the subjective/objective distinction, each case of *must*, *should* and *HAVE to* was marked according to the orientation of the source of the obligation. Coates' (1983:33;36) criteria of an animate subject, along with the inference of speaker involvement (where the speaker is interested in getting the subject to perform the action), were used to mark for a subjective source. The opposite of these criteria (inanimate subject and no speaker involvement), along with the inference that general circumstances are the source of the obligation (see Collins, 2009a:35), were used to mark cases of objective source. Depraetere and Verhulst's (2008:16) distinction between discourse-internal and discourse-external sources where also a fairly productive consideration on this account (cf. § 2.4.1.1), but not in all cases. These analyses for source indeed proved to be less complex than those for force. Example (40) illustrates an objective source of deontic modality, where no direct source can be identified (together with a case marked for median obligation [a rarer case in Period 0], as e.g. indicated by the use of the passive voice (P1), the presence of hypothetical content (P6) and the use of the phrase *our opinion is* (P6), which lowers the intensity

of the obligation).

(40) ...our opinion is, that men **must** be warned as well as edified. (1820s; News)

3.5 CONCLUSION

The research design of this dissertation, i.e. the corpus-based approach, will facilitate reliable empirical analyses (as opposed to anecdotal descriptions) from which hypotheses can be derived. The study population of this empirical study is the linguistic community of English users in South Africa who are descendants of the British Settlers (in the diachronic corpus), except for the input population who are the Settlers themselves, with the minor addition of some English-educated bilingual or multilingual users of other mother tongues (within ICE-SA). The Afrikaans community acts as another study population, with which to draw comparisons with the SAfE community.

Diachronic and synchronic corpora for SAfE were used, but only synchronic corpora for Afrikaans. The diachronic nature of the written historical corpus of SAfE, together with its division into different registers (social and business letters, news, fiction and other narrative texts and non-fiction) enables specific searches in each historical period and within each register. This promotes the construction of a developmental picture across time and context. Synchronic analyses complete the historical picture in both spoken and written registers. As suggested in the previous chapter, the dual chronology of the corpora serves a complementary function in the analysis of text and context.

Quantitative analyses were done by means of various corpus analysis tools such as taggers and parsing software, as well as by using corpus-based techniques, e.g. normalising frequencies. Qualitative analyses were done manually and involved both macro- and microsemantic classifications. Various sources, including e.g. Collins (2009a;b) serve as controls for checking macrosemantic analyses, whereas a newly compiled set of six parameters aids in microsemantic analyses. The results obtained from applying the methodology will be presented in the next chapter.

CHAPTER 4

RESULTS

"When *I* use a word ...it means just what I choose it to mean – neither more nor less." (Lewis Carroll, 1897)

4.1 INTRODUCTION

Humpty Dumpty's words in Lewis Carroll's *Through the Looking Glass: And what Alice Found There* expresses something of the semantic abstruseness of words, as well as the possible degrees of certainty a language user might attain on the continuum of meanings. The previous chapters prepared the way for the focus on SAfE modals in all their semantic complexity, albeit not within a fictional world, but in real usage.

Chapter four presents the quantitative and semantic results of this thesis. The section on general results (§ 4.2) will consider both the diachronic and synchronic trends detected through analysis of the corpus data, as described in the previous chapter. The patterns within and among the frequencies of the modals and quasi-modals of SAfE will be compared to those of British and American English for diachronic data, with the addition of Australian English and New Zealand English for synchronic data. The emphasis will fall somewhat more strongly on comparisons between the trends of these varieties over the 20th century. Furthermore, synchronic Afrikaans modal and quasi-modal frequencies will be compared to those of SAfE, in order to account for unique patterns of frequency.

Section 4.3 will report the diachronic micro-semantic results of the major roleplayers in the obligation/necessity semantic cluster, viz. the modal verbs *must* and *should*, and the quasi-modal verb *HAVE to*. Semantic interpretations for the patterns of

these auxiliaries will be linked to the influence of Afrikaans. This chapter will therefore not only give a detailed historical and contemporary description on SAfE modality and how its patterns compare with those of other Englishes, but also explore possible reasons behind the patterns observed. This will help to reach the goals and objectives stated in § 1.4.

4.2 GENERAL RESULTS

This section will present the frequencies of modals and quasi-modals in SAfE and draw comparisons with patterns in other native Englishes, as well as Afrikaans in the case of synchronic data. The diachronic results will be presented first, followed by the synchronic results, in order to provide chronological continuity in the development of the general argument.

4.2.1 Diachronic results

Section 4.2.1 will discuss the diachronic results of the analyses conducted via the methods described in Chapter 3. After the main quantitative trends observed from the diachronic data are noted in the following introductory section (§ 4.2.1.1), the discussion will turn to individual modals and quasi-modals, with the goal of finding explanations for those trends. Throughout the discussion, the results will be compared with the patterns of other English varieties – especially with the trends reported by e.g. Leech (2003; 2011), Mair and Leech (2006), Leech and Smith (2009), Leech *et al.* (2009) and Myhill (1995) for British English (BrE) and American English (AmE).

4.2.1.1 Introduction

Figure 4 shows some of the kinds of observable patterns of linguistic frequency change, as noted by Leech and Smith (2009:174), which will be used in this chapter. A trend of a steady increase (a) or steady decrease (b) is observable when the

frequency trend line drawn from the relevant data shows a steady, progressing slope between three data points, illustrated in Figure 4 by the two arrows as data points 2 and 3, with the first data point at the unmarked start of the trend line. Examples of trend lines portraying a changing rate of change are illustrated in (c), the deceleration of a decrease in frequency, and (d), the acceleration of an increase in frequency.

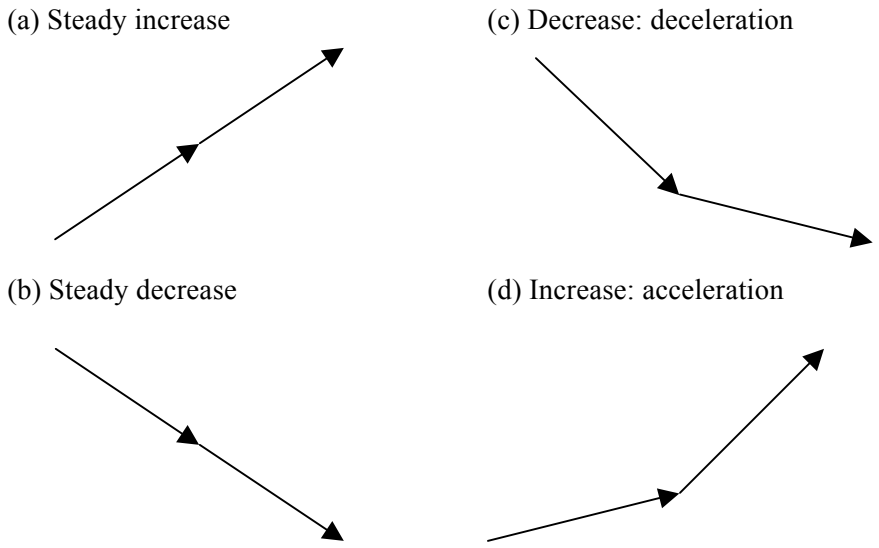

Figure 4
Some patterns of linguistic frequency change

The general diachronic results[125] from the analysis across all the written registers mentioned in Chapter 3 (letters, news, fiction/narrative and non-fiction) are presented in Figure 5 in terms of normalised numbers. To recapitulate from the previous chapter, in this section Period 0 represents the 1820s to 1860s, Period 1 the 1870s to 1900s, Period 2 the 1910s to 1950s and Period 3 the 1990s.

[125] See Appendix 2 for the comprehensive raw and normalised frequencies of the diachronic data.

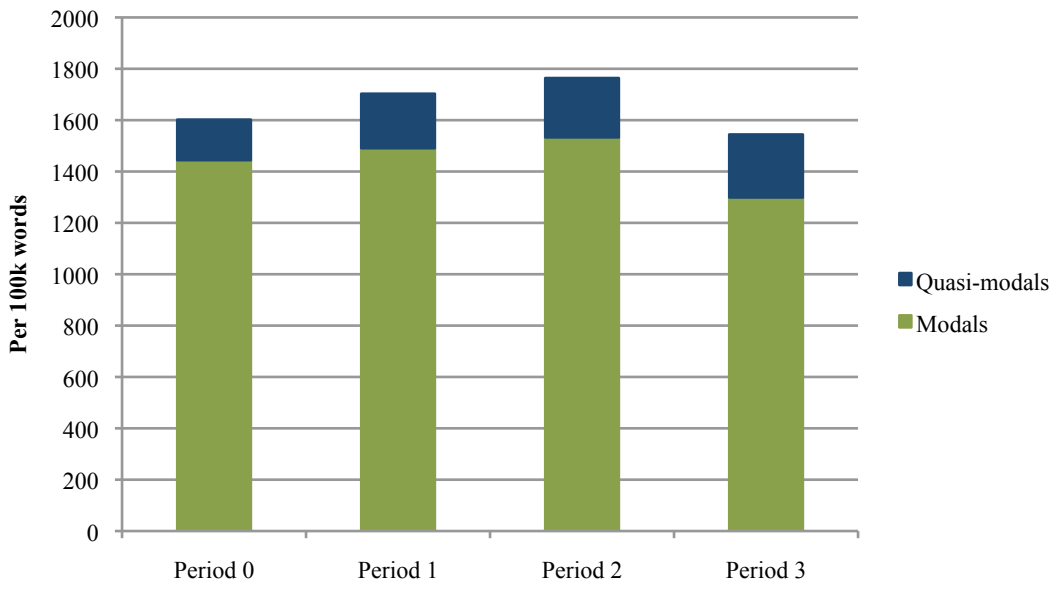

Figure 5

Overall normalised frequencies of modals and quasi-modals across time

In Period 0, a raw total of 2 803 modals and 306 quasi-modals were observed in 194 071 words, for rounded frequencies of 1 444 modals and 158 quasi-modals per 100 000 (henceforth 100k) words (a normalised total of 1 602 auxiliaries[126]). In Period 1, a total of 2 552 modals and 363 quasi-modals occurred in 171 186 words, giving a rounded frequency of respectively 1 491 modals and 212 quasi-modals per 100k words (1 703 normalised auxiliaries in total). In Period 2, a total of 2 164 modals and 326 quasi-modals occurred in 141 165 words, rendering a rounded frequency of 1 533 modals and 231 quasi-modals per 100k words (1 764 auxiliaries altogether). In the final period, Period 3, 1 992 modals and 376 quasi-modals occurred in 151 348 words, representing a rounded frequency of 1 299 modals and 245 quasi-modals per 100k words respectively (together amounting to normalised 1 544 auxiliaries).

The combined pattern for modal and quasi-modal frequency shows a gradual increase from the input data (Period 0) to Periods 1 and 2, but a decline from Period 2 to 3. Some internal variation is present as far as the distribution of modal and quasi-modal auxiliaries are concerned. Figure 6 shows the trend lines for the modals and quasi-modals in SAfE.

[126] Throughout this chapter I use the superordinate 'auxiliaries' to denote combined modals and quasi-modals, even though not all quasi-modals can always be regarded as auxiliaries.

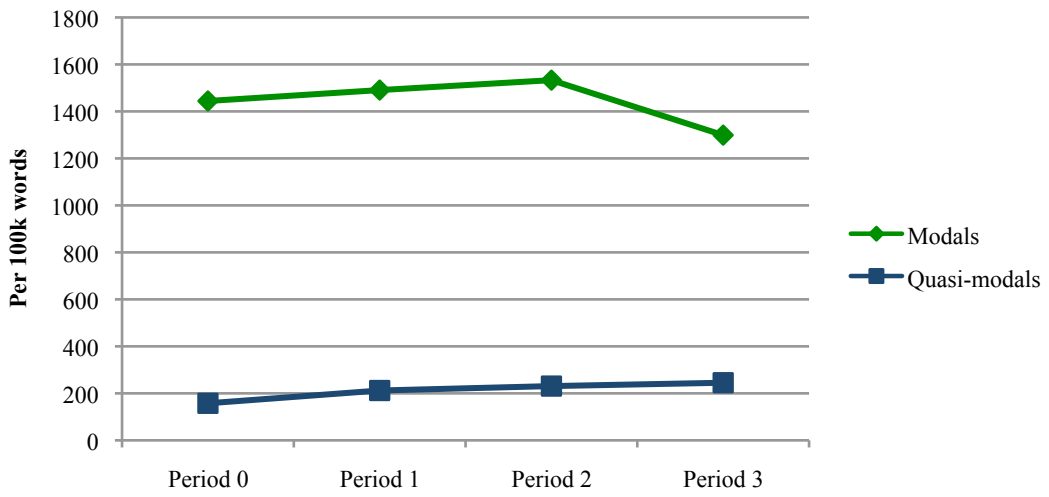

Figure 6

Frequency trends of modals and quasi-modals across time

The modals, being generally higher in frequency than the quasi-modals (in accordance with the tendencies reported for other native varieties (cf. Leech, 2003; 2011; Mair & Leech, 2006), undergo a gradual increase of 47 modals per 100k words from Period 0 to 1, and a slightly decelerated increase of 42 modals from Period 1 to 2 (however still remaining quite stable). The trend of gradual increase however changes from Period 2 to 3, with a decrease of 234 normalised instances between these periods; this carries the trend for the combined decline of auxiliaries from Period 2 to 3, as seen in Figure 5. Overall, a total decline by 145 normalised instances of modals occurs from the period of input to the last period.

This 20th-century pattern for SAfE is similar to the one reported for modals (in Chapter 2) by Leech and his associates for written British English and American English (see Leech, 2003; 2011; Mair & Leech, 2006:327; Leech & Smith, 2009; Leech *et al.*, 2009:72), where the modals were found to steadily decline across the 20th century (from 1931/1961 to 1991/2, which correspond roughly to the eras covered between Period 2 and 3 for SAfE). Nevertheless, when compared to other varieties according to percentages representing rates of decrease, the modals in SAfE decline by 15,3% from Period 2 to 3, whereas Leech *et al*. (2009:73) report that AmE and BrE combined show a decline by 10,6% during corresponding periods (1961 to 1991/2). The 1930s are however more representative of Period 2 – it fits into the middle of the period – and Leech's (2011) results are therefore more comparable with that of Period 2 in SAfE. When percentages are calculated from Leech's (2011:551-3)

data, it shows that from 1931 to 1991 the modals declined by 10% in BrE and from the 1930s to 1990s they declined by 8,6% in AmE. Leech and Smith (2009:187) furthermore report that the BrE variety shows only a slight decrease of modals from 1931 to 1961, with an accelerated decrease from 1961 to 1991. The general trend for the modals in SAfE appears quite similar to the one reported by Leech (2011:551) for BrE, where an increase in modal frequency is also present from 1901 to 1931, the former of which corresponds with the last decade of Period 1 in SAfE. The modals of BrE therefore only decline from about the middle of the 20th century onward (as seen above) – a slight decrease of 0,68% from 1931 to 1961, but an accelerated decrease of 9,43% from 1961 to 1991[127] (Leech, 2011:551) (which corresponds with the above-mentioned findings of Leech and Smith [2009]). In contrast, the modals of AmE already decline from the 1910s and across the entire 20th century, with some accelerations and decelerations in-between, and continue to decline after 1991[128]. Seeing that the periods of the historical corpus of SAfE are not analysed in terms of decades, but rather clusters of decades representing periods, it is difficult to draw exact comparisons with the specific dates of e.g. Leech (2011), but the fact that SAfE shows a somewhat greater loss of modals across the 20th century (5,2% more than in BrE and 6,7% more than in AmE for the 1931-1991 data [Leech, 2011]) is interesting. Despite the increase of modals in SAfE from Period 0 to 2, these auxiliaries still show an overall decline of 10,1% from Period 0 to 3. However, in a more recent study using the extended Brown family of corpora, Mair (2014) reports that the modals of AmE increase by 12% from the 1930s to 1961, but then decrease by 12,2% from 1961 to 1992, which closely resembles the above-mentioned situation in SAfE.

A similar pattern to that of written BrE and AmE emerges for the quasi-modals of SAfE. Leech *et al.* (2009:98), for example, report that the quasi-modals rise steadily during the course of the 20th century in both the above-mentioned large native varieties, and Figure 6 shows that the quasi-modals in SAfE display an overall and steady increase from Period 0 through to Period 3. The quasi-modals in SAfE already rise by 54 normalised instances from Period 0 to 1 and then decelerate in their rise, which occurs by 19 such units into Period 2 and by 14 units into Period 3. Overall, the

[127] Leech (2011:551) also reports that the trend of decline for the modals continues into the 21st century: they decrease by 6.93% from 1991 to 2006.

[128] By 6,45% from the 1990s to 2000s.

quasi-modals of SAfE show a total increase by 87 units per 100k words from the period of input to the last period.

Where percentages representing rates of increase are concerned, the quasi-modals show an increase by 6,2% in SAfE from Period 2 to 3. Leech *et al*. (2009:97) however report that the quasi-modals increase by a larger 9% in BrE and even larger 18,2% in AmE from 1961 to 1991/2, and Leech and Smith (2009:189) report that they show a decelerated increase between these two eras, compared with the greater increase observed from 1931 to 1961. This means that the increase of quasi-modals during the 20th century in SAfE occurs to an even a lesser extent than suggested by the percentages from 1961 to 1991 for the other two native varieties. Furthermore, in SAfE, a gross increase by 55,5% is present for quasi-modals from Period 0 to Period 3, which, with the addition of older data, is not comparable with the other varieties. The general implications are therefore that the overall rising trend of quasi-modals in SAfE corresponds to the overall trend in BrE and AmE, but that there is a slower rate of increase in SAfE across the 20th century when compared to the quasi-modals of BrE and AmE.

Because the modals do not gradually decline across the whole chronological spectrum, which would mirror the gradual increase of the quasi-modals, no interdependency between the trends of these two groups of auxiliaries is present, which corresponds to the findings of e.g. Leech (2003:235) and Mair and Leech (2006:327). This will also become apparent when individual relationships between modals and quasi-modals are explored in later sections. The overall rise of quasi-modals in SAfE could however be evidence of a degree of colloquialisation taking place (see § 2.4.2.1), which is proposed as one of the main reasons for the overall rise in BrE and AmE, as mentioned in Chapter 2 (cf. e.g. Leech, 2003; Mair & Leech, 2006; Leech & Smith, 2009; Leech *et al.*, 2009). However, the extent to which this process is at work across the 20th century is suggested by the SAfE data to be much less than in the other two varieties, but this will be further explored in the discussion on register-internal trends below, as well as in the synchronic results that are presented later in this chapter, where spoken data are included. Democratisation, the other reason for the rise of quasi-modals proposed for AmE and BrE, will be explored when the obligation/necessity group is discussed.

The general picture reported in Figure 5, involving the overall rise of auxiliaries from Period 0 to 1 and from Period 1 to 2, but an overall decline from

Period 2 to 3, corresponds with the general findings of Rossouw and Van Rooy (2012:13). There are however differences in the frequencies of modals and quasi-modals from those reported by Rossouw and van Rooy (2012), due to the fact that the quasi-modals *BE able to*, *(had) better* and *BE to* were analysed additionally, and *used to* was not analysed in this thesis, in order to align the results more with the current work of e.g. Collins (2009a). The much larger corpus size (659 770 words compared to 111 013 words in Rossouw and Van Rooy [2012:12]) further influences the results in the way that the patterns reported in this chapter can be expected to be more stable. The fact that a fourth register, non-fiction, was added to the entire corpus further broadens the scope of data for this thesis, while Rossouw and Van Rooy's (2012) findings are based on only three registers (letters, news and fiction). The data for Period 0 are also different, seeing that Rossouw and Van Rooy (2012) utilised the early Settler correspondence of Mesthrie and West (1995) (mainly comprising business letters) to represent this Period, and, as seen in Chapter 3, the current data for Period 0 exclude this set of correspondence and consist of new data for four registers. These differences relating to data and the selection of quasi-modals for analysis unfortunately necessitate that some aspects reported in Rossouw and Van Rooy (2012) need reconsideration.

The main difference between the results of this study and Rossouw and Van Rooy (2012) involves the frequencies of modals and quasi-modals in Period 0. In the Mesthrie and West (1995) data, only 7 raw tokens of quasi-modals compared to 396 modals in 22 483 words were found, whereas 306 tokens for quasi-modals compared to 2 803 modals were found in the data of this thesis for Period 0 (194 071 words). The low frequency of quasi-modals in Rossouw and Van Rooy's Period 0-data can be linked to the much more formal nature of the business letters in the Mesthrie and West data set. This means that, in the Rossouw and Van Rooy data, the proportion of quasi-modals, comprising 1,7% of the total auxiliaries compared to 98,3% for modals, differs from the proportion illustrated in Figure 5, where the quasi-modals cover 9,8% of the frequency for all auxiliaries and the modals 90,2%. The higher normalised frequency of modals in Period 0 in Rossouw and Van Rooy (2012:13) causes their pattern for modals to drop from Period 0 to 1, before rising again to Period 2, whereas the modals show a gradual increase from Period 0 to 1 and Period 1 to 2 in the data for this thesis. On the other hand, the much lower normalised frequency for quasi-

modals in their data for Period 0 (which converts to 31 instances per 100k words[129]) causes Rossouw and Van Rooy's frequency for quasi-modals to rise much more sharply from Period 0 to 1 (224 instances per 100k words) with 193 instances per 100k words, than in the current study's data for these periods, where quasi-modals rise by 55,5 instances per 100k words.

In the discussion up to this point, two main deviations from expected patterns have been identified in the comparison between SAfE and other native varieties of English, namely that the modals of SAfE decrease more than in BrE and AmE, and that the quasi-modals of SAfE increase less than in these other native varieties across the 20th century. These deviations call for further investigation, for which two strategies will be followed. Firstly, the general trends for the four registers will be considered. Secondly, the frequency shifts of the individual modals and quasi-modals and the effect these might have on the overall trends in Figures 5 and 6, will be explored. The individual modals and quasi-modals will be considered according to their general semantic clusters (as set out in Table 3 in Chapter 3) in order to aid SAfE-internal comparisons. Some references to the role of register-internal variation as well as semantic tendencies for the individual modals and quasi-modals will also be made where appropriate.

When comparing Figure 5 and its illustration of the gradual rise of the combined auxiliaries from Period 0 to 1 and 1 to 2, as well as the general decrease from Period 2 to 3, to the pattern observed for the separate registers, the latter is more complicated, but the trend line of letters follows a pattern largely similar to the general trend in Figure 5, as illustrated in Figure 7.1. This figure depicting overall auxiliary frequency over the various registers is offered before the separate trends for modals and quasi-modals are illustrated, to aid comparison with Rossouw and Van Rooy's (2012:14) similar presentation of combined auxiliary frequencies across registers.

[129] Rossouw and Van Rooy's (2012) data is normalised by 10 000 words, so the raw frequencies mentioned here were re-normalised by 100 000 words to aid comparability with the current study.

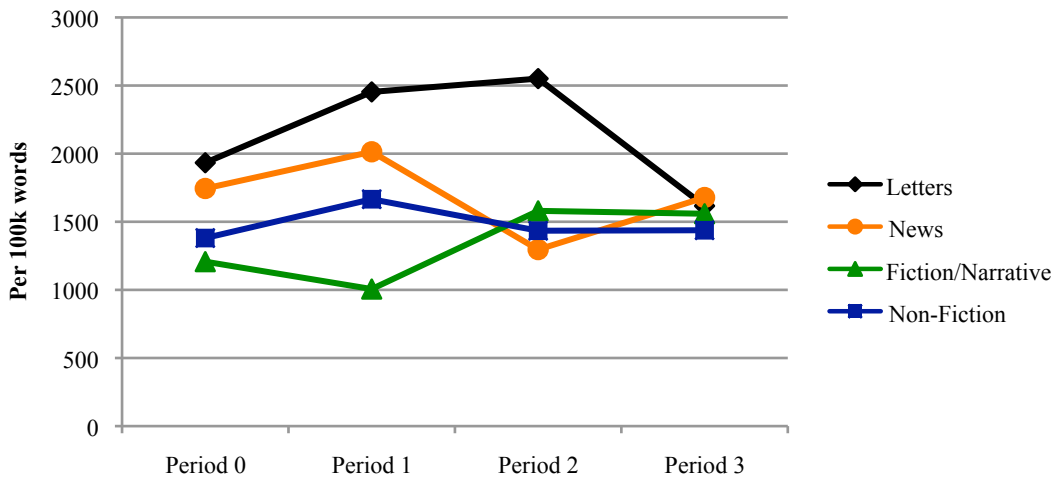

Figure 7.1

Frequency trends of combined modals and quasi-modals in the four separate registers across time

Letters indeed generally contain more auxiliaries than the other registers, except in Period 3 (1 616,4 normalised instances), where news has a slightly higher frequency (1 677,9 normalised instances). From Period 0 to 1, the auxiliaries rise sharply in the letter subcorpus, with a decelerated rise from Period 1 to 2 and a change of direction with a sharp drop from Period 2 to 3 (by 36,6%). It is probable that the sharp rise of letters between the first two periods can be accounted for on the basis that the data for Period 0 contain twice as many formal or business letters as private or social letters, but the data for all other periods contain more social letters. Much the same situation is present in the data of Rossouw and Van Rooy, as noted above, where, based on the findings of Biber *et al*. (1999:486) that the modals and quasi-modals are much more frequent in informal conversation than in news, fiction or academic writing, they assert that "if social letters share some of the properties of spoken registers, it will account for the difference between social letters in Period 1 and official correspondence in Period 0" (Rossouw & Van Rooy, 2012:15). Therefore, where the business letters are regarded to represent a more formal register and social letters a more informal one, it is a plausible explanation behind the auxiliary increase from Period 0 to 1, meaning that this increase is not due to an actual change in modal or register patters, but due to a property of the data sample. Examples (41) and (42) respectively illustrate the more formal nature of the business letters in Period 0 with

fewer auxiliaries, and the more informal nature of the social letters in Period 1 with a higher density of modal auxiliaries.

(41) ...that Your Excellency be pleased to take such other steps ...as Your Excellency **shall** deem meet & the nature of her unfortunate [sic] **can** require. (1820s; Business letters)

(42) Poor Ettie had been very ill, but when she wrote was much better. I hope we **shall** soon [have] good news from them. I **must** close now, dear sister, as it getting dreadfully late. ...I **shall** be so glad to hear from you again. ... Please give my to to Alice and tell her that I have written to her but have mislaid the letter and **can**'t sent it by post. (1870s; Social letters)

The second highest frequency of auxiliaries in Period 0 and 1 occurs in news, but before rising to the highest frequency among the registers from Period 2 to 3, it undergoes a sharp drop from Period 1 to 2 (by 35,7%). The frequency of auxiliaries in non-fiction is the third highest for the first three periods (after rising from Period 0 to 1, but decreasing again from 1 to 2), whereas its frequency in the last period remains stable (only rising by a slight 0,2%) to be the lowest of all the registers in that period. The register with the lowest frequency in the first two periods is the fiction/narrative register. It is also the only register in which auxiliary frequency declines from Period 0 to 1, although it rises from Period 1 to 2 (like letters) and stabilises into Period 3 (only dropping by 1,3% from Period 2 to 3).

It is interesting to note the trend of all registers except letters to move toward closer correspondence in frequency in Period 2 than in the preceding periods, where these registers where more distinct in their respective auxiliary frequencies. Letters however join this trend in Period 3, where all four registers move toward frequencies closer to convergence than in any other period: respectively normalised instances of 1 677,9 in news, 1 616,4 in letters, 1 558,9 in fiction/narrative and 1 437,7 in non-fiction. This trend is similar to the one reported by Rossouw and Van Rooy (2012:14). Even in a smaller corpus they found that the fiction and news registers already move toward similar frequencies in Period 2[130], and that letters (displaying a very similar

[130] The eras covered by Periods 1 to 3 in Rossouw & Van Rooy (2012) are the same as in this study.

trend to the one in Figure 7.1) also move toward this convergence in Period 3. Rossouw and Van Rooy (2012) suggest that "one may want to regard the convergence of the frequencies in different registers as confirmation of the trend across varieties of English towards colloquialisation", but that the decline in the overall frequency of auxiliaries (Rossouw & Van Rooy, 2012:13), as also seen in Figure 5, "is not consistent with colloquialisation, which would favour higher frequencies, and therefore alternative explanations must be considered" (2012:15) (cf. Mair & Leech, 2006). Paragraphs to follow will discuss the trends concerning colloquialisation and a lesser degree of register differentiation in terms of Figures 7.1, 7.2 and 7.3 in more detail.

The trends illustrated in Figure 7.1 are explainable via the trends for modals and quasi-modals in the different registers. Figure 7.2 shows the modal frequencies for the separate registers – displaying a more complex set of trends than the trend for modals in Figure 6. Figures 7.1 (combined modals and quasi-modals) and 7.2 (modals only) appear virtually identical. Because the modals of SAfE are generally higher in frequency than the quasi-modals (see Figures 5 and 6), the modal frequency patterns of the four registers therefore carry the trend lines of the combined auxiliary frequencies per register in Figure 7.1. More specifically, because of the generally high frequency of modals in letters, the modal trend followed by letters in Figure 7.2 can be seen as the main carrier of the trends for both the overall combined auxiliaries in Figure 5 and the modals in Figure 6.

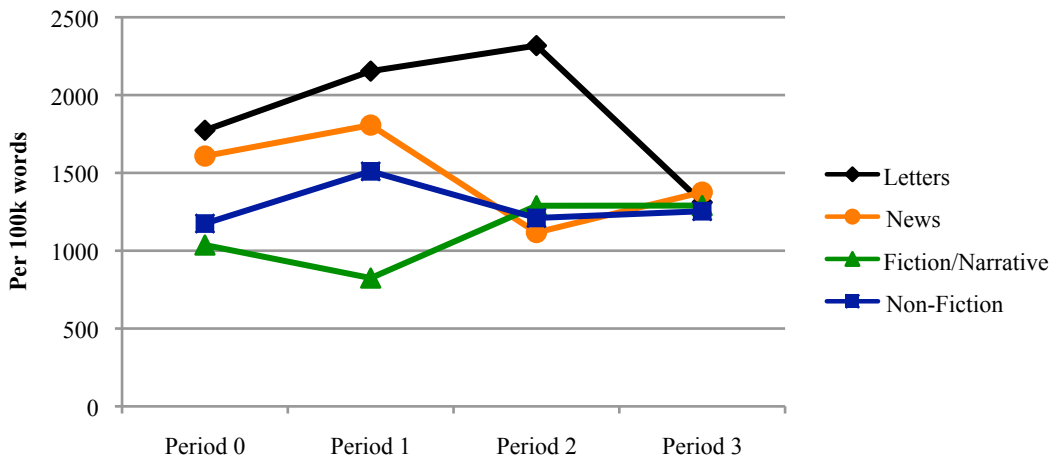

Figure 7.2

Frequency trends of modals in the four separate registers across time

The modal frequencies for the registers in Figure 7.2 show a trend toward converge in Period 3 (with normalised frequencies of 1 376,3 for news, 1 311,1 for letters, 1 290,1 for fiction/narrative and 1 253,7 for non-fiction), after being distinct in Periods 0 and 1, and after already showing a discernable trend toward convergence in Period 2 in all registers except letters. The letters register has the overall highest use of modals for Periods 0, 1 and 2, but it has the lowest frequency among the registers in Period 3 (dropping by a large 43,4% from 2 317,9 to 1 311,1 normalised instances from Period 2 to 3 respectively). This 20th-century drop of modal frequency in letters may be linked to the loss of ritualisation in letter writing, especially since the 19th century (cf. Barton and Hall, 2000; § 2.2.2.1). This particularly involves the decrease of hypothetical *would* in the letter register from Period 2 to 3, as will be discussed § 4.2.1.3.1, but also the decrease of formal *BE to* in letters over the same period (see § 4.2.1.4.3). This may point toward a degree of colloquialisation for letters over the 20th century.

When considering the overall lower frequency of quasi-modals compared to modals depicted in Figures 5 and 6, the effect of the register-internal trends of the quasi-modals is not nearly as influential on the trends of the combined auxiliaries. In fact, the quasi-modal frequencies in the separate registers are internally somewhat unstable, despite the fact that there is still an overall trend towards a slight rise from Period 0 to 3 (except for non-fiction, which drops by 10,7% overall). Figure 7.3 shows the trends for quasi-modal frequency in the different registers.

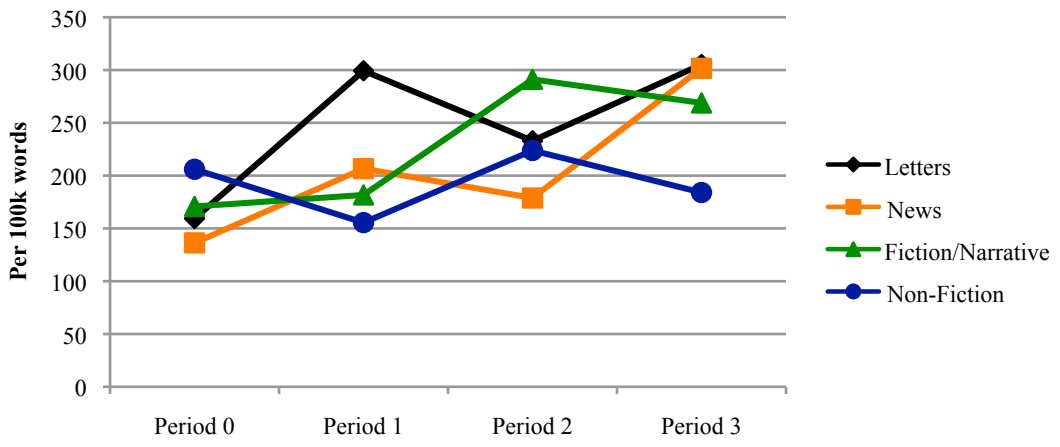

Figure 7.3

Frequency trends of quasi-modals in the four separate registers across time

The trends for letters and news are quite similar, even though the frequencies of quasi-modals in letters are generally higher for all periods than in news: both registers show an increase from Period 0 to 1, a drop from Period 1 to 2 (their frequencies in Period 2 are however still higher than in Period 0), and a rise again from Period 2 to 3 where these registers meet with near convergence of frequencies (305,2 normalised instances in letters and 301,6 in news). The reason behind the similarity of the trends found in letters and news may be that both these registers often contain references to the future, and internal changes in the frequencies of individual quasi-modals carry the trends for these two registers. Carrying the upward trend for Period 0 to 1 in these registers are the quasi-modals *BE to*, *WANT to* and *HAVE to*[131], whereas the decline of *be to* from Period 1 to 2 carries the trend for overall decline between these periods, but the continued increase of *HAVE to* slows down the rate of decline to some extent. The upward trend from Period 2 to 3 in these registers is carried again by rises in the frequencies of *WANT to* and *HAVE to*, as well as *BE going to*. The continued decline of *BE to* into Period 3 also slows down the rate of increase, especially for letters, which inhibits this register to rise to the same extent it did from Period 0 to 1. The slower rate of decline of *BE to* in news into Period 3 however causes the quasi-modals to rise to a larger extent in this register than in letters, with fewer resistance to the general trend of incline. Reasons behind these trends of these individual quasi-modals will be explored in sections to come.

On the other hand, the trends of the fiction/narrative and the non-fiction registers are also similar to each other from Period 1 onward, and contrast with the patterns in letters and news. Even though the quasi-modal frequency of the fiction/narrative register rises slightly from Period 0 to 1, it is not nearly as steep as for letters (or news), but then it rises from Period 1 to 2 to become the register with the highest quasi-modal frequency in Period 2. However, the quasi-modal frequency of fiction/narrative decreases again from Period 2 to 3, where it is again lower than in letters and news. Non-fiction starts out as the register with the highest frequency of all in Period 0 (206 normalised instances), but drops from Period 0 to 1 before rising from Period 1 to 2, and decreasing again from Period 2 to 3, reaching the lowest frequency across all registers in the last period (184 normalised instances); it is therefore the only register that declines in quasi-modal frequency across the entire

[131] See Table 21.2 in Appendix 2.

time span, as also noted above. The similarities in the frequency trends of fiction/narrative and non-fiction may also be accounted for on the basis of the patterns of individual quasi-modals. The different directions of the trends of fiction/narrative and non-fiction from Period 0 to 1 is influenced by the changing relationship between the frequencies of *BE to*: between these periods this quasi-modal rises slightly in the fiction/narrative register (by 12,7 normalised instances), but drops sharply in the non-fiction register (by 123,9 normalised instances); these patterns carry the general trends for these registers between the first two periods in Figure 7.3. Furthermore, the increases found in *WANT to* and *BE able to* for these registers from Period 1 to 2 carry the rising trends for all quasi-modals between the same periods, and the same situation occurs into Period 3, where the declines in be going to and be to for both these registers carry their downward trends for all quasi-modals. The relevance of these individual frequency shifts to the overall trends will be explored after a consideration of the 20th century trends of the registers, as seen in Figures 7.1, 7.2 and 7.3.

To recapitulate from § 2.4.2.1, Mair (2006:183) suggests that there is a trend toward a narrowing stylistic gap between registers in English as an effect of colloquialisation over the 20th century, as registers (mostly broad register categories of written and spoken language) simultaneously move toward a certain level of informality, i.e. their distinctions become 'bleached out'. This narrowing stylistic gap is even found to be evident among written registers by Biber and Finegan (1989) (essays, letters and fiction), the effect of which correlates with Mair's concept of colloquialisation, since the 'narrowing' is directed toward oral styles. Biber and Finegan's study (1989:487) is however based on the multi-dimensional approach, were e.g. such dimensions of language variation as 'information vs. involved production' are measured, whereas this thesis is not. When only the register-internal trends of modals and quasi-modals for Period 2 and 3 (as seen in Figures 7.2 and 7.3) are considered in terms of the general 20th-century trends for English, the following becomes apparent.

From Period 2 to 3, letters and news are the two registers that move toward convergence the most, i.e. their stylistic gap narrows to the greatest extent. As discussed above, this occurs for modals (Figure 7.2) where the frequency in letters drops and the frequency in news rises (the same as for combined modals and quasi-modals [in Figure 7.1]), as well as for quasi-modals (Figure 7.3) where the

frequencies in both letters and news rise. Fiction/narrative and non-fiction remain stable in their modal frequencies, i.e. these registers do not experience much change over the 20th century, which is a different pattern from that of letters and news, even though letters and news meet their frequencies in Period 3. Fiction/narrative and non-fiction also do not follow the same pattern as letters and news in terms of quasi-modal frequency – they decrease – but quasi-modal frequency for the fiction/narrative register is closer to the frequencies in news and letters in Period 3 than for non-fiction in the same period.

If we take the letter register as a representation of more informal language than news per se, the narrowing gap between these two registers for the modals point toward greater convergence of the styles of letters and news, but in the opposite direction of colloquialisation, since it is the frequency of letters that mostly move toward the frequency of news, i.e. a more informal register moves toward the more formal end of the spectrum. Since this is not a multi-dimensional study, it is not possible to prove this argument based on this evidence alone, but it is supported by the fact that modal frequency in letters also moves toward the frequencies of fiction/narrative and non-fiction at the same time, since these registers, especially non-fiction, tend to be more formal and conventionalised due to e.g. the editing process. On the other hand, the ritualised nature of letters in the earlier periods (as noted earlier) could have the effect that the letter register has moved away from more formal language in the 20th century, which is a different view than above. The following extract from a business letter written by Thomas Pringle in 1828 illustrates the formal and ritualised nature of letters in the earlier periods, especially in terms of hedging, as well as opening and closing formulae:

> Sir,
>
> The Rev W PEARS, the gentleman recently appointed minister of the Scotch Church at Baviaans River, Cape of Good Hope, ...requests me to state to you that the sudden death of a sister ...has affected his feelings and disconcerted his arrangements, that it will not be possible for him ...to embark before the 1st week of October. He therefore, most respectfully & earnestly requests that you will, if possible, so order matters in regard to his proposition that he may not be required to embark before the period I have mentioned. ...
>
> Permit me to solicit your obliging attention to his request (which I am well aware is not proffered without urgent cause) & by the favour of an early reply in order to set Rev PEARS' mind at rest on the subject.
>
> I have the honour to be your most obedient servant
>
> Thomas PRINGLE

Furthermore, the fiction/narrative register often contains speech representation (cf. § 2.2.2.1), i.e. depictions of an oral style, so it is relatively balanced between more informal (speech-like) and more formal (written) language. In fact, much the same applies to the news register (as discussed in paragraphs to follow) and the letter register, which contains both business and social letters, even if it contains somewhat more social letters (see Table 19 in Appendix 1). With non-fiction at the most formal end of the spectrum (cf. § 2.2.2.1), it is however still interesting that the modal frequencies for the four registers move toward a narrowed 'target' point that is closer to the frequencies in the edited texts. As noted above, it is quite beyond the scope of this thesis to endeavour to prove that the point of near-convergence in Period 3 (Figure 7.2) is in fact one at the more formal end of the spectrum. On the other hand it may also be that this 'target' point is somewhere in the 'grey' middle area of the formality scale. Either way, the evidence for colloquialisation in terms of the register-internal patterns for modals is not strong.

The only register with opposing patterns of decline and increase among the modals and quasi-modals from Period 2 to 3 is letters, where the modal frequency drops and the quasi-modal frequency rises. It was reported above that the overall rise of quasi-modals in SAfE (Figure 6) suggests a degree of colloquialisation, but that this occurs to a lesser extent than in BrE and AmE (cf. e.g. Leech, 2003; Mair & Leech, 2006; Leech & Smith, 2009; Leech *et al.*, 2009). Hence, the increase of quasi-modal frequency in letters may indeed indicate a level of colloquialisation in that register. The quasi-modal frequency in news therefore apparently rises to 'meet' the frequency in letters over the 20th century, which contrasts with the above-noted tendency for modal frequency in letters dropping to apparently 'meet' that of news. This may point toward a degree of colloquialisation of news, since the quasi-modal frequency in a more formal register (news) moves toward the frequency of a more informal register (letters). By Period 3, the gap in quasi-modal frequency between letters and news has narrowed considerably (Figure 7.3), and fiction/narrative has also moved closer to this 'target' point, with non-fiction lagging behind. When fiction/narrative is taken as a register containing more speech-like language than non-fiction, the narrowing stylistic gap between letters, news and fiction/narrative may indicate that that the 'target' point is at the more informal end of the spectrum, especially since news is rather balanced for levels of formality as well, as discussed below. However, given the fact that letters, news and fiction/narrative are all quite

balanced for varying formality levels, this may also point toward the 'target' being at the 'grey' middle area of the spectrum, just as for modals.

Mair (2006:69) notes that "[n]ewspapers are a genre that displays some stylistic differentiation, both between different newspapers and (within one and the same paper) between more formal subgenres such as foreign news and more colloquial ones such as sports reportage". He adds that news has become more colloquial over the 20th century, with "the intended stylistic effect" making "texts appear more dramatic, interesting, and accessible" and, "[involving] the reader emotionally" (2006:188). The news register of the historical corpus of SAfE is balanced in its different types of reportage on a wide range of subjects, especially in both Period 2 and 3 (including e.g. foreign and local news, as well as e.g. political, legal, social, scientific, entertainment and sports news) so one could argue that news in the 20th century data may contain internally varying levels of formality. The only difference between the news of Period 2 and 3 is that Period 3 contains newspaper editorials that comprise about 15% of all the news texts for that period, which is not really a large enough proportion to create a bias in the data, but perhaps just enough to make the news data for Period 3 slightly more informal than for Period 2. This is based on the report of Westin (2002:163) that contemporary newspaper editorials are becoming more informal (as noted in § 2.4.2.1). As far as this tendency goes, the fact that the news of Period 3 contains some editorials may account for the apparent trend toward more oral styles toward the end of the 20th century, at least as far as this trend involves the higher frequency of quasi-modals, keeping in mind that this trend is perhaps a product of the data sample, and perhaps not. Regardless, evidence for colloquialisation is not overwhelming, but the evidence is perhaps somewhat stronger in terms of the quasi-modals than the modals, especially in the news register.

The register-internal patterns for modals and quasi-modals in SAfE (separately, as well as combined), as carriers for the general trends of modals and quasi-modals in this variety (with modal trends being the main carriers), have been highlighted in the above discussion. As far as the modal and quasi-modal frequency patterns for each register are concerned, the internal fluctuations across the periods are in need of some explanation, which, together with explanations for the more general trends noted earlier, will be explored in terms of the frequency patterns of the individual modals and quasi-modals. Table 5.1 presents the normalised diachronic frequencies of the modal verbs for all registers, together with indications of statistical

significance by means of asterisks, based on log-likelihood calculation[132]. The same information is presented for the quasi-modals in Table 5.2.

Table 5.1

Overall normalised frequencies of modals over time

Period	*can*	*could*	*may*	*might*	*must*	*need*	*ought*	*shall*	*should*	*will*	*would*	Total
0	162,8	153,6	148,9	55,1	98,9	6,7	16,5	85,5	134,5	341,6	240,1	1444,3
1	185,8	143,7	139	74,2	109,2	9,9	13,4	62,5	165,3	313,7	274	1490,8
2	178,5	*178,5	111,9	56,7	124	6,4	7,1	56,7	132,5	345	*335,8	1533
3	***251,7	**131,7	129,8	**29,3	113,5	*1,3	*0,7	****8,5	*100,4	347,6	****184,5	****1299
Overall 0-3	****	ns[133]	ns	**	ns	*	****	****	*	ns	**	**

Table 5.2

Overall normalised frequencies of quasi-modals over time

Period	*BE able to*	*BE going to*	*BE supposed to*	*BE to*	*(had) better*	*HAVE to*	*(HAVE) got to*	*NEED to*	*WANT to*	Total
0	29,4	11,9	2,1	61,3	3,1	43,8	0	0,5	5,7	157,7
1	29,2	11,7	4,1	67,2	2,3	**73,6	*4,1	3,5	*16,4	**212,1
2	41,1	22	2,1	**39	1,4	83,6	3,5	5,7	*32,6	230,9
3	35,9	25,4	3,9	*19,6	2,6	83,5	1,3	***26,7	46,3	245,2
Overall 0-3	ns	*	ns	****	ns	***	ns	****	****	****

When considering the total frequencies of the modals across the periods in Table 5.1 (the same data used for the modals in Figures 5 and 6), the gradual increase in modals from Period 0 to 1 and 1 to 2, as well as the statistically very significant drop of modals from Period 2 to Period 3 (by 234 normalised instances and with a log

[132]Asterisks are used in accordance with their application in e.g. Leech (2011:562) to "represent statistically significant changes with respect to the previous data point: * = < 0.05; ** = < 0.01; *** = < 0.001. Superscripted asterisks indicate a significant fall in frequency; non-superscripted asterisks a significant rise. ... [U]nasterisked figures are ...non-significant." In this thesis, **** = < 0.0001.

[133] Non-significant.

likelihood score of 19.36[134] [above the critical value of 15.13 where $p < 0.0001$]) can be related to individual modals. The same applies to the overall significant decrease of modals from Period 0 to 3 (by 145,3 normalised instances or 10,1%, [as mentioned above] and with a log likelihood score of 7.67 [above the critical value of 6.63 where $p < 0.01$]). The modals that decrease the most across the entire chronological spectrum (Period 0 to 3) are *ought* (-95,8%) and *shall* (-90,1%), both showing a statistically significant drop above the critical value where $p < 0.0001$. *Might* (-46,8%) and *would* (-23,2%) show the second most significant drop (above the critical value where $p < 0.01$), whereas *need* (-80,6%) and *should* (-25,4%) also undergo a significant decrease above the value where $p < 0.05$. Although *could* (-14,3%) and *may* (-12,8%) also decline in general, their decline is non-significant. The only modal that rises with a statistically significant score of 19.24 (above the critical value where $p < 0.0001$) across the entire time frame is *can* (+54,6%), but *must* (+14,8%) and *will* (+1,8%) also increase, albeit below the level of statistical significance.

What is striking about the quantitative changes from Period 0 to 3 is the statistically significant rise of *can*, which is the only such instance among the modals of SAfE. It is however similar to the situation in BrE, where *can* increases with the same level of significance from 1931 to 2006, but contrasts with the situation in AmE, where this modal shows a slight and non-significant rise from the 1910s to the 2000s (Leech, 2011:551). Furthermore, the relative stability of *will* across all periods is noteworthy when compared to its statistically significant rise in BrE from 1931 to 2006, and the contrasting gradual decline of *will* in AmE across the course of the 20th century and into the 21st century, as reported by Leech (2011:551-4). Lastly, but perhaps most surprisingly, the general increase of *must* in SAfE from Period 0 to 2, as well as its non-significant decrease from Period 2 to 3 is exceptionally remarkable. It is indeed the only modal in its semantic cluster to rise from Period 0 to 3 in SAfE, even though this occurs below the level of statistical significance. Contrastingly, as noted in Chapter 2, *must* is reported to gradually decline across the 20th century, more specifically at the $p < 0.001$ level of significance from 1931 to 2006 in BrE, and by the same level of significance in AmE from the 1910s to the 2000s and into the 2010s (Leech, 2001:551-4). Thus, the permission/possibility/ability/ semantic cluster will be

[134] See Appendix 3 for the log likelihood scores of all the modals and quasi-modals across time.

discussed first to account for e.g. the high frequency of *can*, followed by the prediction/volition cluster which includes the stable modal *will*, and lastly the obligation/necessity cluster will receive attention, with a focus on *must* and its competitors. The obligation/necessity cluster is discussed last, because it will receive special attention in the later semantic analysis, and its synchronic patterns will also be highlighted in the next main section.

4.2.1.2 The permission, possibility and ability cluster

This section will focus on the overall and register-internal frequencies of the semantic cluster expressing meanings related to permission, possibility and ability. Reasons behind frequency changes in this cluster will mainly be explored for the modal pairs *can* and *could*, and *may* and *might* in terms of trade-off relationships (on quantitative and qualitative levels), but references to other pairs will also be made, before turning to the only quasi-modal in this semantic cluster, *BE able to*. Comparisons will also be drawn with the 20th-century trends for this group in BrE and AmE. Figure 8 illustrates the diachronic trends of the permission/possibility/ability cluster, which is drawn from the relevant data in Table 5.1.

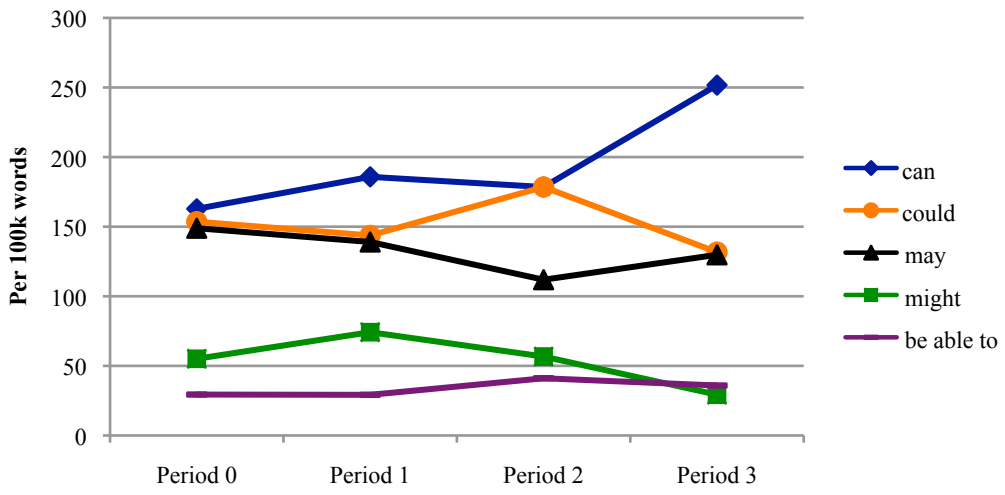

Figure 8

Frequency trends of the permission/possibility/ability cluster across time

From Figure 8 it is apparent that *can, could* and *may* overall enjoy a higher frequency across all periods than *might* and *be able to*, which generally corresponds with the diachronic findings of Mair & Leech (2006:327) and the synchronic findings of Collins (2009a:92). Despite this general quantitative similarity, the frequency trends of the modals and quasi-modals in this semantic cluster are somewhat more complicated.

4.2.1.2.1 *Can* and *could*

This section explores the relationship between *can* and *could*, with a focus on the dynamic use of these modals. The general upward trend of *can* and especially the very significant rise from Period 2 to 3 was already noted above, but the prominence of its high frequency in Period 3 is highlighted against the trends of the other auxiliaries in this semantic cluster in Figure 8. In accordance with its trend towards general incline, the use of this modal increases from Period 0 to 1 (by 23 normalised instances) and from Period 2 to 3 (by 73,2 such instances), despite a slight decrease in frequency from Period 1 to 2 (a decline by only 7,3 normalised instances). Although this high-frequency modal is reported to remain relatively stable in AmE and BrE from 1961 to 1991/2, only slightly increasing in BrE by 2,2% and very slightly decreasing in AmE by 1,5% (Mair & Leech, 2006:327; Leech *et al.*, 2009:74), Leech (2011) reports that it shows a highly significant rise in the earlier BrE data from 1931 to 2006, as also mentioned above – still, this rise only occurs by 2,2% from 1931 to 1991 (calculated from the data of Leech, 2011:511), which is exactly the same percentage of increase for 1961-1991/2 found by Mair & Leech (2006:327). From the 1910s to 1990s (correlating with the start of my Period 2 and the period covering my Period 3) *can* increases by a relatively small 14,8% in AmE according to Leech's (2011:553) data, which is a slightly different picture than the small decline in Mair and Leech (2006:327). Leech's (2011) longer-spanning data as points of comparative reference are therefore closer to the era represented in Period 2 in the SAfE corpus, where *can* is found to significantly increase by 41% in SAfE from the 1910s-1950s to the 1990s (much more than in AmE and BrE), rendering it the second highest ranking modal in

terms of frequency in SAfE for Period 3 – only being outnumbered by *will*.[135] This in itself is unlike the situation in the other varieties, where *would* is the highest-frequency modal by 1991/2, with *will* and *can* next in line (Mair & Leech, 2006:327; Leech *et al*., 2009:72-4). The much higher rise in frequency that occurs for the modal *can* from Period 2 to 3 in SAfE than when compared to the same time frame in the two other native varieties can be linked to its relationship with its original past form *could*.

From the data in Table 5.1 it appears that *can* maintains a trade-off relationship with *could* across every period – with the trend line of *could* almost exactly mirroring that of *can* in Figure 8. *Can* carries the heaviest workload in the permission/possibility/ability cluster in Periods 0, 1 and 3, and *could* only does the same amount of work during Period 2, where the normalised frequencies of *can* and *could* are identical (178,5 instances). As mentioned above, SAfE *could* undergoes a non-significant decline across all periods. However, *could* rises in both BrE (at $p < 0.001$ from 1931 to 2006 in BrE) and AmE (at $p < 0.001$ from the 1910s to 2000s with some small internal fluctuations), according to Leech (2011). This modal however displays a significant decline (at $p > 0.01$) in SAfE from Period 2 to 3 after significantly rising from Period 1 and 2 (by 34,8 normalised instances – only at $p < 0.05$). The fact that *could* does not show an increase from Period 2 to 3 can account for the greater extent to which *can* rises in SAfE between these periods than in other varieties.

Can shows the highest degree of increase (+131,7%) in the news register (from 72,2 to 167,3 instances per 100k words), the second highest rate (+105,3%) in the fiction/narrative register (from 104,7 to 215 normalised instances) and the third highest (+83,2%) in the non-fiction register (from 177,9 to 325,9 normalised instances). Within these periods this modal however decreases by 22,8% in letters, but this is not enough to stop its overall very significant increase, as facilitated by its combined rise in the other three registers. For the same periods *could* shows an even greater decline (-57,1%) in letters (from 177,7 to 76,3 normalised instances) than *can*, which is the highest-ranking decrease for *could* among the registers from Period 2 to 3. Although *could* increases slightly in frequency (+7,7%) in news (from 103,5 to

[135] In their smaller corpus, Rossouw and Van Rooy (2012:16) found that *can* was the third most frequent modal in SAfE, after *will* and *would*. However, given the much larger corpus used in the current study, one can expect the frequencies of the modals to be more stable.

111,5 normalised instances), it is also not enough to counter its general downward trend, which, apart from its large drop in letters, is also motivated its large decline by 48,3% in non-fiction (from 153,7 to 79,5 normalised instances) and slight decline by 5,7% in the fiction/narrative register (from 287,9 to 271,6 normalised instances). The general trend of decline in *could* from Period 2 to 3 therefore results from its decline in letters and non-fiction, whereas the general trend of increase in *can* occurs due to its high rates of increase in firstly news, and also in fiction/narrative and non-fiction. The trade-off relationship between the two modals across these periods can therefore be deduced to be motivated by the register of non-fiction, where the rates of increase and decrease are proportionally closer to each other (-48,3% for *could* and +83,2% for *can*) than for the only other register where *can* increases and *could* decreases (fiction/narrative).

As far as general preference of *can* and *could* in the registers are concerned, it is apparent from the data in Table 21.1 in Appendix 2 that *could* enjoys higher frequencies than *can* in the fiction/narrative register for all the periods, but that *can* is generally preferred in letters across the entire time span, which is also true for news (except in Period 2, where *could* has a slightly higher frequency). The overall preference for *could* in fiction/narrative is due to the fact that narrative writing generally prefers the past tense by convention, as mentioned in § 2.2.2.1 (cf. Banfield, 1982:169-170). This use can be seen in examples (43) and (44), which here represent Periods 0 and 2 respectively.

(43) These expressions, beyond doubt, resulted from sincere and honest feelings; but they **could** not avoid tinging [sic] them with the flattery and adulation usually employed when addressing a Chief or Headman. (1830s; Fiction/narrative)

(44) We saw white face after face, all bared in smiles, and their laughter surrounded the thin screams of the child, until one **could** no longer believe that what one heard was truly a scream of fear. (1950s; Fiction/narrative)

In both these examples *could* has a purely temporal function, distancing the "situation from the reality of the present moment" (Collins, 2009:105).

Non-fiction is less definite in its preference – after *can* and *could* have similar frequencies in Period 0 (158 and 151,1 normalised instances in *can* and *could* respectively), the ratio of preference fluctuates, but remains in favour of *can*, which a agrees with the findings of Biber *et al*. (1999:489) that *can* is used more frequently in academic writing than *could*, and conveys both ability and possibility. The ratios in non-fiction for *can* and *could* respectively range from 1,05:1 in Period 0, to 2,08:1 in Period 1, to 1,16:1 in Period 2 and 4,10:1 in Period 3. Numbers (45), (46) and (47) exemplify the use of *can* in the non-fiction of Period 3, where it was found to convey meanings of ability and possibility – often ambiguously as part of the dynamic group of meanings (cf. Collins, 2009a:101-104) – and often with high density in informative language.

(45) The validity of the warning or notice **can** be proved by registering the letter. It **can** also be delivered personally and its validity **can** be proved with a sworn statement. (1990s; Non-fiction [W2D-001])

(46) [H]ydroxycitronellal is likely to position itself at the interface or michelle's surface, where it **can** significantly effect [sic] emulsion stability, or lower the viscosity. Phenylethanol and vanillin may find their way into the aqueous phase, whereas amylalcohol **can** squeeze into a liquid-crystal structure, as **can** other materials of a polar nature. (1990s; Non-fiction [W2A-023])

(47) Any additional load inertia will increase rotor acceleration time and **can** severely limit the ability of conventional synchronous motors to achieve synchronous operation. (1990s; Non-fiction [W2B-033])

The ratio change from Period 2 to 3 into a higher proportion for *can* confirms the trade-off relationship between these modals mentioned above, which motivates the trends of *can* and *could* for these periods in Figure 8.

The results therefore suggest that from Period 2 to 3 *can* has taken over the functions that *could* had earlier in the non-fiction register. As explained in Chapter 2, *could* and the other original past forms of the modals (*might, should, would*) often express a layer of incompleteness or a hypothetical meaning (the *irrealis*) (see § 2.4.1.2, as well as Bybee [1995:503] and Larreya [2003:21]). This is indeed a common meaning expressed in the non-fiction register, where its argumentative and

informative nature often includes the discussion of future implications, (see § 2.2.2.1). However, as also noted in Chapter 2, the English modals, not unlike those of other Germanic languages, have been on a path of semantic change since earlier times where they have passed along an axis of meanings ranging from knowledge to ability, from ability to permission, from permission to obligation and finally to futurity (see § 2.2.2.1, cf. Conradie, 1987:171 and Traugott, 1972:198-9).

The fact that the rise of *can* correlates with the drop of *could* into Period 3 and that this trade-off is carried by the non-fiction register suggests that *can* increasingly expresses future meanings and is therefore more often used in the hypothetical tenor (linking to the dynamic meaning, which includes theoretical possibility, ability or dynamic implication [cf. Collins, 2009a:101-4]). Theoretical possibility is exemplified by (45), (46) and (47) above, which all express the sense of existential possibility, replacing the regular candidate for these kinds of hypothetical statements, *could*. For *can*, examples (48), (49) and (50) respectively show the other sense of theoretical possibility, namely rational possibility (where "actualization is licensed by more abstract factors pertaining to what is generally considered to be reasonable, legitimate, societally/culturally acceptable", etc. [Collins, 2009a:103]), as well as the ability (linked with the notion of 'potentiality' by Collins [2009a:103]) and dynamic implication meanings ("used in the formulation of an indirect speech act" [Collins, 2009a:104]). All these uses are ultimately classified as dynamic (Collins, 2009a:101-104).

(48) Bread baking is a long and involved story, dating right back to the Stone Age, when grain was crushed between two stones, mixed to a dough with water, shaped into thin cakes and baked on a stone over a fire. As **can** be imagined, these bread cakes were hard and very chewy. (1990s; Non-fiction [W2D-012])

(49) Only rich farmers own tractors. Most farmers **can** only afford to hire them for ploughing. They **cannot** afford to hire them again for harrowing, discing and cultivation. (1990s; Non-fiction [W2D-016])

(50) The other kind describes various projects around southern Africa so that you **can** see how other people organise themselves. (1990s; Non-fiction [W2D-016])

The theoretical possibility meaning (specifically the existential possibility meaning) is by far the most frequent dynamic meaning for *can* in other native varieties of English (Collins, 2009a:101) and the same applies to SAfE for the non-fiction data set of Period 3, although ample instances of other dynamic uses were found, as exemplified directly above. Moreover, dynamic *can* as a whole is much more frequently used than deontic or epistemic *can* in AmE, BrE and AusE (Collins, 2009a:98) and in fact covers a combined 81% of the semantic space or in Leech *et al.*'s terms ecology of this modal across all these varieties. This also corresponds with the finding of Leech *et al.* (2009), as noted in § 2.4.2.2, that *can* has moved toward expressing more root possibility meanings than in the past. As mentioned in § 2.4.1.1., the term 'root' or 'event' modality that includes both deontic and dynamic uses is preferred by e.g. Palmer (2001); in essence, the term 'root possibility' translates as 'dynamic' because the meaning of possibility rules out the deontic meaning associated with permission. Collins (2009a:98) reports that the deontic meaning only covers 9,9% of the semantic space of *can* for the combined varieties, with only 1,1% left for epistemic possibility meanings (8,1% is indeterminate).

The reason why *could* in non-fiction has a higher frequency in Period 2 than in Period 3 is that this component of Period 2 contains larger sections of narration, as non-fiction often does (see § 2.2.2.1; cf. Fontaine and Glavin, 1987:170), which tends toward the use of the past form, as also noted above (cf. Banfield, 1982:169-170). There are fewer narrative sections in the non-fiction register of Period 3. Examples (51) and (52) illustrate the use of *could* in narrative parts of non-fiction in Period 2; (51) is from a non-fiction work of Herman Charles Bosman, in which he remembers his time in prison and expresses a hypothetical meaning within his narration, but *could* in (52), which is simply used to narrate past events (being similar to the temporal uses in [43] and [44]), is from an article in the magazine, *Outlook on a century: South Africa 1870-1970*.

(51) I wanted to go as slow as possible; those moments that we were outside the prison had to be stretched as long as days and hours and years; oh, we hadn't to go at more than a snail's pace. And I saw to it that we didn't. The warder **could** shout his head off. This dreadful ecstasy had to linger. (1940s; Non-fiction)

(52) The vessel heeled and very few boats **could** be launched, and by the time she righted the water was level with the decks. (1910s; Non-fiction)

Furthermore, the non-fiction part of the data for Period 3 contains more scientific language, which, according to e.g. Collins (2009a:102), often contains more possibility meanings. The corpus makeup therefore contributes to inflate the differential between the frequencies of *can* and *could* in Period 3 to some extent.

To conclude on the discussion of *can* and *could*, the increase of *can* into Period 3 can be linked to it taking over some of the hypothetical possibility functions of *could* in the non-fiction register, but this picture may be inflated due to the larger portion of scientific language in Period 3, which implicates a higher occurrence of dynamic uses. Furthermore, the apparent drop of *could* into the same period can be ascribed to the larger portion of narration in Period 2, which somewhat inflates the differential between these modals in Period 3, as seen in Figure 8. Despite the degree of inflation of trends due to register makeup, the fact that *can* has moved into the hypothetical domain of *could* still suggests that it is the more productive choice to express dynamic possibility meanings.

4.2.1.2.2 *May* and *might*

Moving on to a mid-frequency modal in other varieties, *may,* it is evident from Figure 8 that this modal verb maintains a trade-off relationship with *could*. Just as *can* and *could* mirror each others' trends, *may* and *could* do too, but *may* follows a very different path from *can*. The frequencies of *may, can* and *could* in Period 0 are grouped closely together (respectively 148,9, 162,8 and 153,6 instances per 100k words), but when *can* rises from Period 0 to 1, *may* follows the same path as *could* (-6,4%), declining by a very similar rate of 6,6%. Yet, this relationship changes profoundly from Period 1 to 2, where *may* maintains the position of lower-ranking modal in relation to *can* and *could*, with an accelerated decline by 19,5%. On its own, however, a downward trend for *may* is discernable across the board (despite the increase from Period 2 to 3), which corresponds with the general trends of AmE and BrE since the early 20th century to the early 21st century (Leech, 2011:551-4).

When limiting the analysis to Period 2 and 3, however, the frequency of *may* increases by 16% in SAfE, which projects *may* in a totally different direction than in other native varieties, seeing that this modal decreases by 17,4% in BrE from 1961 to 1991/2 and an even larger 34,1% from 1931 to 1991 (calculated from the data of Leech [2011:551]), as well as by 32,4% for 1961 to 1991/2 in AmE (Mair & Leech, 2006:327). This increase resembles the results for *may* in Millar's study of the TIME corpus, since he found this modal to rise from the 1920s to the 2000s (Millar, 2009:199) – a time span corresponding to an extent with the time covered by my Period 2 to slightly beyond my Period 3. SAfE *may* does however not rise to the same extent as in TIME (+59,7%) (2009:199). In Period 3 the frequencies of *may* (129,8 normalised instances) and *could* (131,7 normalised instances) in SAfE are very close, resembling the situation in Periods 0 and 1.

The rise of *may* from Period 2 to 3 is supported by register-internal patterns. Between these periods *may* increases by a large 96,1% in non-fiction (which appears to be a trade-off with *could* in this register) and by 70,4% in news, where in both cases it was found to frequently convey meanings of epistemic possibility, as seen in (53), (54) and (55), despite a decrease by 48,7% in letters, and by 42,2% in fiction/narrative. The above-mentioned findings for *may* in SAfE correspond to the findings of Leech *et al*. (2009:83) and Millar (2009:203), as noted in § 2.4.2.2, that *may* displays a trend towards monosemy in BrE and AmE, in that it most often expresses epistemic possibility meanings, which can sometimes be ambivalent with dynamic possibility (as in [54] and [55]). Despite the ambivalence, these instances are analysed as epistemic on the grounds that the speaker is uncertain as to whether "a situation whose potential for occurrence is not in doubt will be actualized" (Collins, 2009:94), which is often reinforced by epistemic expressions, e.g. *may be possible* in (54) and *may well* in (55).

(53) In drier fynbos[136] areas other rodent species **may** have had a similar effect on the growth pattern and possibly also even on the architecture of vygies[137] and other plants. (1990s; Non-fiction [W2B-024]).

[136] A vegetation type unique to South Africa.

[137] An indigenous plant species of South Africa.

(54) It **may** be possible to produce such losses by taking a spouse and/or children into the business. (1990s; Non-fiction [W2D-004])

(55) ...with a more free-market versus trade-unionist/socialist approach to economic policy, differences **may** well emerge. This is where internal ructions **may** come to the fore but it is still early days... (1990s; News [W2C-002])

Although there are more permission meanings expressed in the letters of Period 2 than those of Period 3, which could contribute to account for the decrease of *may* in letters from these two periods, the permission meaning (which more often occurs in social letters), as exemplified in (56), is not the most prevalent in the letters of Period 2, which suggests that the decrease of deontic meanings of this modal is not as influential on its trends across the 20th century. An example of the more frequent formal, formulaic/polite permission meaning of *may* in Period 2 is given in (57) and (58), which is somewhat ambiguous with expressing a wish correlating with dynamic possibility. No such formulaic uses were found in Period 3, contributing to the apparent decrease in this period.

(56) There **may** be degrees of loyalty, but it must not be blind. (1950s; Social letters)

(57) I hope I **may** hear from you that you are well & "not dismayed". (1940s; Social letters)

(58) I should like to try and re-draft one or two of the chapters in a style and length suitable for the S.A. Outlook, and then submit them to you again, if I **may**. (1940s; Social letters)

However, in Period 2 the most prevalent meaning by far in letters is already that of epistemic possibility, where a high frequency of such meanings occur in a small subset of the letters, particularly offering advice and instruction on various areas of life, as in (59).

(59) Sometimes you **may** make a mistake in your assessment, and find yourself on the giving side of the association. (1950s; Social letters)

In Period 3, *may* often expresses polite wishes to an addressee in social/private letters, as in (60) and (61), where conventional good wishes are expressed for a birthday and for Christmas by means of a kind of dynamic possibility, but epistemic possibility is still the general meaning conveyed by *may* in letters in general, which corresponds to the 20th-century findings of Leech *et al*. (2009:84-5) that *may* "is becoming predominantly an epistemic modal". Only two instances of permission (with some degree of residual ambiguity with epistemic possibility)[138] is found in the letters of Period 3 – which happens to be from a particular business letter addressed to an applicant for a home loan, where the 'we' refers to the bank extending the loan, as seen in (62).

(60) Tania Hope you're [sic] had a very special birthday! **May** God bless you lots in this brand new year. (1990s; Social letters [W1B-005])

(61) Enjoy your Christmas and **may** 1998 bring you all the love and fulfilment you so well deserve. (1990s; Social letters [W1B-005])

(62) Kindly note that the mortgage bond must be registered within two months from the date of this letter, failing which we **may** exercise our rights to withdraw from the loan agreement. ... Notwithstanding the terms and conditions of any mortgage bond/s heretofore registered, it is agreed that we **may**, at any time, increase the rate of interest specified above on all amounts... (1990s; Business letters [W1B-017])

There is at least some evidence of a shift away from permission meanings for *may* in letters over the 20th century, which would support a trade-off with *can* on a qualitative level, but since the decline of such meanings in *may* is not prominent, only a slight degree of trade-off with *can* is suggested on a quantitative level, as will be shown below. The fact that the frequency of *may* decreases in letters from Period 2 to 3 can therefore not be directly attributed to a move away from permission meanings.

[138] Charnock (2009:177) finds that *may* "is generally used to confer a discretionary or enabling power" in English statutes, which links to the legal nature of example (62). In such contexts *may* "is ...seen as conferring a right, rather than as imposing an obligation", but it could also be interpreted as imposing a duty on the authority that holds the power, insofar that it has "a coercive meaning similar to *must*" (2009:177).

The decrease however does not impede the rising trend of this modal toward Period 3, which in itself is found to also correlate with a higher use of epistemic possibility meanings in non-fiction and news.

In written language one would, however, not only expect to often find permission meanings in social letters due to their somewhat conversational quality, but perhaps more so in the dialogue representation of fiction/narrative (see § 2.2.2.1; cf. Banfield, 1982), and seeing that *may* decreases in its frequency in the fiction/narrative register from Period 2 to 3, this register is worth considering in this regard. Nevertheless, the raw tokens of *may* in the fiction/narrative register in the last two periods are very few (9 and 6 in Period 2 and 3 respectively)[139], and out of all of these only 4 instances each are found in dialogue representation in Period 2 and 3, all of which, except one in Period 3, have permission meanings, as in (63). The rest of the instances in fiction/narrative in Period 2 and 3 have epistemic possibility meanings, save one instance of dynamic (theoretical) possibility in Period 2.

(63) "Excuse me, mama,' she broke in, '**may** I please have a word with the mfundisi[140]?' (1990s; Fiction/narrative [W2F-003])

When considering the low frequency of raw *may* in fiction/narrative in the last two periods, making generalisations such as that the permission meaning has increased into Period 3, would be unfounded. Therefore, on this score, there is also little or no evidence for the decline of permission meanings for *may* in the fiction/narrative register[141], which is similar to the situation in letters noted above. This corresponds to the findings of Leech *et al*. (2009:85) for BrE and AmE that the permission meaning of *may*, "losing its traditional bolstering by pedants, pedagogues and parents, has not declined" from 1961 to 1991/2. Yet, when considering SAfE *can* on its own, there is some evidence as to a higher frequency of permission meanings in the fiction/narrative register in the last two periods.

---

[139] See Table 20.1 in Appendix 2.

[140] An isiXhosa word meaning 'master'.

[141] It is of course possible that the permission meaning of *may* might have been declining since before Period 2, but this is unimportant considering the trend of increase in the permission meanings of *can* over the 20th century.

In consideration of the finding mentioned in § 4.2.1.2.1 that the second highest rate of increase for *can* from Period 2 to 3 occurs in the fiction/narrative, the meanings of *can* in this register were analysed in order to test the possibility that the permission meaning of *can* has risen throughout the 20th century, despite little evidence of a reciprocal relationship with *may.* Out of the 32 raw instances of *can* in the fiction/narrative register in Period 2, only 1 is epistemic and the rest, 31, have dynamic meanings, 25 of which are ability meanings (the remaining 6 denoting theoretical possibility). No instances of a deontic (permission) meaning are present here, but only 10 out of the 32 instances are embedded in speech representation. On the other hand, 58 of the 76 instances in the Period 3 data for fiction/narrative contain speech representation. Out of the 76 raw tokens for Period 3, 63 have dynamic meanings (47 ability meanings, and 16 other dynamic meanings [theoretical possibility and dynamic implication]), only 10 have deontic (permission) meanings, as in example (64), despite the high occurrence of speech representation in this register, and the remaining 3 have epistemic meanings. These findings suggest that the permission meaning of *can*, which is not present in the fiction/narrative of Period 2, has been added to the meanings of *can* in Period 3. The frequency of speech representation in this register does however not seem to influence the occurrence of permission meanings, as supposed earlier. The analysis of *can* in the fiction/narrative register over the 20th century therefore provides qualitative support for a trade-off with *may* regarding the permission meaning, but with so few raw instances of *may* to analyse in the fiction/narrative register in Periods 2 and 3, the evidence is not strong on a quantitative level, as mentioned above. The following example is taken from the same text as (63), in order to illustrate the similarity of the permission meanings of *may* and *can* in close proximity to each other in a text ([63] is a request for permission, whereas [64] gives permission).

(64) ‘The priest says you **can** come, mama,’ she said. ‘He is waiting for you.’
(1990s; Fiction/narrative [W2F-003])

Most of the instances of *can* in fiction/narrative therefore have ability meanings, and seeing that this meaning is not available for *may* (Collins, 2009a:101), this once again suggests that dynamic meanings support the frequency trends of *can*

in Figure 8, just like was found to be true for the non-fiction register in § 4.2.1.2.1. The idea of a trend toward monosemy (dynamic meanings) for *can* is indeed strengthened by these above findings, whereas, especially in letters, the trend is also present for *may* (epistemic meanings). These trends agree with the trends noted for these modals in AmE and BrE by Leech *et al.* (2009) (see § 2.4.2.2). Hence, the permission meaning is not a real factor in the frequency trends of *can* and *may*, which agrees with the findings of Collins (2009a:92-8) that this meaning occurs infrequently for both modals.

The rise of *may* from Period 2 to 3 can also be attributed to register-internal trade-offs with its original past tense form *might*. *Might* displays a gradual downward trend when Period 0 is excluded (it increases by 34,7% from Period 0 to 1). Its low frequency in Period 0, at first disrupting this trend, appears to be due to a trade-off with *may*, which generally has much higher frequencies in all of the registers. Despite decreases in the frequencies letters and non-fiction, *might* increases most notably in the news register from Period 0 to 1 (+130,1%), followed by the fiction/narrative register (+61,4%). For the other periods displaying the downward trend, *might* declines by 23,6% between Period 1 and 2 and 48,3% between Period 2 and 3, where it ultimately enjoys the lowest frequency of all the auxiliaries in this semantic cluster. The latter percentage of decline is much higher than for both AmE and BrE, which decrease by 4,5% and 15,1% respectively (Mair & Leech, 2006:327).

The reason behind the more drastic decline in SAfE *might* from Period 2 to 3 is that its frequency declines by 100% in letters[142] (from 48 to 0 normalised instances), by 55.3% in non-fiction and a similar 55% in fiction, as well as by a smaller 6,9% in news. A trade-off situation between *may* and *might* in their non-fiction frequencies in the last two periods is therefore present, which would account for their differing overall trends during this time. The decline of *might* in non-fiction from Period 2 to 3 occurs as an effect of its frequent expression of the future possibility meaning in a retrospective or historical context (in the past tense) during Period 2, as seen in (65), for which such contexts are fewer in the texts of Period 3 (only one such usage for *might* is detected). In Period 3, *might* however almost exclusively expresses future possibility meanings, as seen in (66). However, both

[142] This may point to a degree of colloquialisation, seeing that *might*, a more tentative and hence formalised option (in terms of hedging) in its semantic cluster, declines in letters over the 20th century.

these possibility meanings (past and future temporality) are hypothetical and hence epistemic (cf. Collins, 2009a:110), which corresponds with the finding of Collins that contemporary *might* most frequently conveys epistemic meanings (77,2%) in AmE and BrE.

(65) ...the English minority, conscious always of being a minority, were for long reluctant to embrace a nationalism that **might** cut them off from their tap-root in Britain... (1940s; Non-fiction)

(66) Methods which **might** work well in the ideal settings of well-resourced campuses with small student numbers, **might** not work as well in our environment... (1990s; Non-fiction [W2A-003])

The last auxiliary in need of discussion in this section is *BE able to*, which is the only quasi-modal in the permission/possibility/ability cluster and the overall lowest ranking auxiliary across all periods except in the last, where *might* is 6,6 normalised instances less frequent than this quasi-modal.

4.2.1.2.3 *BE able to*

From Table 5.2 and Figure 8 it is clear that *BE able to* remains stable in its low frequency from Period 0 to 1, but rises by 40,8% from Period 1 to 2, before decreasing somewhat (-12,7%) into Period 3. The effects of these frequency fluctuations on the overall trend line is however not drastic, and this quasi-modal still displays an overall rise in frequency, if non-significant, of 22,1% from Period 0 to 3. The 20th-century trend of decline for *BE able to* does however not correspond with the slight increase in frequency it undergoes in BrE and AmE combined (Leech *et al.*, 2009:97). Support for the degree of decline that *BE able to* undergoes into the 1990s is found in its decrease in all registers (-27,8% in letters, -25,8% in fiction/narrative and -19,1% in news), except in non-fiction (+19,6%).

One reason for the larger frequency of *BE able to* in Period 2 is that this quasi-modal is mostly found in private/social letters for this period (as in [67]), and that it is almost exclusively found in formal/business letters in Period 3 (save one instance). From the assumptions that social letters are regarded as more informal in nature and

that quasi-modals are regarded as more common in informal conversation as opposed to more formal registers (cf. Biber *et al*. [1999:486]), this could account for the higher frequency of *BE able to* in Period 2. Since the ratios for social and business letters in data for Periods 2 and 3 are almost equal (1:0,73 and 1:0,75 respectively), this does not occur due to imbalances in the corpus, but the authors of the private letters in Period 3 appear to have preferred *can* to express dynamic, ability meanings (as in [68]) rather than *BE able to*, which suggests a register-internal trade-off relationship between *can* and *BE able to*, influencing the rise of *can* and the drop of *BE able to* from Period 2 to 3. Thus, these results suggest that *BE able to* is not really a colloquial alternative to *can* in SAfE, but quite the other way around, which problematises the notion that quasi-modals are more often associated with informal, colloquial language than modals (cf. Collins, 2009a:9; § 2.4.2.1).

Another reason for the higher frequency of *BE able to* in Period 2 lies in the different patterns of usage of *BE able to* in the fiction/narrative register in Period 2 and 3. In Period 3, *BE able to* is more productive in hypothetical, predictive expressions projected into the future (as in [69]) than in Period 2, where this quasi-modal is more often used to convey past ability and possibility (as in [70]). This is consistent with the findings of Collins (2009a:199-120) that contemporary *BE able to* mostly expresses theoretical possibility and ability in the other native varieties, which he groups together under the dynamic meaning. This quasi-modal therefore displays a similar semantic trend in SAfE as for AmE and BrE.

(67) ...some day I hope I'll **be able to** come out & see you all again. (1920s; Social letters)

(68) He is also having a driving lesson (!!!) tomorrow, to make sure he **can** pass his driver's test soon. (1990s; Social letters [W1B-003])

(69) "No ways he will **be able to** describe the Chinese. They all look alike to a guy like him." (1990s; Fiction/narrative [W2F-009])

(70) He had done fairly well at his job, and **was able to** hand his wife quite a decent parcel of money. (1930s; Fiction/narrative)

It is however interesting to note that *can* in SAfE is 7 times more frequent than *BE able to* in the 1990s, while *can* is around 10 times more frequent (over 4000 units) than *BE able to* (over 400 units) in the written corpora for both AmE and BrE in 1991/2 (Leech *et al*., 2009:72;97), which suggests that *BE able to* is somewhat more productive in SAfE relative to *can* than in other varieties in the late 20th century.

To conclude the discussion on the permission/possibility/ability cluster, the trends noted in Figure 8 were firstly linked to various trade-off relationships in the register-internal frequencies. The trade-offs that are most clear occur between *can* and *could*, *could* and *may* and *may* and *might*, – especially in the non-fiction register from Period 2 to Period 3, where internal changes in the frequencies of expressing permission/possibility/ability in this register were found. Trade-offs were also noted between *can* and *BE able to*, as well as between *can* and *may*, but only on a qualitative level regarding permission meanings across the last two periods. Instead, trends toward monosemy similar to those noted by Leech *et al*. (2009) for other native varieties of English were highlighted: *can* tends to express more dynamic possibility meanings, whereas *may* more frequently express epistemic possibility.

4.2.1.3 The prediction and volition cluster

The trends of *will, would, shall, WANT to* and *BE going to*, which are illustrated in Figure 9, will be discussed here in terms of overall frequencies and, where necessary, in terms of register-internal frequencies. Some reasons behind noteworthy changes in frequency will be explored, and a quantitative comparison of the patterns of the prediction/volition cluster in BrE and AmE will be undertaken. On a larger scale, the discussion of these modals and quasi-modals will aid in explaining the general trends visible in Figures 5 and 6. *Will* is the first modal that warrants attention in this regard, along with its original past form *would*.

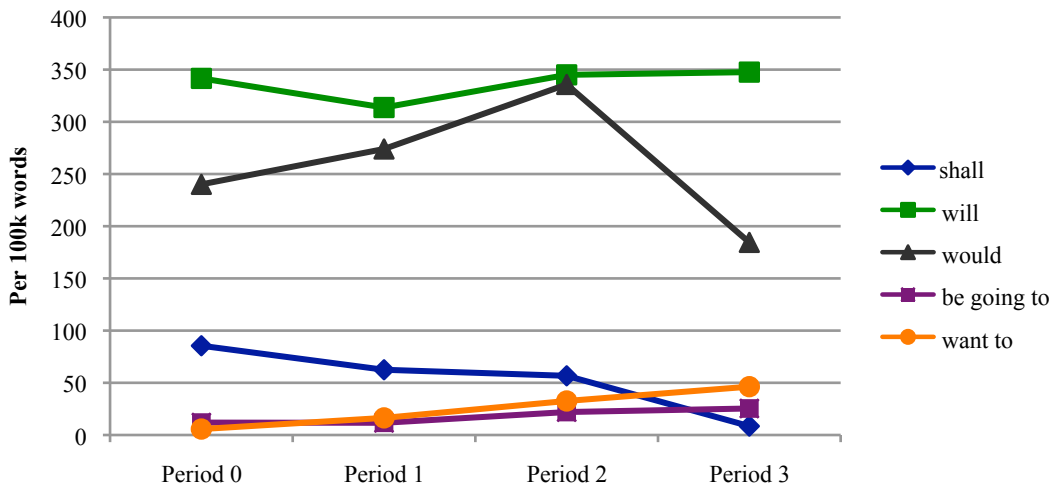

Figure 9

Frequency trends of the prediction/volition cluster across time

4.2.1.3.1 *Will* and *would*

As noted above, *will* is relatively stable across all periods in SAfE when compared to its statistically significant rise (at $p < 0.001$) in BrE from 1931 to 2006 and on the other hand its gradual decline (at $p < 0.001$) in AmE from the 1910s to the 2000s (Leech, 2011:551;4). In Table 5.1 *will* stands out as being especially stable in its frequencies for Periods 0, 2 and 3, varying between 341,6, 345 and 347,6 instances per 100k words respectively, so very minor rises are present here (+1% from Period 0 to 2 and +0,8% from 2 to 3), resulting in an overall increase by just 1,8% from Period 0 to 3. The only small deviation from this pattern occurs from Period 0 to 1, where it decreases slightly and non-significantly by 8,2%. Nevertheless, the above rates of frequency change are small, hence maintaining the relatively stable pattern, especially from Period 2 to 3. The decrease of *will* from Period 0 to 1 is mirrored by an increase in *would* by 14,1% (from 240,1 to 274 normalised instances) across the same periods, which will receive attention in paragraphs to come.

The following consideration of the register-internal frequency patterns of *will* aims to assess whether its stability is supported by these patterns. *Will* has the highest usage frequency in the letters register from the early 19th to mid-20th century (Period 0 to 2), but news contains more instances of this modal in the later 20th century (Period 3). The trends of *will* from Period 0 to 1 and Period 1 to 2 (a decrease followed by an

increase in frequency) are in both cases supported by two registers, but also at the same time slowed down by two other registers. The slight dip in frequency from Period 0 to 1 is supported by the fiction/narrative register, where *will* decreases by 37,8%, and to a lesser extent by the news register, where this modal decreases by 6,2%. There are however increases in frequency within letters (+6,6%) and especially non-fiction (+113,4%), but despite their slight offset, these increases do not succeed to counteract the overall pattern of decrease from Period 0 to 1. On the other hand, the increase of *will* from Period 1 to 2 is supported by its frequency increase within the fiction/narrative register (+87,5%) and also letters (+7,6%), but the two other registers show a decrease in frequency (news [-6,4%] and non-fiction [-45,7%]). These fluctuating register-internal patterns from Period 0 to 2 seem to be somewhat counteractive, i.e. they cancel each other out, and consequently create some stability in the frequency of *will* over these periods. This kind of alternating frequency pattern for the registers is also present form Period 2 to 3 – a pattern which ultimately contributes to the larger degree of stability of *will* across these two last periods. Into the last period, *will* decreases in the letters register by 30,9% and slightly increases in fiction/narrative by 4,4%, but increases more in the news register by 48,1% and by a greater 77,1% in non-fiction.

The data suggest that the alternating directions of the trend lines of *will* from respectively Period 0 to 1 and Period 1 to 2 are largely influenced by this modal's frequencies in the fiction/narrative register. It was found that many uses of dynamic *will* with a future volition/intention reading (with some degree of ambiguity with prediction) (cf. Collins, 2009a:131-4) are present in the fiction/narrative writing of Period 0, where narrators describing their explorative expeditions frequently express their intentions for their journey's next move (as in [71]). Also, the narrations of this period are often written in direct address to a reader, either a superior that the narrator has to report to, a family member or a generic reader public (72), and it also frequently happens that the narrator reports past speech acts containing *will* (73) – all of these cases have dynamic readings. Some formulaic uses in the dynamic category were also found in the speech representation of Period 0 (74).

(71) ...we **will** start as soon as I hear that the Great Fish River is likely to be practicable for the wagons to pass. (1830s; Fiction/narrative)

(72) It **will** not surprise the reader that a happy connexion with the settlers of Salem, which had lasted for nearly four years, ... could not be dissevered without mutual pain and regret. (1830s; Fiction/narrative)

(73) ... now who is the man that **will** velantear [sic] to go? T Keen steped [sic] forverd [sic] and said: I **will** go[143]. (1820s; Fiction/narrative)

(74) I responded, 'if you **will** bring them up to my father-in-law's.' He did so, and the bargain was struck. (1860s; Fiction/narrative)

Furthermore, epistemic readings of future prediction are as present, albeit slightly less so than dynamic readings, and are usually related to the narrator's predictions about the places and people they encounter (75) or their hopes for a situation (76) (cf. Collins, 2009a:126-131).

(75) And now we have traversed what is undoubtedly the largest gold-fields on the face of His earth... That it **will** soon lie opened up to European enterprise is a certainty. (1860s; Fiction/narrative)

(76) Certainly the present is not the time for penetrating that country; but I hope the present turbulent spirit of that people **will** soon begin to subside... (1830s; Fiction/narrative)

To some degree, the nature of the textual subgenres in the fiction/narrative of Period 0 (that tends toward dynamic meanings) contributes to the high frequency of *will* in this period. The fact that more texts containing fictional writing rather than narratives of journeys are present in the data of Period 1 contributes to the decrease of *will* into Period 1, where this modal more often has an epistemic reading, respectively the central epistemic sense (objective predictability) in (77) and the temporal sense of futurity (subjective prediction) in (78) (cf. Collins, 2009a:126). Both these examples are from speech representation, but *will* does not occur in speech representation as frequently in this period as in the previous one, as a result of an overall lower

[143] *The Chronicle of Jeremiah Goldswain*, from which this example was taken, is, among other things, noted for its eccentric spelling, as is visible here.

frequency of speech representation in the texts of this period, which also accounts for the slight drop in frequency.

(77) ...as they sat twisting wicks for the candles together, she said to him, 'You **will** be eighteen on your next birthday, my son... (1900s; Fiction/narrative)

(78) Just forty years ago these Namaquas[144] first saw a plough... Their astonishment on seeing it set to work was great. The old chief stood upon a hill looking on in silence for a long time. At length he said to his councillors, "Look at this strange thing which Mynheer[145] has brought; how it tears up the ground with its iron mouth! If it goes on so all the day, it **will** do more work than ten wives!" (1870s; Fiction/narrative)

In Period 2, more instances of the dynamic use resembling those in Period 0 are present than in Period 1, mostly due to a higher frequency of speech representation in fiction proper here, but epistemic uses are still also present to a large extent, contributing to the slight increase in frequency of *will* in this period. Period 3, on the other hand only contains fiction proper in its texts for this register, and the frequency of speech representation in relation to narration resembles that of Period 2 (and 0), contributing to its stability in this last period. The dynamic meaning is still frequent, but the epistemic meaning ([79] and [80]) is more popular in the fiction/narrative register in Period 3 than in previous periods. This corresponds with the report of Collins (2009a:126) that the epistemic meaning is the most frequent among the uses of *will* across contemporary AmE, BrE and AusE, accounting for 59,2% of uses in these varieties combined.

(79) But more likely they **will** fall into some trap. (1990s; Fiction/narrative [W2F-005])

(80) That **will** shut her big mouth and give my hot ears a rest. (1990s; Fiction/narrative [W2F-002])

---

[144] An indigenous ethnic group of the Northern Cape, Namibia and Botswana of Khoisan descent.

[145] The spelling is a representation of the Dutch form of address *mijn heer* (later Afrikaans *meneer*) meaning 'my lord' or 'sir'.

The internal proportions of the subgenres of fiction and narrative texts and their frequencies of speech representation, together with the accompanying occurrences of dynamic and epistemic meanings, are therefore found to be responsible for the small fluctuations in the frequency of *will* from Period 0 to 2, but because of the sole presence of fiction proper in Period 3, these small differences are balanced out and manifests in a more stable trend line for this modal toward the 1990s. Furthermore, because of the counteractive force obtained by the simultaneous frequency rise and drop of *will* from Period 2 to 3 in two registers in each case, the balance between these two periods is reinforced. Although these quantitative fluctuations contribute to the stable trend line of *will* over the 20th century, there is no clear trend for one register or the other on the longer term, so, on a qualitative level, the continuity in the most frequent senses of *will* from Period 2 to 3, i.e. dynamic and epistemic meanings, can be said to support its stable trend to a greater extent.

The overall stable pattern of *will* is not present in the case of *would*; after rising slightly by 14,1% from Period 0 to 1 (as noted above), *would* goes on to rise significantly from 1 to 2 (with a log likelihood score of 6.31 – above the value where $p < 0.05$) with an accelerated increase of 22,6%, whilst its trend of increase undergoes a dramatic change of direction when it decreases into Period 3. The frequencies of *will* and *would* are almost in agreement in Period 2, with a difference of only 9,2 units per 100k words in the favour of *will*. The abrupt change in direction for *would* in Period 3 however steers the trend lines of these two modals in very different directions. In this last period the decrease of *would* by 45,1% is statistically very significant with a log likelihood score of 44.40 (above the value where $p < 0.0001$), which is the highest log likelihood score for all modals across chronologically consecutive periods (*shall* has a higher score for its decrease from Period 0 to 3[146]). This makes *would* the modal that decreases the most from Period 2 to 3 (from 335,8 to 184,5 normalised instances), even more so than *shall*, although these modals share the same statistical level of significant decrease between these last two periods – making them the most prominent carriers of the general downward trend of modals from Period 2 to 3 as seen in Figure 6. The overall decrease of *would* from the first to the last period is also significant, with the log likelihood score of 7.14 (above the critical value where $p < 0.01$).

[146] See Table 26.1 in Appendix 3.

SAfE *would* at first displays a similar trend to other native varieties, with its rising frequencies from Period 0 to 2. Leech's (2011:551) results for BrE show that *would* fist increases in frequency from 1901 to 1961 (roughly corresponding with my Period 1 to Period 2), but then it decreases again to 1991 and 2006 (comparable with Period 3), which is very similar to its trends in SAfE. However, in both these later periods (1990s and 2006) in BrE, *would* still has a somewhat higher frequency than in 1931 (i.e. a non-significant increase), which is again different from the situation in SAfE. A more comparable picture is indeed present for AmE, for which Leech (2011:553) found that *would* increases in frequency from the 1910s and 1920s to the 1930s up until the 1950s (during my Period 2), where after a gradual decline, albeit non-significant, occurs over the decades leading to the 1990s and 2000s (corresponding with Period 3) – once again very similar to the trends of SAfE *would*. Mair and Leech (2006:327) also found this modal to decrease in these varieties in their earlier study. The high-frequency modal *would* declines by 11% in BrE and by 6,1% in AmE from 1961 to 1991/2, and despite this modal maintaining the highest frequency of all modals in AmE, *will* has a higher frequency in BrE by 1991. Ergo, by Period 3, the relationship between *will* and *would* in SAfE is more comparable with that of BrE by 1991/2, but the fact that SAfE *would* decreases by a much higher 45,1% into Period 3 (as noted above), together with the stability of *will* into this last period, suggests that *will* is maintaining its position in the prediction/volition cluster and *would* is losing ground to *will* and other auxiliaries.

Naturally all four registers show a decline in their frequency of *would* from Period 2 to 3, but letters and non-fiction support this trend to the greatest extent. Between these two periods *would* declines by 48,8% in letters (from 420,3 to 215,1 normalised instances) and by a larger 65,9% in non-fiction (from 256,1 to 87,3 such instances). Its decrease in letters can be linked to a decline in hypothetical uses conveying excessive degrees of politeness and tentativeness, and consequently a lesser degree of formality, which links with the overall concept of colloquialisation. The use of *would* to express polite suggestions or indirect speech acts of requesting, both classed under the tentative hypothetical meaning by Collins (2009a:141-2), are indeed very prevalent in the social and business letters of Period 2. A very large amount of these polite uses are found in the form of suggestions and advice in the private letters from father to son in the 1950s (as in [81] and [82]), which is also mentioned in earlier examples. In (81) *suggest* is used performatively, with a degree

of diffidence added by tentative *would* (cf. Collins, 2009a:142). Indirect speech acts of requesting are less present, but share equal frequencies in social and business letters. These uses often occur with extra hedging in the form of collocates such as *kindly* and *oblige* (83), and by formulating a request in the declarative mood it is also often made more syntactically indirect, with the diffidence and politeness being reinforced by e.g. *perhaps* and *kindly* (84) (cf. Collins, 2009a:142).

(81) Everyone should have a philosophy for life. John, I **would** suggest that yours be summed up in one word: Simplicity. (1950s; Social letters)

(82) But there is no virtue in letting bad types take advantage of you if you can avoid it. For that reason, I **would** like to indicate how to recognise good from bad. (1950s; Social letters)

(83) Enclosed also please find communications... I have mislaid the address of the Secretary. **Would** you kindly oblige me by directing them to him. (1910s; Business letters)

(84) Perhaps you **would** kindly let me know at an early date if this is acceptable to you... (1950s; Social letters)

In Period 2 *will* is also used for indirect speech acts of requesting (similar to the use of *would* in [83] and [84]). This use does not only occur in social letters, where conventionalised hedging (connected with past modal forms) is less needed to keep one's face (see § 2.2.1.2.2 and § 2.2.2.1), yet is still applied to some extent with collocates such as *kindly* (85), but also in business letters, where a slightly higher level of hedging is found with verbs like *grant* (86). These uses are classified as dynamic by Collins (2009a:132) and adhere to the notion of willingness. In Period 3 however, only one such request hedged by *please* is found in social letters (87). No such requests using *will* are however present in the business letters of Period 3, and only one request using *would* is (88), but none are present in the social letters of the last period.

(85) **Will** you kindly give me some more particulars or else get him to send in an application... (1930s; Social letters)

(86) **Will** you kindly grant us permission to print these songs... (1910s; Business letters)

(87) By the way, **will** you please post all my Xmas letters for me when G & R hand them to you... (1990s; Social letters [W1B-004])

(88) **Would** you kindly complete it and return it to us... (1990s; Business letters [W1B-018])

To summarise at this point, formulaic requests using *will* and *would* in Period 2 occur in both social and business letters to similar extents[147], but in Period 3 *will* is not present in the requests of business letters and only has one such use in social letters, whereas *would* is not present in the requests of social letters and only has one such use in business letters. To the extent at least where such low frequencies can account for trends, the data for *would* suggest that the changing frequencies of formulaic requests are not the main reason behind its drop, seeing that both *will* and *would* display similar degrees in the loss of this use. The set of 1950s private letters from father to son contains the most instances of *would* by far when compared to the private and business letters of Period 2 – nearly all of them adhering to the hypothetical use of e.g. polite suggestions, as exemplified in (81) and (82), but also other hypothetical uses with mostly epistemic meanings in future predictions regarding hypothetical situations (as in [89]), or as conditional constructions (as seen in [90]).

(89) I doubt, however, whether any South African **would** be expected to leave these shores again, except as a volunteer. I cannot see South Africa ever starting a war, but there is always the possibility of our having one forced on us. It **would** be difficult not to resent the enemy. (1950s; Social letters)

(90) If there were no doubt, however, that an organization or even my family were behaving in an uncompromising, evil fashion which my conscience could not accept, I **would** resign from associating with them. (1950s; Social letters)

[147] In Period 2 formulaic requests with *would* occur in 5 instances each in social and business letters, and with *will* they occur in 5 instances in social letters and 3 in business letters.

Hence, the frequency of hypothetical *would* in these particular letters, usually carrying polite suggestions and other epistemic meanings, accounts for the decrease of this modal in the letters register from Period 2 to 3, seeing that the letters of Period 3 do not offer advice and make suggestions to even remotely the same extent.

Considering the fact that *will* is more frequently used in letters than *would* across all periods, the proportional use of these modals changes to some extent in the last two periods: the ratios for *will* and *would* respectively for letters in Period 2 is 1,46:1, but this ratio changes to 1,97:1 in Period 3. This suggests that the decrease of hypothetical *would* in letters has enlarged its proportional difference to *will* since Period 2, where the latter is now almost twice as frequent in the letters of Period 3. This is therefore not so much of a trade-off as it is a symptom of internal genre makeup.

Despite the decline of such hypothetical meaning in letters influencing the decline of *would* from Period 2 to 3, it is non-fiction that indeed contains an extent of a quantitative trade-off between *will* and *would* between these two periods (as seen in Table 21.1 in Appendix 2). This is however not so much supported on a qualitative (semantic) level as on the level of register-internal makeup. Non-fiction is the register that shows the biggest increase for *will* from Period 2 to 3 (from 175,2 to 310,3 normalised instances), which cross-matches the decrease of *would* in the same register between the same periods. The ratio of *will* and *would* in non-fiction changes respectively from 1:1,46 in the favour of *would* in Period 2, to 3,55:1 in the favour of *will* in Period 3. As it happens, the non-fiction of Period 2 contains many texts adopting either a historical or political view. The textual functions of narration and argumentation can respectively be identified here. In the pieces on history, the tone is often narrative and the approach retrospective, resulting in some purely temporal uses of *would* (in the case of [91] the backshifted epistemic use), with similar meanings to its present time counterpart *will.* For the pieces on politics, the argumentative tone motivates the very frequent occurrence of the hypothetical (epistemic) meaning of *would* (as in [92] and [93]) (cf. Collins, 2009a:139-142), which in turn supports its high frequency in the 1910s-1950s. This finding corresponds with the findings of Collins (2009a:140) that the epistemic meaning is the most frequent in BrE and AmE. Contrary to this, the texts in the non-fiction of Period 3 contains much fewer such hypothetical uses and tend toward a more informative style, contributing to the large drop of *would* into this last period.

(91) They claimed that their own edifice **would** be solid and homogenous and strong... (1940s; Non-fiction)

(92) The Nationalists could quote history to prove that even samewerking[148] was so risky as to be highly undesirable, while samesmelting[149] **would** be a pointless national suicide, the destruction of Hertzog's great work at Hertzog's own hands. (1940s; Non-fiction)

(93) Yet I suppose if Christ came on earth to-day His so-called followers **would** be ashamed to be seen in His company. (1920s; Non-fiction)

Because the non-fiction texts of Period 3 contain much less such historical material (resulting in the temporal use of *would*) and are generally more informative than argumentative (i.e. fewer hypothetical meanings), *will* is preferred instead of *would*. These frequency changes do therefore not necessarily provide firm evidence of language change, but rather show that particular uses are more likely to be found in particular text types/genres, e.g. informative language. In other words, as the genre changes, the frequencies of these particular uses are affected.

For the same formulaic use of *would* as a means of making a polite request via an indirect speech act (as in [83] and [84]), Bowerman (2004b:477) maintains that *won't* is often used as a directive softener in SAfE, as noted in § 1.1.2. As far as his assertion is concerned, no instances of such usage were found in the historical corpus, not even in the social letters or in speech representation in fiction or other narratives, where friendly requests are more likely to arise than in other written registers. The raw tokens of *won't* in the corpus were however few, so it should not be taken as conclusive evidence that this use does not exist (respectively 0, 12, 16 and 12 raw tokens in Period 0, 1, 2 and 3). The tokens found in the historical corpus rather expressed the conventional meaning of *will not* – that of negated prediction (94), and in rarer cases, volition (95).

[148] An Afrikaans term (noun) meaning 'collaboration', or literally 'working together'.

[149] An Afrikaans term (noun) meaning 'merging', or literally 'melting together'.

(94) "I guess it **won't** take more'n a foot measure to get to the bottom of your feelings, tho' you are long enough to be a telegraph pole," snorted Uncle Abe. (1890s; Fiction/narrative)

(95) Dear Sarie, I **won't** try to write more now – But I hope that I may have a line from you... (1950s; Social letters)

Hence, for a more balanced report on *won't*, this issue will receive more attention in the synchronic analysis for contemporary SAfE where spoken SAfE is analysed and many more instances of *won't* were found (§ 4.2.2.1).

To conclude this section, the stable trend of *will* was found to be supported by small fluctuations in the frequency of *will* in the fiction/narrative register due to register-internal makeup up until Period 2, but the balancing out of these fluctuations stabilises *will* in Period 3 - the stability being reinforced by the simultaneous frequency increases and decreases of *will* in two registers in each case for both Period 2 and 3. A frequency trade-off between *will* and *would* between Period 2 and 3 in the non-fiction register is present, where *would* in Period 2 is often used in texts adopting a historical perspective or argumentative tone view and *will* in Period 3 is more often used in texts containing less historical material and more informative language, strongly influencing the decrease of *would* over the last two periods. On the other hand, this decrease, which is statistically very significant, was linked to the registers of non-fiction and letters. In letters, a decline in the hypothetical meaning between Period 2 and 3 was found to be present, especially in the form of polite suggestions and advice based on future predictions of hypothetical situations in the letters written by a father to his son in the 1950s, further contributing to the frequency drop of *would* in Period 3, which is the drop that carries the downward trend of modals from Period 2 to 3 to the greatest extent. Finally, no evidence for the use of *won't* as a directive softener was found in the diachronic corpus.

4.2.1.3.2 *Shall*

Although *shall* shares the same level of highly significant decrease as *would* from Periods 2 to 3 (above the critical value where $p < 0.0001$), *shall* decreases by a greater level of significance than *would* (also above the value where $p < 0.0001$) from Period

0 to 3 (-90,1% [from 85,5 to 8,5 normalised instances]). In fact, this is the most significant decrease of any modal from Period 0 to 3, as mentioned above, having the highest[150] log likelihood score (72.16) present for all the modal and quasi-modal data. The decline of *shall* is also fairly constant across all periods (see Table 5.1 and Figure 9), in contrast with the drastic change in direction of *would* across the last two periods. It decreases from 85,5 to 62,5 instances per 100k words from Period 0 to 1 (-26,9%), but shows a decelerated decrease into Period 2 to 56,7 such instances (-9,3%). A greatly accelerated decrease however occurs again into Period 3 to 8,5 instances (-85%), which is highly significant (above the value where $p < 0.0001$, with a log likelihood score of 38.92) as mentioned above, contributing to the general decline of modals from Period 2 to 3 as seen in Figure 6. A decline over the course of the 20th century is also the pattern reported for *shall* in BrE and AmE (Mair & Leech, 2006; Leech, 2011).

Mair and Leech's (2006:327) data show that this modal declines by 43,7% in BrE and by 43,8% in AmE from 1961 to 1991/2, which is a slower rate of decrease than in SAfE from Period 2 to 3[151]. For a somewhat more fair comparison, Leech's (2011:551) data for BrE, reaching further back, shows a very steady overall decline for *shall* from 1901 to 1931 (toward the end of my Period 1 and the middle of my Period 2) and again from 1961 to 1991 (and 2006), with a steady level of significant decrease (where $p < 0.001$) for each consecutive period from 1931 to 2006. In AmE he also found this modal to decline steadily and significantly from the 1910s to the 1990s (and 2000s) (corresponding with my Periods 2 to 3) (above the critical value where where $p < 0.001$). The levels of statistical significance reported in these studies therefore indicate a slightly less significant decline for *shall* in the other varieties than in SAfE from Period 2 to 3, which corresponds with the above inference made from Mair and Leech's (2006) data that the other varieties enjoy a slower rate of decrease in terms of percentages[152]. SAfE *shall* hence declines to a somewhat larger extent than BrE and AmE during the course of the 20th century.

---

[150] Only *ought* shares the same statistical level of decrease from Period 0 to 3, but *ought* has a lower log likelihood score (17.23).

[151] BrE however has a slightly higher log likelihood score (44.3) than SAfE for the decline of *shall* across the 20th century, but in AmE the log likelihood score (33.1) is lower than in SAfE.

[152] The slower rate of decrease for *shall* in BrE and AmE than for SAfE may point toward the logical endpoint to the trajectory of decline in these varieties, where is tends to slow down as it approaches zero.

Collins (2009a:135) and Leech's (2011:551-3) data agree that contemporary *shall* is about twice as frequent in BrE than in AmE. In the ICE-corpora, which are more comparable with the data used in my Period 3, Collins' (2009a:135) pmw frequencies calculate to 22,3 instances per 100k words in BrE and 10,2 in AmE. This agrees with the above finding that SAfE *shall* declines more than in these varieties, with a frequency of only 8,5 per 100k words – slightly lower than in AmE, but almost three times less frequent than in BrE. As noted in § 2.4.2.2, the decline of *shall* in the large native varieties was found to be related to the trend toward "paradigmatic atrophy" by Leech *et al*. (2009:80) (where "the number of the various grammatical contrasts which constitute the paradigm of an English verb is reduced, so that only some of the possibilities are available, or at least are likely to occur"), where in contemporary English this modal is virtually restricted to first-person subjects or to archaic uses, i.e. it displays "distributional fragmentation" (mostly in quotations or administrative language) (Leech *et al*. (2009:81), and as a reinforcing factor, the rarity of the negative construction *shan't* contributes to the scarcity of *shall*. These trends correlate with the overall idea that modals are becoming more monosemous over the 20th century. The possible reasons behind the decline of *shall* in SAfE will be investigated in the next few paragraphs.

To begin with, no instances of *shan't* were found in the entirety of the diachronic corpus, not even in the older data, which is interesting because Collins (2009a:135) still found a few tokens in the contemporary data, mostly in BrE. Thus, the absence of *shan't* in SAfE can be deduced to reinforce the overall low frequency of *shall* when compared with those of *will* and *would* in the prediction/volition cluster, because it is less inclined to negation than its rivals – apparently even more so than in other varieties. Moreover, the restriction of *shall* to first-person subjects is certainly present in SAfE. Out of the 166 raw tokens of *shall* in Period 0, 112 collocate with first-person subjects *I* and *we* (67,5%), 4 with second- (*you* and *ye*) and 50 with third-person subjects (e.g. *he, she* and *it*). In Period 1, 84 of the 107 raw tokens are used with first-person subjects (78,5%), only 1 with second- and 22 with third-person subjects. Among the 80 raw tokens of this modal in Period 2, 67 collocate with first-person subjects (83,8%), whereas only 3 with second- and 10 with third-person subjects are present. By Period 3, 10 of the 13 raw tokens have first-person subjects (76,9%) and the remaining 3 all have third-person subjects. The small number of raw tokens in Period 3 skews the implication of the latter percentage to

some extent, but as for the other three periods, the increase of first-person collocates with *shall* rises by 16,3% from Period 0 to 2. This confirms the trend for SAfE *shall* to be virtually restricted to first-person subjects, just as in BrE and AmE, and SAfE even has a higher proportion of these than the other varieties (compare 76,9% in SAfE by Period 3 with 68,2% in Collins' [2009a:137] combined data for all varieties), which accounts at least in part for its decline over these periods. Register-internal patterns may shed more light on this.

From Period 2 to 3 *shall* in SAfE declines strongly in all four registers. In terms of percentages, the greatest decline occurs in news, where it is totally eliminated in Period 3 (by 100% from 18,8 to 0 instances per 100k words), followed by the declines in fiction/narrative (-89,1%), non-fiction (-67,2%) and letters (-60,3%). Perspective on these rates of decrease is gained when considering the raw frequencies of this modal in Period 3, which are indeed very low (see Table 20.1 in Appendix 2), with only 7 tokens in letters, 0 in news, 2 in fiction/narrative and 4 in non-fiction. The demise of *shall* is therefore obvious. A link between this and a trade-off relationship with firstly *will* and secondly *BE going to* can be drawn.

It was mentioned above that from Period 2 to 3 *shall* declines the most in the news register. As it happens, all 6 raw tokens of *shall* in the news of Period 2 are from the 1930s, 5 of which are from the *Diamond Fields Advertiser* and the remaining 1 being from a quote from Olive Schreiner in a piece about her burial in *The Outspan* (96). The piece mentions that this quote was spoken in 1894 (Period 1), which itself is therefore an archaic use. Three further instances of *shall* in direct quotations are present among the other 6, respectively being from the Bible (another archaic use), a quote attributed to the King of Spain (both of the former do not count as SAfE) and also a quote from the formal speech of the South African Minister of Posts and Telegraphs. The two remaining instances are used in a piece expounding on future legal measures, which imply formal usage in administrative, as exemplified in (97).

(96) Before they went down she said: "...I **shall** buy one morgen[153] of this top and we must be buried here." (1930s; News)

[153] A Dutch loanword denoting a former measure of land (later Afrikaans *morg*).

(97) Mr Mowbray thought they might propose an amendment to this effect: In no case **shall** a call on an employer exceed the amount an insurance company would accept the risk at... (1930s; News)

Similar restrictions of usage reported for *shall* in other varieties are therefore already present in SAfE in the news of Period 2, and seeing that this modal disappears altogether in the news of Period 3, these restrictions certainly have an adverse effect on its frequency in this register. Once again, to the extent that the disappearance of *shall* in Period 3 can be taken as evidence, a trade-off relationship with *will* in news can be discerned across the last two periods: the ratio of *shall* relative to *will* is 1:19 in Period 2, and of course the ratio altogether disappears with the 0 instances of *shall* in Period 3 against the 529,7 normalised instances of *will*.

The 6 raw tokens of *shall* in the news of Period 2 mentioned above divide equally between deontic, dynamic and epistemic meanings (here used with the conditional in [98]), and in the 209 raw tokens of *will* in the same register for Period 3, the epistemic meaning is the most frequent (99). To some extent at least this suggests that the shift of the epistemic meaning to *will* may be a reason behind the decrease of *shall* in this register, but the evidence is not conclusive. A somewhat more striking trade-off with quasi-modal *BE going to* is also apparent in news across the 20th century, where no instances of this quasi-modal are present in the news of Period 2 (just as there are for *shall* in Period 3), but 32,9 normalised instances of *BE going to* are present in the news of Period 3. A shift in workload from *shall* to *BE going to* to express volition/prediction meanings can hence be supposed, with *BE going to* almost being twice as productive in the news of Period 3 as *shall* was in Period 2. Out of the thirteen raw tokens of *BE going to* in the news of Period 3, eleven are epistemic (100) and the remaining two are dynamic. This strengthens the evidence for the shift of the epistemic meaning to other auxiliaries (*will* and *BE going to*) being responsible for *shall*'s decrease in news.

(98) ...as long as [they are] insured in a proper company and they hand in their receipts they **shall** be free. (1930s; News)

(99) A domestic worker who is dismissed unfairly **will** be able to get her job back and/or obtain compensation from her employer. (1990s; News [W2C-020])

(100) Different numbers of senior posts **are going to** be assigned to schools in accordance with their gradings. (1990s; News [W2C-001])

As further noted above, the register with the smallest loss of *shall* from Period 2 to 3 is letters, and indeed a lesser degree of trade-off with *will* can be detected in this register. The proportional frequencies of *shall* and *will* respectively change from 1:5 in Period 2 to 1:8,7 in the letters of Period 3, which does show some weight moving towards *will*, albeit less so than in fiction/narrative, where the ratio changes from 1:3,3 to 1:31,8, and even less so than in non-fiction, where it changes from 1:9,3 to 1:50. A bigger trade-off with *BE going to* is present in letters, with the ratio of *shall* relative to *BE going to* being 10,2:1 in Period 2 changing to 1,2:1 in Period 3, which indicates near-equal proportions in Period 3, against the ten times more frequent *shall* of Period 2. It was found that most of the 51 raw tokens of *shall* in the letters of Period 2 express volition as part of the dynamic meaning (101), whereas, despite this meaning still being present in the seven raw tokens in the letters of Period 3, four still expresses dynamic and the remaining three deontic meanings. Out of the 5 raw tokens of *BE going to* in the letters of Period 2, only two have dynamic (volition) meanings and the other three express epistemic meanings; this changes in Period 3 where this quasi-modal has 6 raw tokens in letters (only one less than *shall* in this period), where 5 of these have volition meanings (102) and only one epistemic meaning persists. To the extent that the changes within such a small amount of tokens can indicate trends, it would seem as if the shift of the dynamic meaning of volition from *shall* to *BE going to* is one of the reasons for the decline of *shall* in the letters of Period 3.

(101) I was not able to talk to you at length so I **shall** tell you now what I would have like to have said before leaving. (1950s; Social letters)

(102) But I **am going to** sign off now as it is past bed time and I want to do some clinic work before lights out... (1990s; Social letters [W1B-004])

The register besides news showing the most evidence of a trade-off between *shall* and *will* is shown above to be non-fiction. In Period 3 *shall* only has 4 raw tokens in non-fiction, adding to the proportional difference, but the fact that this modal decreases in this register by 67,2% from Period 2 to 3 and *will* increases here

by 77,1% suggests a trade-off nonetheless. In Period 2 the meanings of the 7 tokens of *shall* in non-fiction are divided as follows: 1 instance found in a quote expresses a dynamic (volitional) meaning and the remaining 6 all express epistemic meanings, as in (103) (cf. Collins, 2009a:136).

(103) In each of the four Provinces of the Union the laws referred to emphasise the necessity of three things: – Clothes, houses, and work-factors for civilisation which, as we **shall** see, were in America among the first things which the negroes – without much alternative is true-adopted to their ultimate benefit. (1910s; Non-fiction)

On the other hand, of the 4 tokens of *shall* in the non-fiction of Period 3, three conveys the dynamic meaning expressing volition or intentionality (cf. Collins, 2009a:137), twice as part of a writing formula in academic writing, as in (104), and once in a threat found in a legal document, as in (105). The remaining token conveys the deontic meaning (the "constitutive/regulative use" [Collins, 2009a:137; cf. Huddleston & Pullum, 2002:194]) in another legal document (106). No epistemic readings were found.

(104) I **shall** now focus on the most current thinking in English on theories of teaching... (1990s; Non-fiction [W2A-003])

(105) I **shall** institute an action against you in the small claims court for the recovery of the purchase price... (1990s; Non-fiction [W2D-001])

(106) ...the law provides a table of precedence as to which of your relatives **shall** inherit. (1990s; Non-fiction [W2D-004])

For the other native varieties (BrE, AmE and AusE) combined, Collins (2009a:135) reports that the deontic meaning for *shall* has the highest contemporary frequency of occurrence, and to the extent that very small raw frequencies can be indicative of trends, the deontic meaning can be regarded as a point of difference between these varieties and the non-fiction of SAfE, seeing that this meaning only occurs once in Period 3 and not at all in Period 2. Collins further reports that the

dynamic meaning enjoys near-equal prevalence to deontic meanings in the other native varieties, and epistemic meanings have the lowest frequency. The fact that the epistemic meaning of *shall* is lost in the non-fiction of Period 3 and the dynamic meaning takes over does therefore correspond with Collins' (2009a) findings to some degree. As mentioned earlier, I am careful to draw firm conclusions from such small amounts of data, but to some extent at least, the loss of the epistemic meaning in non-fiction contributes to the trend of general decline in *shall* from Period 2 to 3 being more dramatic than in the other varieties. This provides evidence for a trend toward a degree of monosemy for *shall*, moreover to express dynamic meanings. The trade-off with *will* in non-fiction is indeed related to the epistemic meaning: in Period 3 *will* often expresses the epistemic meaning in this register, as here illustrated in an informative account from an identification guide of South African spiders (107). This, together with the elimination of the epistemic meaning for *shall* in the same period indicates that its potential for conveying epistemic meanings has totally been shifted to *will*[154].

(107) Whatever moves in her vicinity she **will** attack and either kill or drive off; because of this, the males have evolved a complex and amazing variety of courtship rituals... (1990s; Non-fiction [W2B-023])

The trade-offs for 20th-century *shall* discussed above, namely the shift of epistemic meanings to *will* and *BE going to* in news, the shift of dynamic (volition) meanings to *BE going to* in letters and the shift of epistemic meanings to *will* in non-fiction, can be linked to the increasing rarity of *shall*, due to paradigmatic atrophy, distributional fragmentation and a degree of monosemy. Firstly, seeing that first-person subjects are infrequent in informative writing like news and non-fiction, where third-person subjects are more often found, the potential of *shall* to be used in these registers is impeded by its restriction to the first-person. Secondly, the epistemic meaning often occurs in such informative texts as these, and seeing that *shall* often expresses dynamic meanings by the late 20th century, and is somewhat restricted to this use in contemporary SAfE (as the deontic meaning is not as prevalent as in other varieties), the potential for other auxiliaries (*will* and *BE going to*) to express this

[154] In the case of (107), *shall* could not be used to denote the habitual action that *will* does.

meaning is increased. Lastly, the register that retains many first-person subjects to its clauses is letters, but even in this register *shall*, despite its restriction to the first-person, is replaced by the quasi-modal *BE going to*. As letters are of a more interpersonal nature and often contain more informal language (especially in its social component) than other registers, the dynamic meaning of volition occurs the most – a meaning for which *BE going to* has a propensity. The rise of a quasi-modal at the expense of a modal in a more interactive register can of course be linked with colloquialisation.

As a final point of discussion, considering the reference in the text on Cape English quoted in §1.1.3 that *will* and *would* are replacing *shall* and *should*, there is evidence for the former in the historical data, as seen above regarding the trade-off between *will* and *shall*. Yet, this cannot be regarded as an exclusive feature of SAfE considering the general decline of *shall* in BrE, AmE and AusE noted above. The latter suggestion that *would* is replacing *should* will be considered in § 4.3.3.1.3. The two quasi-modals of the volition/prediction cluster, *BE going to* and *WANT to*, will be discussed in the next section.

4.2.1.3.3 *BE going to* and *WANT to*[155]

It is already noted above that in the news register *BE going to* has gradually taken over the expression of volition/prediction meanings from *shall* across the 20th century (from Period 2 to 3), but the overall patterns of *BE going to* remain to be discussed. The data in Table 5.2 show that after the frequency of *BE going to* maintained stability from Period 0 to 1 with 11,9 and 11,7 instances per 100k words respectively, this quasi-modal almost doubles in frequency in Period 2 with 22 normalised instances (+88%), and continues to rise (with deceleration) into Period 3 with 25,4 normalised instances (+15,5%), which causes the overall rise of *BE going to* from Period 0 to 3 to be statistically significant (where $p < 0.05$) with a log likelihood score of 4.67. This log likelihood score is much lower than the one signifying the increase of *BE going to* from 1961 to 1991/2 in AmE, which is 23.5, where it increases by 51,6% (Mair &

[155] No instances of the reduced forms with non-standard spellings *gonna* and *wanna* were found in the historical corpus, but this is not so surprising for written language, given that Leech *et al.* (2009:106) almost exclusively found these forms in the spoken corpora of AmE and BrE.

Leech, 2006:328). The trend in BrE on the other hand is that of relative stability, seeing that *BE going to* decreases, if only by 3 instances pmw (-1,2%) from 1961 to 1991/2 (2006:328). This means that the upward trend for this quasi-modal in SAfE is more comparable with that of AmE, but there is still a 36,1% difference in the rate of increase between these varieties, so it can therefore be said that the trend for SAfE *BE going to* more or less maintains the middle position between the trends of AmE and BrE. Despite the difference in the trends of SAfE and BrE, these two varieties have very similar frequencies of *BE going to* in the late 20th century when the normalised frequency of *BE going to* in BrE (pmw) from Mair and Leech (2006:328) is recalculated to per 100k words to match the data for SAfE: *BE going to* in SAfE and BrE respectively have normalised frequencies of 25,4 in the 1990s and 24,5 in 1991/2. This reveals that, when compared with the recalculated normalised frequency of 33,2 for AmE *BE going to* in 1991/2 (see Mair & Leech, 2006:328), the frequencies of SAfE and BrE *BE going to* are equally different from that of AmE.

The deceleration in increase from Period 2 to 3 is supported by the fact that, although *BE going to* rises in two registers (letters and news), as was mentioned regarding the degree of trade-off with *shall* (§ 4.2.1.3.2), it simultaneously drops in two registers (fiction/narrative and non-fiction). The greatest decrease is found in fiction/narrative, where *BE going to* drops by 33,2%. Out of the 22 raw tokens of *BE going to* in the fiction/narrative register of Period 2, 11 (50%) are from speech representation and 11 (50%) from narration, so there is equal division among these uses. All of the 11 instances of *BE going to* that are used in narration take a preterite form of *be* to describe a past event, as seen in example (108), whereas in all but one of the instances in speech representation *BE going to* takes a present form of *be* to describe a future event, as in example (109). The situation changes in Period 3, where 11 (64,7%) out of the 17 raw tokens of *BE going to* were found in speech representation and 6 (35,3%) in narration. Hence, the changing ratios between the uses of *BE going to* in narration and speech representation respectively from Period 2 (1:1) to Period 3 (1:1,8) can account for the deceleration of the increase of *BE going to* between these periods, which is supported by the fact that the amount of raw instances (eleven) in speech representation remains the same from Period 2 to 3 and only the amount of instances in narration drops.

(108) "The blight of the English is over South Africa," Gerhardus Grobbelaar said. "... We must go away somewhere where there is not the Englishman's flag." ... We **were going to** trek across the Kalahari into German territory. (1940s; Fiction/narrative)

(109) And she said again: "... If I marry old Jan Redlinghuis he will let the water into my father's furrow, and the lands of Zeekoegatt will be saved. I **am going to** do it, and God will help me." (1920s; Fiction/narrative)

On a quantitative level, there is some indication of a frequency trade-off with *will* in the fiction/narrative register from Period 2 to 3, where the ratio of *will* relative to *BE going to* in fiction/narrative changes from 2,4:1 (173,4:72 normalised instances) to 3,8:1 (181,1:48,1 normalised instances) between these periods, i.e. *will* becomes more frequent at some expense of *BE going to*. Among the 64 raw tokens of *will* in Period 3, 2 (3,1%) are used in direct thought representation, 9 (14,1%) in narration and the remaining 53 (82,8%) in speech representation, which, when compared with the distribution of *BE going to* among these literary modes or forms of conventional literary language, is similar in terms of the predominant use in speech representation. When the percentages are converted to ratios, on the other hand, it is apparent that the proportion of speech representation relative to narration in Period 3 differs between 5,9:1 for *will* and 1,8:1 for *BE going to*, which indicates that *will* is the preferred choice in the representation of speech.

This is perhaps surprising given the fact that quasi-modals occur more frequently in speech than in writing, where direct speech representation can be taken as an imitation of spoken language (cf. Biber *et al.* 1999:486). The fact that *BE going to* is somewhat more frequent in narration than *will*, can simply be linked to the propensity of *BE going to* to take preterite, i.e. past, forms of *be*, e.g. *was/were going to* (as seen in example [108]), which are the more prevalent forms in narration (see § 2.2.2.1; cf. Banfield, 1982), and naturally, *will* most often has a future (prediction) reading (see § 4.2.1.3.1; cf. Collins, 2009a:126) and is therefore somewhat counterproductive for use in narration, where its past form *would* indeed appears more often (as seen in [110]).

(110) (She) let herself in to the safe quietness of her home. Both Lex and Denise **would** be back later. (1990s; Fiction/narrative [W2F-005])

Besides these quantitative findings, no semantic trade-off with *will* seems to be present, seeing that both *will* and *BE going to* most often convey the epistemic meaning in the fiction/narrative register in Period 3, as also noted for *will* in § 4.2.1.3.1. *BE going to* expresses the epistemic meaning in 14 cases out of 17 (82,4%), which largely corresponds with the findings of Collins (2009a:144), where more than half of the instances of *BE going to* were found to denote the epistemic meaning in the other three native varieties. Epistemic *BE going to* in the fiction/narrative of Period 3 is exemplified in (111), where it is used with a present tense form of *be* in speech representation. The meanings of *BE going to* and *will* therefore overlap in their expressions of future prediction, but there is a slightly different nuance expressed by *BE going to*, in that it is not restricted to being a marker of futurity in its present *be* form, but can express a general present reference to "situations that are on the point of occurring or are already in train"[156] (Collins, 2009a:144) (as in [111]). This is slightly different to the more common epistemic meaning of *will* that expresses future prediction (as in [112]). The difference in nuance is illustrated here via uses with similar themes.

(111) "It's okay, Johnny. You **are** not **going to** die. It's okay my brother, I'm taking you home." (1990s; Fiction/narrative [W2F-001])

(112) "The person who caused the death of your brother is after you. You **will** die the same way as your brother." (1990s; Fiction/narrative [W2F-003])

The dynamic meaning of *BE going to* is the other meaning that is present in the fiction/narrative of Period 3, albeit much less so than the epistemic meaning, with only 3 such uses (all from speech representation) among the 17 instances (17,6%). This corresponds with Collins' (2009a:144) findings for BrE, AmE and AusE, where dynamic uses are the second most common for *BE going to*. In the following examples

[156] This corresponds with Palmer's (1990:144) concept of "current orientation" (cf. Collins, 2009a:144).

the dynamic expressions for this quasi-modal convey a sense of intention rather than willingness (cf. Collins, 2009a:147).

(113) "Please don't hit me." The man laughed. "I'**m** not **going to** hit you," he said, "but you can't stay here". (1990s; Fiction/narrative [W2F-019])

(114) "One shushu[157] day, when I'm nice and well-done fed up, I'**m going to** tell her to her face." (1990s; Fiction/narrative [W2F-002])

WANT to, unlike *BE going to* that remains stable from Period 0 to 1, rises across all periods, and with much significance. This quasi-modal increases from 5,7 to 16,4 instances per 100k words from the period of input to Period 1 (+187,7%). The increase is statistically significant with a log likelihood score of 4.72 (above the critical value where $p < 0.05$). A further increase with the same level of significance, even if decelerated, is present from Period 1 to 2, where the frequency of *WANT to* basically doubles from 16,4 to 32,6 normalised instances (+98,8%). The rising trend line of this quasi-modal continues into Period 3, but a deceleration once again occurs from 32,6 to 46,3 normalised instances (+42%), rendering the increase non-significant. Overall, however, its continued rise from the first to the last period is highly significant, with a log likelihood score of 34.89 (above the value where $p < 0.0001$). The rise of *WANT to* therefore contributes considerably to the general rising trend for quasi-modals seen in Figure 6.

The trends for *WANT to* are different across SAfE, AmE and BrE. *WANT to* rises much less in SAfE between Period 2 and 3 than in AmE, where it increases with much significance by 70,9% from 1961 to 1991/2. Contrary to this, *WANT to* rises more in SAfE than in BrE, where it only increases by 18,5% (still significant) from the mid to late 20th century (Mair and Leech, 2006:328). The absence of significance for the trend of *WANT to* in SAfE from Period 2 to 3 is therefore a notable difference between this variety and the other large native varieties on a diachronic level, and the fact that the final frequency of SAfE *WANT to* in Period 3 (46,3 instances per 100k words) is slightly higher than the frequency of this quasi-modal in BrE in 1991/2

[157] A colloquial, onomatopoeic expression literally meaning 'hot' (with 'shushu' mimicking a burning sound like that of a fire), as in 'one hot day', but the phrase 'one shushu day' would here translate as the more standard expression 'one of these days'.

when recalculated to 100k words (42,3 normalised instances), and lower than the frequency for AmE when recalculated to 100k words (55,2 normalised instances), matches the concept that SAfE is maintaining the 'middle ground' in terms of the frequency of *WANT to* at the end of the 20th century.

As for the initial spurt of increase for *WANT to* from Period 0 to 1, the register in which *WANT to* increases the most, supporting the general increase during these periods to the greatest extent, is letters, where the frequency of *WANT to* increases from 6,2 normalised instances in Period 0 to 48,7 such instances in Period 1 (+685%). The frequency of this quasi-modal in letters stabilises somewhat in Period 2, with a drop to 40,8 normalised instances, but a rise occurs again in Period 3 with 55,5 normalised instances. *WANT to* in the three remaining registers, news, fiction/narrative and non-fiction, after remaining relatively stable from Period 0 to 1, however experiences drastic increases in frequencies from Period 2 onward, supporting the general trend of increase for this quasi-modal. Letters can therefore be said to be the accelerating factor from Period 0 to 1 and the decelerating factor later on, especially from Period 1 to 2.

There are only three raw occurrences of *WANT to* in the letters of Period 0, which contain a total of 48 313 words, all of them from social letters, and all of them expressing the dynamic volition meaning (as exemplified in [115]).

(115) ...thair [sic] is a small peace [sic] of ground next to Mr Boose is ground wich [sic] I **want to** get you must not let it be marken [sic] for he **wants to** by [sic] it him self [sic][158]... (1840s; Social letters)

The frequency of *WANT to* in social letters increases dramatically in Period 1, supporting the equally dramatic overall rise of this auxiliary from the period of input to Period 1. In Period 1 there are 19 raw instances of *WANT to* in social letters and one in a business letter, totalling at 20 occurrences within the 41 090 words present in the letters register for this period. All but one of these 20 uses of *WANT to* in the letters of Period 1 denote the dynamic meaning (as in [116]), just like in Period 0. The remaining one expresses a deontic meaning (as in [117]). This corresponds with the

[158] This example is from a letter from J. Tramithick to his friend Mr. Jennings at Somerset West, Western Cape, and the peculiar spellings occur exactly as they are quoted.

findings of Collins (2009a:151) where *WANT to* is reported to expresses dynamic meanings in 98% of cases and deontic meanings in 1,1% (the remaining 0,9% is divided among epistemic and indeterminate uses) in the three native varieties he studied.

(116) Death'll come and I've done nothing, and I've so much I **want to** say and so much I **want to** paint. (1880s; Social letters)

(117) He appears to be under a cloud. Innes and Sauer **want** me **to** take his place. (1880s; Social letters)

To conclude the discussion on the prediction/volition cluster, the overall stable for *will* as seen in Figure 9 was firstly linked to small fluctuations in its frequency in the fiction/narrative register from Period 0 to 2, secondly to the balancing out of these fluctuations in Period 3 and thirdly to its simultaneous frequency increases and decreases in two registers in each case from Period 2 and 3. *Will* and *would* maintain a frequency trade-off in non-fiction between Period 2 and 3, which was linked to the textual functions (e.g. argumentative, informative) and the perspective (e.g. historical) of the texts in each period. The very significant decrease of *would* over the last two periods was linked to non-fiction and letters, where a decline in the hypothetical meaning between Period 2 and 3 was reported. No evidence for Bowerman's (2004b:477) claim for *won't* as a directive softener was found.

Furthermore it was found that the general downward trend for *shall* is supported by a movement toward monosemy, where it mostly expresses dynamic meanings by the 1990s: a shift of epistemic meanings to *will* and *BE going to* in news, a shift of dynamic (volition) meanings to *BE going to* in letters and a shift of epistemic meanings to *will* in non-fiction were reported. *Shall* was also reported to be restricted to first-person subjects (paradigmatic atrophy), which hampers its potential to be used in registers containing a lot of informative writing (news and non-fiction), where third-person subjects prevail (distributional fragmentation). *Shall* is also being replaced by *BE going to* in letters to express the dynamic (volition) meaning, even if this register contains many first-person subjects – this was linked with the process of colloquialisation.

As for the quasi-modals of this semantic cluster, the deceleration in the increase of *BE going to* from Period 2 to 3, despite its overall increase from Period 0

to 3, which was linked to the changing ratios of narration and speech representation in the fiction/narrative register between Period 2 and 3. The fact that *BE going to* is somewhat more productive in narration due to its propensity to express past time accounts for its decrease as an effect of the decrease of narration, and hence its decelerated increase from Period 2 to 3. *BE going to* most often conveys epistemic meanings and some dynamic meanings in this register in Period 3, which is not different from the case of the other native varieties. It was also reported that a very significant increase of *WANT to* occurs from Period 1 to 2 and that this quasi-modal continues to rise across the entire time span. The register of letters, and especially social letters, supports the initial upward surge for *WANT to* from the period of input to the period surrounding the late 19th century to the turning of the 20th century (Period 0 to 1), where volition meanings were found to be the most frequently expressed.

4.2.1.4 The obligation and necessity cluster

Section 4.2.1.4 will discuss the frequencies of the quasi-modals *HAVE to, (HAVE) got to*, *NEED to*, *BE supposed to, (had) better,* and *BE to* and the modals *must, should, need* and *ought* of the obligation/necessity semantic cluster in terms of overall and register-internal frequency. I will offer explanations for noteworthy frequencies and comparisons with the patterns of the obligation/necessity cluster in BrE and AmE. The discussion of these modals and quasi-modals, being the largest collection of auxiliaries in a semantic cluster, will shed light on the general trends seen in Figure 6. As mentioned in the introductory section on the diachronic results, the results presented in Table 5.1 show that *must* undergoes an overall increase from Period 0 to 3, even if the increase does not meet the criterion for statistical significance (see § 4.2.1.1). To recapitulate, this increase was found to be remarkable when compared to the statistically very significant decrease of this modal across the 20th century in both BrE and AmE (Leech, 2001:551-4). This section will therefore explore *must* and its competitors in the obligation/necessity cluster more closely in terms of frequency shifts and register variation.

In the previous sections describing the trends of the modals and quasi-modals of the permission/possibility/ability and prediction/volition clusters (§ 4.2.1.2 and §

4.2.1.3), some semantic observations within isolated registers were made where they were relevant to general trends relating to e.g. trade-offs between modals, significant changes or noteworthy accelerations and decelerations of the rate of increase or decrease. In this particular section such semantic observations will only be restricted to the modals *need* and *ought* and the quasi-modals NEED *to*, *(HAVE) got to, BE supposed to*, *(had) better* and *BE to* where they support general trends, seeing that Section 4.3 will be entirely devoted to the semantics of *must, should* and *HAVE to*. Consequently, all of the examples of the uses of these three auxiliaries will be provided in Section 4.3 where they are relevant to the discussion. Figure 10 illustrates the diachronic frequency shifts of the auxiliaries categorised under this semantic cluster.

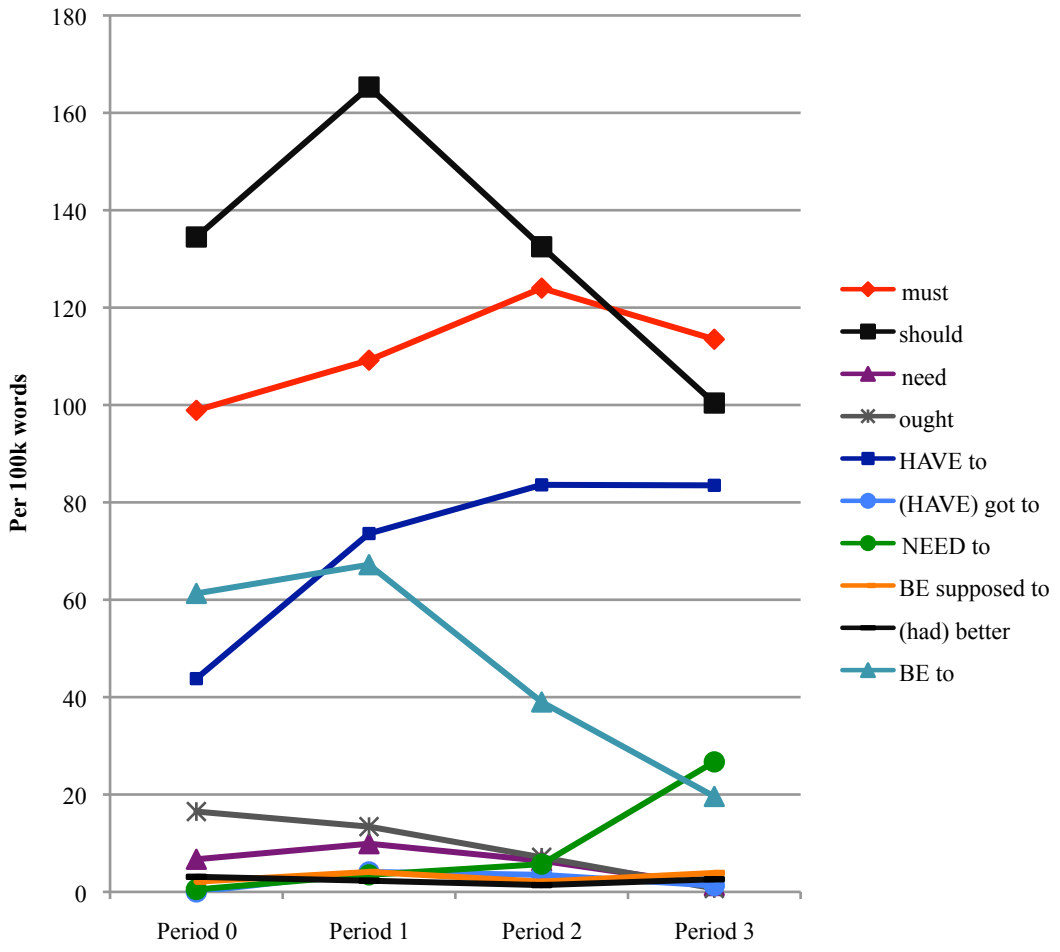

Figure 10

Frequency trends of the obligation/necessity cluster across time

4.2.1.4.1 *Must*, *should* and *HAVE to*

The three auxiliaries with the highest frequencies in Period 3 and indeed across the entire time span (except for the case of *BE to* in Period 0) in the obligation and necessity semantic cluster are *must*, *should* and *HAVE to*. Their interaction on a quantitative level, as discussed in this section, is the incentive for the in-depth analyses of their macro- and microsemantics in § 4.3. *Must,* rising from 98,9 to 109,2 instances per 100k words from Period 0 to 1 (+10,4%), with an accelerated rise from 109,2 to 124 normalised instances from Period 1 to 2 (+13,6%), decreases slightly from Period 2 to 3, finally having 113,5 such instances (-8,5%). None of these changes are statistically significant, as the log likelihood scores here are all below 3.84 (where $p < 0.05$). It is already mentioned that the overall increase of *must* from Period 0 to 3 occurs by 14,8%, which is also non-significant. It is also pointed out above how striking the trend for SAfE *must* is in comparison with its trend in other native varieties (if we take the data to mean, very firmly, that *must* does not decrease in frequency), where this modal shows a salient and sharp decline – Leech *et al*. (2009:114) call the fall of *must* "extraordinarily steep" – that has indeed been noted in various diachronic corpus-based studies (Leech, 2003; Mair & Leech, 2006; Leech & Smith, 2009; Leech *et al*., 2009; Millar, 2009 and Myhill, 1995). As § 2.4.2.1 reported, the frequency of *must* from 1961 to 1991/2 decreases by 29% and 34,4% in BrE and AmE respectively, according to the findings of Mair and Leech (2006:327), and Leech (2011:551-4) confirms its highly significant decline over the course of the 20th century and into the 21st century in both these varieties.

When comparing the rates of decline for *must* in other varieties with matching periods in the SAfE data, as calculated from the data of Leech (2011:551), the following becomes apparent. In BrE *must* at first increases slightly by 4,1% from 1901 to 1931; these periods are comparable with the end of my Period 1 and the middle of my Period 2, where SAfE *must* showed a greater increase by 13,6%. BrE *must* however decreases from 1931 to 1961 by 7,6% to a lower frequency than in 1901, and shows an accelerated decrease into 1991 by a rate of 29%, which is the same rate reported by Mair and Leech (2006). Its decline in BrE continues into 2006 with a further accelerated rate of 34,1%. The longer period from 1931 to 1991 is closer to the time covered between my Period 2 and 3, where this modal decreases by

34,4% in BrE, and when compared to this trend, SAfE *must* only decreases by 8,5% from Period 2 to 3, rendering this modal's decline in SAfE 25,9% less than in BrE.

Where *must* slightly increased in BrE in the early 20th century, Leech (2011:553) reports that AmE *must* maintains a steady decrease[159] from the 1910s to the 1990s (by overall 55,2%) and into the 2000s. From the 1930s to 1960s this modal decreases by 12,9% in AmE, with an accelerated decrease rate of 35,3% from the 1960s to 1990s, being ever so slightly higher than the rate reported by Mair and Leech (2006). From the 1930s to 1990s however, a decline by 43,6% is noted in AmE, which means that AmE *must* declines 35,1% more than SAfE *must* does from Period 2 to 3. Millar (2009:199; 251) confirms the decline of *must* in AmE in the TIME corpus, along with the low-frequency modals *shall* and *ought*, although he generally found the other modals not to decline (cf. § 2.4.2.1), which makes his result of a declining *must* even more noteworthy. Indeed Myhill's (1995:162) results for AmE are not much different, but he reports the decline of *must* since the early 19th century: this modal moves from covering 10% of the frequency amongst all the other modals in 1824 (corresponding to the beginning of my Period 0) to only 2% in 1947 (corresponding to the later part of my Period 2). This pattern of course is totally different from the trend of slight increase in SAfE *must* from Period 0 to Period 2[160]. To summarise the difference in rates of decrease over the mid- to late 20th century, BrE *must* declines 25,9% more and AmE *must* declines 35,1% more than SAfE *must* (cf. Leech, 2011:551-3).

The evidence for a trend toward democratisation, which is argued for as one of the reasons behind the decline of *must* in the other varieties (see § 2.4.2.1), therefore at first appears not to be as strong for SAfE *must*, simply because it declines to a lesser extent, but it might not be as simple as this. If democratisation as a social tendency is present in South Africa, which it undoubtedly would be in recent periods considering its political past and present (see § 2.3.3), then some other linguistic means are the carriers of this trend, or indeed there is something else at work in the

[159] Mair (2014), based on the extended Brown family of corpora, however reports that *must* does increase in frequency from the 1930s to 1961 in AmE (+6,6%), but then decreases by 30,4% from 1961 to 1992.

[160] The percentages as expressions of each modal's proportional frequency among all the other modals that are used in Myhill's (1995:162) data are not really comparable with the normalised frequencies in the data of this study. But to the extent that these changes in proportion suggest frequency changes, a decrease in the role of *must* is detectable, indeed implying a decrease in its frequency.

case of *must*. Bowerman's (2004:477) claims that *must* has less social impact in SAfE than in other varieties and that it is often used as a substitute for polite *should* (see § 1.1.2) suggest the possibility of a semantic overlap between these two modals, i.e. a synonymous relationship, which will be explored in § 4.3.2.

The register-internal trends for *must* in SAfE yielded some interesting results. News is the only register to follow a stable downward trend from Period 0 to 3 (-36,6%), with a slight deceleration of the downward trend from Period 1 to 2, but its overall decline is not as influential as the trends of the other registers on the general rising trend of *must* from Period 0 to 3. The other three registers all undergo a change in the direction that their trends follow in the time between Period 0 and 1 (an increase in letters and non-fiction, and a decrease in fiction/narrative); all of these trends turn in the opposite direction between Period 1 and 2, and continue to follow in those directions between Period 2 and 3, suggesting that some stability in the register-internal trends is maintained from the time between Period 1 to 2 and onward[161]. In Period 2 the frequencies of *must* in all four registers are closer together than in other periods, suggesting that the registers moved toward the bleaching out of register differentiation during the early to mid-20th century regarding the use of *must*. In Period 2 the frequency of *must* varies between 100,4 normalised instances in news, 104,7 in fiction/narrative and 107,8 in non-fiction, with only letters having a generally higher frequency of 170,5 normalised instances. By comparison, the rival modal of *must* in its semantic cluster, *should*, only moves toward more balanced frequencies for the four registers in Period 3. This would also be the case for *must* in Period 3, if it were not for fiction/narrative, which contains almost twice as many instances of *must* by this time (192,4 per 100k words) than the other three registers (97,1 per 100k in letters, 93,6 in non-fiction and 81,1 in news). This suggests that *must* has gained an overall degree of stability in its frequencies across registers by Period 2 and maintains that stability into Period 3, with the only exception being the case of fiction/narrative in Period 3, and to some extent also letters in Period 2. What is interesting about Myhill's (1995) older AmE data is that the trends for individual modals are not influenced by the varying frequencies for some periods, but remain rather clear. The same pattern is visible in SAfE, especially for *must*, where the longer-term trends tend to remain clear, despite sporadic fluctuations.

[161] See the data in Table 21.1 in Appendix 2.

The overall rising trend of *must* from Period 0 to 1, its acceleration from Period 1 to 2 and its slight decline from Period 2 to 3 can be linked to the internal trends for the four registers in the data. From the period of input to Period 1 the frequency of *must* increases the most in the non-fiction register (+191,9%), but also increases to a large extent in letters (+83,8%), which supports its rising trend between these periods. Decreases in frequency are however also noted for news (-14,7%) and fiction/narrative (-33,9%), but these occur to a smaller extent and thus do not counter the influence of the other two registers on the general trend between Period 0 and 1. The accelerated trend of increase from Period 1 to 2 is supported by both the increase of *must* in fiction/narrative by 101,7% (whereas the trend for this register between the preceding periods was one of decrease) and by the relative stabilisation in the trends for letters (-2,7%) – a change in direction, but a low rate of decrease nonetheless – and news (-8,1%) – a lesser extent of decrease as between the preceding periods. The greatest change of direction is however found in the non-fiction register: where this register showed a large increase in the frequency of *must* from Period 0 to 1, it now decreases by 23,2% from Period 1 to 2, but this does not counter the influence the other three registers and their trends have on the overall trend for *must* during these Periods. From Period 2 to 3 no changes in the direction of trends occur, but general deceleration of these trends do occur in two registers: fiction/narrative (+83,8) and non-fiction (-13,2), whereas the acceleration of trends occurs in letters (-43%) and news (-19,2%). These rather counteractive forces of deceleration and acceleration, along with the fact that no changes in the directions of trends occur, contribute to the attainment of some stability and balance in the frequencies of *must* in the four registers from Period 2 to 3. The changing frequencies of *must* each in the four registers over the 20th century will be linked to semantic changes in § 4.3.2.1. The conclusion at this point is that diachronically *must* behaves very differently than in other native Englishes, which shows the need for *must* to be more comprehensively discussed in the section on semantic results (§ 4.3.2). The modal with the most interesting relationship with *must* is *should*, which will be discussed next. It is indeed within this relationship (and the relationship with *HAVE to*) that the uniqueness of SAfE *must* can be appreciated to the full.

The rival modal of *must* in SAfE, viz. *should,* increases from the period of input to Period 1 from 134,5 to 165,3 instances per 100k words (+22,9%), but this trend basically reverses into Period 2, where the normalised frequency (132,5) is

somewhat lower that it was in Period 0, which means that this modal decreases by 19,8% from Period 1 to 2. The downward trend continues from Period 2 to 3, but with an accelerated decline by 24,2% (from 132,5 to 100,4 normalised instances), which is statistically significant with a log likelihood score of 4.69 (above the critical value where $p < 0.05$). A decline by the same level of significance occurs across the entire time span for *should* – an overall decrease of 25,4% – which is even more than the overall increase rate for *must* reported above (14,8%). *Should* also declines significantly in both AmE (-32,5%) and BrE (-21,4%) across the 20th century (from the 1910s-1990s for AmE and from 1931-1990s for BrE), as the data of Leech (2011:551-3) reveal (cf. Mair & Leech, 2006:327), which means that SAfE *should* decreases less over the 20th century than in AmE and more than in BrE (maintaining the 'middle ground' between these two varieties), but in these other native varieties *should* declines less than *must*. As seen in Figure 10, *should* is the highest-frequency modal in the obligation and necessity cluster in all periods but the last, where it is outranked by *must*, that is to say in Period 3 *should* has a lower normalised frequency than *must* (with a difference of 13,1 normalised instances). This situation arises after the trend lines of these two rival modals intersect between Period 2 and 3, rendering *must* the auxiliary in the obligation/necessity cluster with the highest frequency by the late 20th century.

This is certainly very different from the situation in the other native varieties, where *should* is universally more frequent than *must* by the end of the 20th century in both AmE and BrE (Leech, 2011:551-3; Mair & Leech, 2006:327). Mair and Leech's (2006:327) data show that *should* is already more frequent than *must* in BrE in 1961 and stays more frequent in 1991/2. Leech's (2011:551) data reveal that *should* has been more frequent than *must* in BrE even since as early as 1901, and indeed maintains this relationship in all the periods represented in their data (1931, 1961, 1991 and 2006). A different situation is present for AmE, where *must* maintains a higher frequency than *should* from the 1910s all the way to the 1960s, but this pattern changes in the 1970s, where *should* has the higher frequency and continues in this manner in the 1980s, until the 1990s and into the 2000s (Leech, 2011:553). This is confirmed by Mair & Leech's (2006:327) data, where AmE *must* is shown to be more frequent in 1961, but *should* is more frequent by 1991. The main differences between SAfE, BrE and AmE in this regard are thus that while in BrE *should* has a higher frequency than *must* across the entire span of the 20th century, in AmE the crossover

from *must* to *should* as the higher-ranking modal is only made in the latter half of the 20th century. It is hard to say exactly when the crossover in the opposite direction than AmE happens in SAfE, since the historical corpus does not contain data for the period between the 1960s and 1980s, but as far as the trend lines in Figure 10 can suggest an approximate period of overlap, it would seem as if the crossover from *should* to *must* as the higher-ranking modal occurs somewhere around the 1960s or 1970s, seeing that the data for Period 2 end in 1959 and the data for Period 3 start in 1990. This suggests that SAfE and AmE move in their opposite directions in this regard at more or less the same time, which makes the situation in SAfE even more remarkable. Despite the fact that *must* is not such a strong marker of obligation in SAfE as in other varieties (as Bowerman [2004b] suggests), this does not necessarily contradict an explanation based on democratisation, since the meaning of *must* has itself become less authoritarian in SAfE than in BrE and AmE, so there would be no reason for it to decline to avoid 'undemocratic' language. This possibility however remains to be fully explored in § 4.3.

The trend of increase for *should* from Period 0 to 1 is supported the most by its increase in letters (+121,9%), but also by news (+16,1%) and non-fiction (+22,2%), while the decrease in fiction/narrative (-34,7%) does not inhibit the overall pattern. From Period 1 to 2, the trends of *should* in letters (-12,2%) and news (-68,6%) reverse completely, whereas the trends of this modal in fiction/narrative (-15,1%) and non-fiction (+19,2%) continue in the same directions as between the previous periods. Fluctuations are once again present for between Period 2 and 3, where *should* continues to decrease in letters with much acceleration (-54,6), but changes the direction of its trends for all three other registers: *should* decreases in non-fiction by 25,5%, but it increases in news (+43,9%) and fiction/narrative (+20,3%), but these increases do not counter the downward trend between these periods. The acceleration of the decrease of *should* in letters and the change in direction from an increase to a decrease in non-fiction are hence the two factors contributing the most to the overall accelerated and significant decline of *should* from Period 2 to 3. The changing frequencies of *should* each in the four registers over the 20th century will be linked to semantic changes in § 4.3.3.1.

As mentioned in § 2.4.2.2 the popularity of *should* over *must* in BrE and AmE is mainly linked to the increasing monosemy of *should*, where during the 20th century it gradually moves to predominantly express weaker deontic meanings and fewer

epistemic meanings (cf. Leech *et al*., 2009:86-7; Millar, 2009:203). The epistemic meaning is reported to be used more and more by *must* in these native varieties, but simultaneously *must* has also become restricted to stronger deontic meanings, leading to its decrease to avoid such meanings (i.e. democratisation) (cf. Smith, 2003:242; Leech *et al.*, 2009:114-116). If *should* does not increase as a beneficiary of the decrease of *must*, as is the case in SAfE, then some other semantic tendencies are at work here – a topic which will receive thorough investigation in § 4.3.

Figure 10 shows that in SAfE the auxiliary with the third highest usage frequency in this semantic cluster for all periods except Period 0 is *HAVE to*. This quasi-modal rises from Period 0 to 1 and Period 1 to 2, and its frequency stabilises from Period 2 to 3. Its frequency most noticeably changes from 43,8 instances per 100k words in Period 0 to 73,6 instances in Period 1 (+68%), which is very significant with a log likelihood score of 7.71 (above the critical value where $p < 0.01$). It continues to rise by a further 13,6% to 83,6 normalised instances in Period 2 (a decelerated increase), but remains entirely steady with 83,5 normalised instances in the last period (+0,1%). Overall, from Period 0 to 3, the rise of *HAVE to* (+90,6%) is statistically very significant with a log likelihood score of 12.71 (above the critical value where $p < 0.001$). *HAVE to* is only outnumbered by *BE to* in Period 0, but from Period 1 onward it is the highest frequency quasi-modal in SAfE. Despite its constant high frequency after Period 0, it is not the quasi-modal that has increased the most across the entire time frame, since *NEED to* and *WANT to* showed the greatest level of increase from Period 0 to 3 (see § 4.2.1.4.2 and § 4.2.1.3.3 for more detail). The steady trend of SAfE *HAVE to* from Period 2 to 3 is much the same as the trend for this quasi-modal in AmE from 1961 to 1991/2, where it increases by a rate of only 1.1%, but its rate of increase in BrE during the same periods is a larger 9%, rendering the tendencies of AmE and SAfE more similar. Despite the fact that *HAVE to* has not undergone dramatic frequency changes in BrE and AmE across the 20th century, some ongoing semantic developments are reported. As noted in § 2.4.2.2, Leech *et al.* (2009:115) report that *HAVE to* also displays a trend towards monosemy in these other native varieties, where it moves toward expressing more general root (deontic and dynamic) meanings that are less face-threatening and have an objective source of modality (cf. Mair & Leech, 2006; Collins, 2009a). Investigations into the semantics of *HAVE to* in terms of its most frequent meanings, as well as its obligative force and source are reported in § 4.3.4.

The upward trends of *HAVE to* over the course of Period 0, 1 and 2 are sustained by the constant increase of this quasi-modal in all four registers, except in the fiction/narrative register from Period 1 to 2. The significant increase of *HAVE to* from Period 0 to 1 is strongly supported by its high rate of frequency increase in all the registers during this period in time, with non-fiction displaying the highest rate of increase from 13,7 to 68,3 instances per 100k words (+398,5%), but with letters (+105,7%) and news (+103,6%) also having high rates of increase. Fiction/narrative is once again the exception, since it only increases by a smaller 14,1% (from 78,5 to 87,6 instances per 100k words), and as noted above this register changes direction in its frequency from Period 1 to 2 to decrease by 10,4%, which renders its normalised frequency in Period 2 (78,5 normalised instances) very close to its original frequency in Period 0 (76,8 such instances). These back-and-forth fluctuations in frequency for *HAVE to* in fiction/narrative contribute to decelerate its rise from Period 1 to 2 in respect to the former period of increase, along with the dramatically decreased rates of increase present in letters (+20%), news (+27,1%) and non-fiction (38,2%) between Period 1 and 2. The stable trend from Period 2 to 3 on the other hand is supported by the simultaneous decrease and increase of *HAVE to* in two registers in each case (which is similar to the case of *will*, where the same pattern contributed to its stability [see § 4.2.1.3.1]). The frequency of *HAVE to* in letters and news continues to increase between these periods by 44,3% and 39% respectively, and fiction/narrative continues its downward trend, but with a deceleration to 6,2%. Non-fiction is the only register to change direction to decrease by 30,6% between Period 2 to 3 (from 94,4 to 65,5 normalised instances), which is almost the rate of increase it maintained for *HAVE to* during the previous periods, meaning that this quasi-modal nearly returns to what its frequency had been in Period 1 (68,3 normalised instances).

It is interesting to note that despite these above-mentioned fluctuations present in the frequencies of especially fiction/narrative from Period 1 to 2, but also non-fiction from Period 2 to 3, the frequencies of *HAVE to* in letters and news are very similar in each Period, with some instances of similarity in the other two registers at intervals. In Period 0 letters and news have respectively 33,1 and 30,3 normalised instances of *HAVE to* and rise to 68,1 and 61,7 normalised instances in Period 1, where it also has a similar frequency in non-fiction, namely 68,3 such instances. This tendency continues into Period 2, where *HAVE to* has 81,7 and 78,4 instances per 100k words respectively in letters and news, which agrees with its frequency in

fiction/narrative (78,5 normalised instances). Ergo, a lesser degree of register differentiation in terms of the frequencies of *HAVE to* in each of these three periods is present, but a slightly more differentiated pattern occurs in Period 3. During the 1990s the register-internal frequencies of *HAVE to* share less consistency, but the frequencies in letters and news (117,9 and 109 instances per 100k words) are still closer together than those of fiction/narrative and non-fiction, and vice versa for the two latter registers, as their frequencies are also grouped closer together (73,6 and 65,5 normalised instances). This suggests that there is not a very exclusive preference for *HAVE to* in more speech-like registers. The changing frequencies of *HAVE to* in each in the four registers over the 20th century will be linked to semantic changes in § 4.3.4.1.

The above exposition broadly encapsulates the diachronic frequency trends of *must, should* and *HAVE to* as shown in Figure 10, and highlights the uniqueness of especially *must* among the large native Englishes, but to fully appreciate this unique frequency pattern of *must* in SAfE, Figure 11 shows the frequency changes of the three main carriers of obligation and necessity meanings from Period 0 to 3 in clustered column form. This graphic format permits the data to be presented in such a way as to aid the comparison of the frequency of *must, should* and *HAVE to* alongside one another in each period.

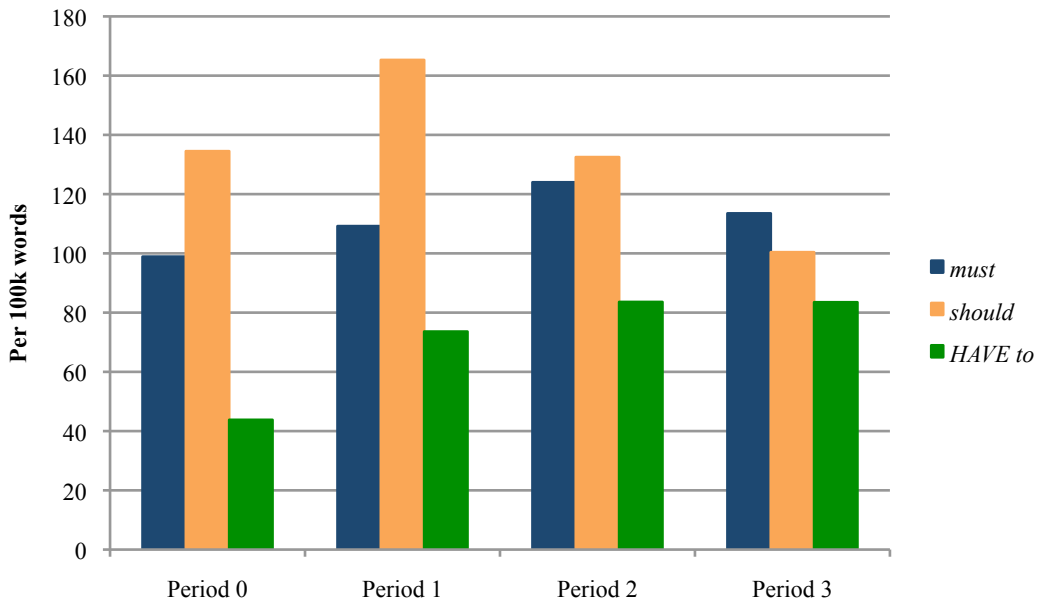

Figure 11

Frequency changes of *must, should* and *HAVE to* in SAfE over time

Figure 11 illustrates the shift in frequency well, in that it shows the movement from overall more differentiated frequencies from Period 0 and 1 to Period 2, to a lesser degree of frequency differentiation among *must, should* and *HAVE to* in Period 3. It also shows the gradually changing relationship between *should* and *must*, where, after the lower frequency of *must* against *should* in the first two periods changes its extent in Period 2, *must* finally becomes the higher-ranking modal in the 1990s, even if its frequency is slightly lower than in the preceding period. The less differentiated frequency pattern of these three auxiliaries in Period 3 suggests a greater degree of correspondence in the quantitative trends of *must, should* and *HAVE to* than in previous periods. In Period 3 more evenly proportioned differences between all three auxiliaries are also present than in previous periods (where there was e.g. a larger difference in frequency between *should* and *HAVE to* in Period 2), which corresponds to the findings reported above that all three these auxiliaries are gaining a degree of stability in the 20th century.

For a comparison with these patterns between Period 2 and 3 in SAfE, Figure 12 illustrates the changing frequencies of *must, should* and *HAVE to* in BrE and AmE from 1961 to 1991/2. The graph is generated from the data of Mair and Leech (2006:327-8) (the frequencies are re-normalised per 100 000 words, to aid comparison with Figure 11).

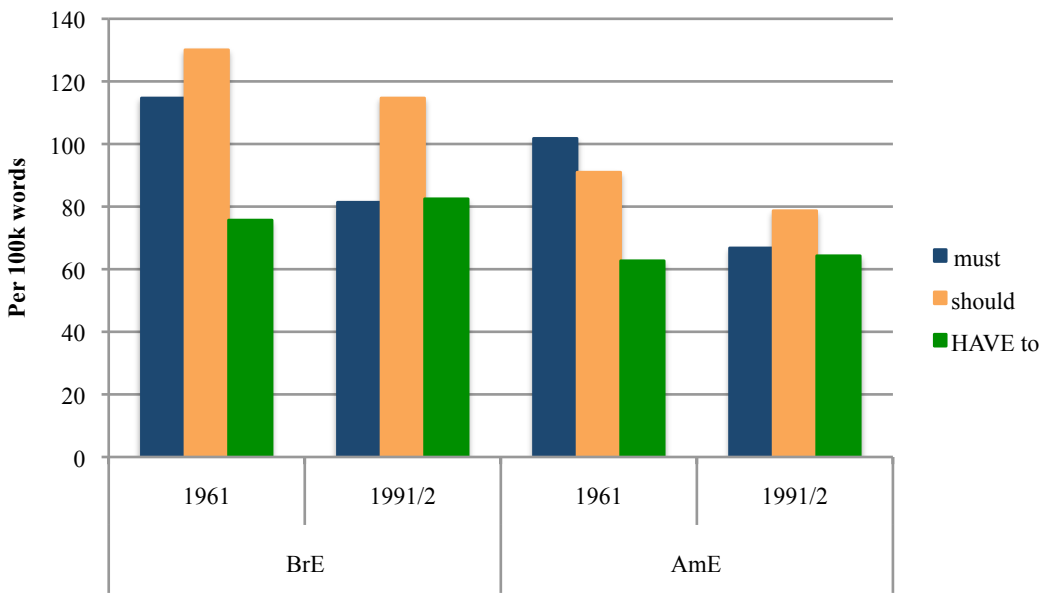

Figure 12

Frequency changes of *must, should* and *HAVE to* in BrE and AmE over time

Figure 12 firstly highlights the extent of the difference in diachronic frequency patterns between BrE and SAfE from the mid- to late 20th century. Here *must* indeed stands out as the auxiliary undergoing the largest degree of change in both varieties. The frequencies of the three auxiliaries for BrE in 1961 is to some extent similar to the pattern for SAfE in Period 2, where in both varieties the frequency of *must* is below that of *should*, but a larger extent of this is present in BrE, i.e. in SAfE the frequency gap between *must* and *should* is not so pronounced as it is in BrE. *HAVE to* in BrE is also much less frequent than its rivals in 1961, just as in SAfE during Period 2. The situation changes rather drastically in the late 20th century. In 1991/2 the patterns are very different in BrE than in SAfE, where in BrE *should* is the forerunner in frequency and the frequency of *must* has dropped to below that of *HAVE to*, which itself is still much lower than that of *should.* As noted above and seen in Figure 11, SAfE *must* is the forerunner in frequency over both *should* and *HAVE to* in Period 3.

A different pattern than for SAfE is also present in AmE, where in Figure 12 it can be seen that the frequency of *must* in 1991/2 is close to that of *HAVE to*, and that *should* is more frequent than both its rivals. This differs from the pattern in BrE firstly in terms of the overall frequency of these three auxiliaries, where in AmE they are generally lower, secondly in terms of AmE *should* not being as decisively more frequent than *must* and *HAVE to* as it is in BrE, and thirdly regarding the more frequent use of *must* than *HAVE to* in AmE, where the reverse is true for BrE. The fact that *HAVE to* has a lower starting point in AmE in 1961 than BrE and ends at a lower point than BrE in 1991/2 is surprising, but not part of an unnoticed tendency in literature. Leech *et al*. (2009:104) refer to what they call a "puzzling aspect of written AmE": the general tendency for quasi-modals in AmE to "[start] from a lower point and [end] at a lower point than [written] BrE" (even if AmE shows a larger increase overall) – here referring to the Brown family of written corpora used by e.g. Mair and Leech (2006). Leech *et al.* (2009:104) offer a "conjectural explanation", which is "that written AmE ...is more subject than written BrE to prescriptive influence". The overall effect of this tendency "may be to make the 'prestige barrier' between spoken and (published) written language less permeable, so that colloquialization is resisted where spoken forms are felt to lack respectable credentials in the written language" (Leech *et al.*, 2009:104). However, Leech *et al*. (2009:105) also mention that the higher-frequency quasi-modals *NEED to, WANT to* and *HAVE to* "do not seem to suffer from the same degree of inhibition" as the other less frequent quasi-modals. As seen

in Figure 12, although this tendency is present for *HAVE to*, it does indeed not display a high degree of materialisation, where its frequency only changes from 757 pmw in BrE and 627 pmw in AmE in 1961, to 825 pmw in BrE and 643 pmw in AmE in 1991/2. Figure 12 further shows that one of the respects in which AmE and BrE does concur is the prominence of *should* over *must,* and this is the one respect in which SAfE differs the most from these varieties. To some extent, both the fact that the 1961 frequency pattern for AmE is similar to the pattern for SAfE in Period 3 with regard to the higher frequency of *must* over its rivals, and the fact that that the 1991/2 pattern for AmE is somewhat similar to the pattern for SAfE in Period 2 regarding the higher frequency of *should* over its rivals, suggest that the 20th-century trends for AmE and SAfE are slightly reversed. As mentioned above, this does not necessarily mean that democratisation is not a factor in the development of SAfE, but at this point I will defer this issue to § 4.3, and turn to the other modals and quasi-modals in the obligation and necessity semantic cluster.

4.2.1.4.2 *NEED to, need* and *ought*

This section explores the frequency trends of the quasi-modal *NEED to* and the marginal modals *need* and *ought*. The quasi-modal in the obligation and necessity cluster that rises the most across the entire data set is indeed *NEED to*, but, as mentioned above, it is outranked by *HAVE to* as the quasi-modal with the highest frequency in this semantic cluster and indeed across all semantic clusters across the entire time span (except for *BE to* in Period 0). *NEED to* starts out with a minuscule normalised frequency of 0,5 in the period of input, representing only one raw token of this quasi-modal found in the news register. Its normalised frequency however rises to 3,5 (6 raw tokens) in Period 1 (+600%) and the rise continues but decelerates with 5,7 normalised instances (8 raw tokens) in Period 2 (+62,9%). A striking acceleration of increase however occurs into Period 3 with 26,7 normalised instances of *NEED to* (41 raw tokens) – a rise by 368,4% – which, with a log likelihood score of 14.45 is very significant above the critical value where $p < 0.001$. Across all four periods *NEED to* increases with great statistical significance, having a log likelihood score of 30.19 for the data from Period 0 to 3 (above the value where $p < 0.0001$) and a stupendous

increase rate of 5 240%, but, given the base frequency of only one raw token in Period 0, this extremely high rate of increase has to be interpreted with some caution.

When compared with the frequency changes in the other native varieties across the 20th century, the level of significance with which *NEED to* rises in SAfE from Period 2 to 3 (at $p < 0.001$) is lower than for both BrE and AmE, seeing that the log likelihood scores for the rise of *NEED to* from 1961 to 1991/2 in these varieties are respectively 83.0 and 33.3 (the levels of significance are both at $p < 0.0001$) (Mair & Leech, 2006:328). Nevertheless, this does not render the increase of *NEED to* in SAfE over the 20th century any less noteworthy. SAfE indeed has a higher rate of increase (+368,4% as mentioned above) from Period 2 to 3 for *NEED to* than these varieties, which are 249,1% and 123,2% for BrE and AmE respectively between 1961 and 1991/2 (2006:328). Leech *et al.* (2009:94;97) confirm that *NEED to* "spectacularly [increases] in frequency" in the Brown family of corpora (BrE and AmE combined) over the 20th century. The register-internal changes of *NEED to* in SAfE between Period 2 and 3 support its general increase between these periods.

The frequency of *NEED to* increases in all four registers from Period 2 to 3. By Period 3, *NEED to* has the highest frequency in the news register with 48,2 instances per 100k words, followed by non-fiction with 26,5 normalised instances, fiction/narrative with 11,3 normalised instances and letters with 6,9 such instances. The register with the highest rate of increase for SAfE *NEED to* from Period 2 is fiction/narrative with a rate of 251,1% (followed by non-fiction [+145%]), but as mentioned above, it is not the register in with the highest frequency by Period 3, so the 100% increase rate for news, increasing from 0 to 48,2 instances per 100k words, is somewhat more noteworthy. *NEED to* in news has fluctuated slightly in its frequency across the four periods, being the only register in which this quasi-modal present in Period 0 (with 1,4 normalised instances representing only 1 raw instance). The frequency of *NEED to* doubles but stays low in Period 1 with 2 raw instances represented by 4,7 normalised instances, before disappearing entirely in Period 2 with zero instances and rising to 19 raw instances (48,2 normalised instances) in Period 3, as also noted above.

Out of these 19 raw instances in Period 3, eleven (57,9%) convey dynamic necessity meanings, were the need is mostly located in external circumstances, as in example (118), seven (36,8%) convey deontic meanings (as in [119]) where there is a low, non-face-threatening level of deontic strength or obligative force and "the

required action is merely being recommended for the doer's own sake" (Collins, 2009a:73), matching the findings for BrE and AmE (cf. Leech *et al.*, 2009:115), and the remaining one instance (5,3%) conveys an epistemic meaning (as in [120]) (cf. Collins, 2009a:73-6).

(118) ...the school allegedly refused to pay the teachers, saying it was the responsibility of the new owners. **Needed to** relocate at the same time, the new owners wanted to relocate... (1990s; News [W2C-014])

(119) The public **needs to** bear in mind that without legitimising the SAPS, that is, improving community police relations, it will not be possible to improve the effectiveness of the police. (1990s; News [W2C-002])

(120) ...there seems every likelihood they will **need to** do so again in the current financial year. (1990s; News [W2C-003])

This situation roughly corresponds with the findings of Collins (2009a:73) that *NEED to* more often conveys dynamic meanings in the native varieties in his study (a combined 62% for AmE, BrE and AusE), followed by deontic meanings (28,9%) and epistemic meanings (2,5%) (6,6% of cases were indeterminate).

Van der Auwera *et al*. (2012:62) note that *NEED to* in BrE and AmE is more often used in "positive polarity contexts", but has also increased in negative uses, although modal *need* remains the more regular option in negative contexts. Indeed Leech *et al*. (2009:80-82), Collins (2009b:289) and Van der Auwera *et al*. (2012) all note that in both AmE and BrE *need* is basically restricted to negative clauses by the late 20th century, i.e. it almost always occurs in the negative form *needn't* in assertive contexts (see § 2.4.2.2). Out of all the 56 raw occurrences of SAfE *NEED to* across all periods 50 (89,3%) have positive polarity and 6 (10,7%) negative polarity. The one raw instance of *NEED to* in Period 0 is positive, and all but one of the 6 raw instances of this quasi-modal is positive in Period 1, but by Period 2 an equal amount of positive and negative uses occur in the 8 raw instances (as seen in [121]). The situation changes again in Period 3, where 39 out of the 41 raw instances of *NEED to* are positive and the remaining two are negated. By the end of the 20th century the use of *NEED to* with positive polarity (as in example [122]) is therefore decisively the most prominent.

(121) And I showed my ticket to a whole lot of convicts, and they worked it out, also, not with paper and pencil, because they didn't **need to**. (1940s; Non-Fiction)

(122) I want to say that since all my people have problems that they **need to** sort out, I will no longer accommodate anyone coming to me with their own political ideas... (1990's; Non-fiction [W2B-001])

By comparison, out of the 41 raw instances of modal *need* across all four periods, 14 (34,1%) have positive polarity and a larger 27 (65,9%) have negative polarity. In Period 0, 5 positive and 8 negative (see example [123]) uses are present in the 13 raw instances of *need*, and in Period 1, 5 positive against 12 negative uses are present in 17 raw instances. The diminishing diachronic frequency of *need* leads to the occurrence of 4 positive and 5 negative uses out of only 9 raw instances in Period 2, whereas all positive uses disappear in Period 3, where the only two raw instances are both negated (as in [124]). The restriction of *need* to negative contexts leading to its decline over the 20th century in SAfE therefore corresponds with the findings for AmE and BrE mentioned above to be experiencing paradigmatic atrophy (Leech *et al*., 2009:80; cf. Van der Auwera *et al*., 2012).

(123) This event has awakened many old reminiscences and melancholy reflections, which I **need** not enter upon. (1830s; Social letters)

(124) ...if common sense and good planning prevail, these **need** not add too much bulk and weight to your pack. (1990s; Non-fiction [W2D-014])

As for the frequency patterns of *need* in SAfE, this modal undergoes an overall decline from Period 0 to 3 (-80,6%), with some internal fluctuations, but this general decline is still statistically significant with a log likelihood score of 5.06 (above the value where $p < 0.05$). After the number of raw instances of *need* in the data per period is noted above, it is worthy to report on its normalised frequencies and rates of increase or decrease between each period. As seen in Figure 10, the frequency of *need* first increases slightly from 6,7 to 9,9 instances per 100k words (+47,8%) from Period 0 to 1, but then the trend reverses almost completely when this modal

decreases to 6,4 normalised instances (-35,4%) in Period 2, rendering its frequency at this point slightly lower than in Period 0. An acceleration of the decrease (-79,7%) however occurs into Period 3, where the frequency of *need* falls from 6,4 to 1,3 normalised instances, which is statistically significant with a log likelihood score of 3.96 (above the critical value where $p < 0.05$). By Period 3 non-fiction is the only register to contain instances of *need*, which are only two, whereas it has a frequency of zero in the other three registers.

Mair and Leech (2006:327) report that *need* declines by 40,2% in BrE from 1961 to 1991/2 and by 12,5% in AmE, and only the decline in BrE is statistically significant to the level where $p < 0.01$ – a higher level of significance than in SAfE from Period 2 to 3. When percentages are calculated from the data of Leech (2011:551) it is visible that in BrE modal *need* diminishes by 41,3% from 1931 to 1991, which is closer to the chronological areas covered by my Period 2 to 3. This rate is close to the one reported above for the period from 1961 to 1991. From the COHA data (§ 2.4.2.1) it is clear that *need* in AmE drops by a larger 59% from the 1930s to 1990s (Leech, 2011:553). The decrease rate of *need* between Period 2 and 3 in SAfE is therefore higher than in BrE and AmE, but the extent of decline is closer to that of AmE.

Collins (2009a:57) reports that the most frequent meaning of *need* in contemporary AmE, BrE and AusE is dynamic necessity (63,2%), which makes this modal semantically similar to quasi-modal *NEED to*, but modal *need* conveys more epistemic (logical necessity) meanings and consequently fewer deontic meanings than *NEED to.* In SAfE during the period of input modal *need*, which has 13 raw instances as noted above, conveys equal amounts of deontic and dynamic meanings (5 cases each, i.e. 38,5% each) and epistemic meanings in 3 cases (23%), whereas the two raw instances in Period 3 both convey dynamic meanings, making contemporary SAfE *need* comparable with the other varieties to the extent that such small raw numbers can give evidence. Examples (125), (126) and (127) respectively illustrate the deontic, dynamic and epistemic meanings of *need* in Period 0.

(125) ... let them refrain from all equivocal measures; and we **need** not hesitate to affirm that their authority will be respected... (1830s; News)

(126) ...and as it is intended to publish the trials as space permits, particulars **need** not be stated here. (1840s; News)

(127) In relieving these, exercise the appropriate office of benevolence, that of mitigating the partial evils arising from general laws; and in this direction of our charity, therefore, we **need** not apprehend any ill consequences. (1820s; Non-fiction)

The other modal along with *need* with an extremely low frequency among the other auxiliaries of the obligation and necessity cluster is *ought*, also called *ought (to)* in some of the literature, as seen in Figure 10. *Ought* is the modal, together with *shall*, that declines with the most statistical significance across the entire chronological spectrum, i.e. from the period of input to the 1990s. Despite starting out with higher frequencies than *need* in Period 0, 1 and 2, *ought* has the lowest normalised frequency of any modal in Period 3, and seeing that it declines by 95,8% throughout the entire chronological span (with a highly significant log likelihood score of 17.23 – above the critical value where $p < 0.0001$), its downward trend is more steady than that of *need*, even though there are some accelerations present in the trend line. *Ought* firstly decreases form 16,5 normalised instances in Period 0 to 13,4 such instances in Period 1 (-18,8%), before the decrease accelerates as the frequency drops to 7,1 instances in Period 2 (-47%) and accelerating again into Period 3, finally reaching 0,7 instances (-90,1%), which is a significant decline above the critical value where $p < 0.05$ (see Table 26.1 in Appendix 3). The only one raw instance of *ought* left by the 1990s can be found in the news register (example [128]) (see Table 20.1 in Appendix 2). The decline of *ought* is also reported for AmE and BrE to occur significantly over the course of the 20th century (Leech, 2011:551-3), which corresponds to the SAfE trend.

(128) The law and the people **ought** to be spared the indignities now to be seen in Utah. (1990s; News [W2E-003])

Collins' (2009a:53) data show that contemporary *ought* expresses deontic meanings 94,6% of the time, close to the meaning of median-degree *should,* across AmE, BrE and AusE combined. Collins adds that "[l]ike *should*, *ought to* can never

be as strong as *must*[162], but nevertheless may convey a forceful representation of what the speaker regards as appropriate or right". Indeed the one instance left by Period 3 in SAfE carries the deontic meaning in the above sense (128). In the period preceding this, representing the early to mid-20th century, all ten raw instances of *ought* also convey the deontic meaning (as in [129]), so this is a stable pattern for 20th century SAfE. These deontic meanings of *ought* in the 20th century are notably not very strong in their conveyance of obligative force, corresponding to the above report of Collins (2009a), and indeed such median-degree uses are the most frequent in the 19th century as well (as in [130] further down).

(129) We had nineteen absconders in twenty-six days, & I felt I **ought** to go back. (1930s; Social letters)

On the same note, it is mentioned in Chapter 2 (§ 2.4.2.2) that *should* and *ought* have enjoyed a high level of synonymy since OE, but that this marked semantic overlap only recently sent *ought* on the path of elimination (cf. Goossens, 1987:130; Harris, 1987:184). In SAfE there is positive evidence of a shift in ratio between the frequencies of *ought* and *should* during the chronological span of the data. A gradual increase in relative difference between these two modals are present from the early 19th century to the mid-20th century, where the ratio of *should* against *ought* changes from 8,2:1 in Period 0 to 12,3:1 in Period 1, and finally to 18,7:1 in Period 2. In the 1990s, however, the ratio of *should* relative to *ought* changes to 143,4:1 – a major widening in difference. Since SAfE *must* is more frequent than *should* by Period 3, ratios were also calculated for *ought* relative to *must*, in order to establish whether *should* is really the modal responsible on a quantitative level for the near annihilation of *ought* in this variety by the late 20th century. The ratios representing the relative frequency of *must* and *ought* stays smaller than those between *should* and *ought* in Period 0 and 1, with ratios of 6:1 and 8,1:1 respectively, but the ratio in Period 2, namely 17,5:1, is closer to that of *should* relative to *ought* than in previous periods. Yet, in Period 3 *must* has a higher relative frequency to *ought* than *should* has, namely 162,1:1, and in consideration of the fact that *must* has been reported to be on the

[162] This is of course based on the assumption that *must* is stronger (i.e. expresses a higher degree of obligation) than *should* in English, moreover in BrE, AmE and AusE, as noted in § 2.4.2.2.

decline due to its restriction to stronger obligative meanings in the other native English varieties, this ratio is truly striking. Based on this account it would appear as if SAfE *must* is more responsible for the decline of *ought* from the mid- to late 20th century than *should* – a assumption purely based on quantitative patterns, but which suggests, once again, that there is something afoot on a qualitative level within the semantic workings of *must* in SAfE (see § 4.3.2). If indeed the meaning of weaker, (i.e. median-degree) obligation is conveyed by *must* in contemporary SAfE, as *ought* does in (128), this will provide some evidence of a semantic trade-off between *must* and *ought* over the course of the development of this variety.

Leech *et al*. (2009:80-2) also mentions the increasing rarity of the negative construction *oughtn't* in BrE and AmE, which contributes to the scarcity of *ought* (as mentioned in § 2.4.2.2). In SAfE *ought* has only 7 negated uses among the 32 raw tokens in Period 0 (21,9%) (as seen in [130]), yet in Period 1, 2 such uses occur in 23 raw instances (8,7%) and by Period 2 only 1 instance with negative polarity remains among the 10 raw instances, but the one instance left in Period 3 has positive polarity (example [128]). This indeed shows a gradual increase in the rarity of negated *ought* in SAfE, just like for the other native varieties.

(130) Why has our old intimate correspondence broken off? It **ought** not to be so. (1830s; Social letters)

The fact that both *need* and *ought* in SAfE only remain present in one register each by the 1990s (non-fiction and news respectively) suggests that their restriction to such formalised or conventionalised published contexts has had an impact on their overall frequency. The colloquialisation process strongly appears to be a factor here, since the restriction of these modals to more formal registers has apparently lowered their usage frequency by the end of the 20th century. Other auxiliaries with low frequencies by Period 3 in the obligation/necessity cluster are *(HAVE) got to*, *BE supposed to*, *(had) better* and *BE to*, which will be considered next.

4.2.1.4.3 *(HAVE) got to*, *BE supposed to*, *(had) better* and *BE to*

Leech *et al.* (2009:105) note that in the Brown family of corpora (representing AmE and BrE [see § 2.4.2.1) "[t]he semi-modals *(had) better*, *have got to* and possibly *be supposed to* occur rather rarely ...because they are associated with personal, emotive, interactive modal functions. These do not easily occur in published written sources except where speech is being imitated..." The fact that the diachronic data of this thesis contain only written texts would hence also account for the low frequencies of these quasi-modals in SAfE.

(HAVE) got to, the lowest-frequency quasi-modal in SAfE by Period 3, is not present in the data for the 1820s-1860s in any of the registers, but first appears in Period 1 with 4,1 instances per 100k words, which is a significant rise (log likelihood score = 5.55 [above the value where $p < 0.05$]). This quasi-modal had indeed only became grammaticalised in English during the 1800s, as noted in § 2.4.1.3 (cf. Biber *et al*., 1998; Krug, 2000:119), and the first uses of *(HAVE) got to* in Myhill's AmE data only occurs in 1896, so it is no surprise that only appears in SAfE for the first time in the 1870s-1990s. Despite this initial rise, *(HAVE) got to* decreases somewhat in Period 2 to 3,5 instances (-14,6%) and continues to decrease at an accelerated speed to 1,3 instances in Period 3 (-62,9%), which is non-significant. The overall change from Period 0 to 3 for this quasi-modal is also non-significant, as this is only a slight rise from 0 to 1,3 normalised instances. This quasi-modal has exactly the same normalised frequency in Period 3 as modal *need*, and only *ought* is even less frequent.

(HAVE) got to also decreases in BrE from 1961 to 1991/2 (Mair & Leech, 2006:328) by 34,1%, which is a much slower decrease rate than in SAfE between Period 2 and 3, but the trend for AmE *(HAVE) got to* is one of increasing frequency by 15,6% from 1961 to 1991/2. Both the decrease in BrE and increase in AmE are non-significant. It is plain that the SAfE results indicate a much lower level of productivity for SAfE *(HAVE) got to* than in AmE and also in BrE, even though the trends of this quasi-modal move in opposite directions in AmE and BrE. Once again an auxiliary in SAfE maintains the 'middle ground' between the trends of BrE and AmE, as in the case of e.g. *WANT to*, which points toward a degree of stability for SAfE *(HAVE) got to* in comparison with its trends in the other varieties.

The only register in which *(HAVE) got to* remains present from its introduction into SAfE in Period 1 to Period 3 is fiction/narrative. 2,4 normalised instances of

(HAVE) got to occur in the letters of Period 1 and 2,7 normalised instances occur in the non-fiction of Period 2 (one raw instance in both cases), but this quasi-modal is not present in these registers in any of the other periods, and *(HAVE) got to* also never appears in news during the full span of the data. It is moreover the fluctuating frequencies of this quasi-modal in fiction/narrative between Period 1 and 3 that support its more general trends, where *(HAVE) got to* first rises from Period 1 to 2 by 35,1% (from 9,7 to 13,1 instances per 100k words) and then declines by 56,5% from Period 2 to 3 ending with only 5,7 normalised instances in Period 3.

All six raw instances of *(HAVE) got to* in the fiction/narrative register in Period 1 convey deontic meanings in speech representation, as in (131), and the one raw instance in letters in Period 1 conveys the dynamic meaning, where a need imposed by external circumstances is expressed, as in (132) (cf. Collins, 2009a:69-71).

(131) "See, Sonny! If you been useter goin' round like a dawg with a tin in ain't any wonder you seen nothin'. You **got ter** walk soft an' keep yer head shut!" (1900s; Fiction/narrative[163])

(132) I don't know if you will be able to read this dreadful writing but we have no ink and **have got to** write with tea put in an old ink pot... (1870s; Social letters)

In Period 2, among the four raw instances of *(HAVE) got to* in fiction/narrative (all in speech representation), only one of those instances expresses a deontic meaning (the deontic meaning is also found in the one raw instance in non-fiction), another instance conveys an epistemic meaning (see example [133]) and the remaining two convey a dynamic meaning similar to (132), where circumstances impose necessity (see example [134]).

(133) ...she thought a moment, and then appealed, with the simplicity of fact: But they **got to** come against it sometime, Mam. (1950s; Fiction/narrative)

(134) "But the English are here now, and we'**ve got to** live with them," Koos answered. (1940s; Fiction/narrative)

[163] This specific example is from *Jock of the Bushveld* by Sir Percy Fitzpatrick, spoken by the character 'Rocky'.

By Period 3, however, the only two raw instances of *(HAVE) got to* found in the data, which are both also from speech representation in the fiction/narrative register, convey deontic meanings, as in (135) and (136).

(135) "Esther," Mrs McCord said, jerking her head in Katie's direction, "tell her she'**s got to** do as I say. I don't think she understands." (1990s; Fiction/narrative [W2F-018])

(136) "You'**ve** just simply **got to** make that extra bit of a effort to keep up standards these days..." (1990s; Fiction/narrative [W2F-008])

Collins (2009a:68) reports that *(HAVE) got to* conveys deontic meanings 84,1% of the time in contemporary AmE, BrE and AusE combined, with dynamic meanings being the second most frequent (12,8%) (minuscule percentages are reported for epistemic and indeterminate meanings). With so few raw instances of *(HAVE) got to* in the SAfE data, especially for the 1990s, which corresponds to the period the ICE-corpora cover in Collins' (2009a) study, it is only possible to say that SAfE *(HAVE) got to* also mostly conveys deontic meanings as far as these small numbers permit.

The fact that *(HAVE) got to* almost exclusively appears in speech representation in the fiction/narrative register, together with the fact that this is the only register in which this quasi-modal consistently occurs from Period 1 onwards, may point to its association with spoken language, as indeed the data from Biber *et al*. (1999) for BrE and AmE show: *(HAVE) got to* is reported to occur in conversation far more than in written registers (1999:489). Indeed Leech *et al*. (2009:94) also note that *(HAVE) got to* and *(had) better* are "virtually being excluded from the written language, except in simulations of speech". The same is found to be true for quasi-modal *BE supposed to* in SAfE, but where *(HAVE) got to* lowers in frequency into the 1990s, the frequency of *BE supposed to* increases into this period.

BE supposed to maintains a low frequency across all the periods, with respectively 4, 7, 3 and 6 raw tokens in Periods 0, 1, 2, and 3. The frequencies of this quasi-modal in Period 0 and 2 are similar (both represented by a normalised frequency of 2,1), and the same is almost true for Period 1 and 3, where normalised frequencies amount to 4,1 and 3,9 respectively. In Period 1 *BE supposed to* has exactly the same normalised frequency as *(HAVE) got to* (4,1), but unlike *(HAVE) got to*, *BE*

supposed to is already present in Period 0, and despite these two quasi-modals following the same direction into Period 2 (a trend of decline), *BE supposed to* increases again in Period 3, as mentioned above. Overall, *BE supposed to* increases by 85,7% from Period 0 to 3, which is a non-significant rise, and since the frequencies of this quasi-modal in Period 0 and 2 are the same, as pointed out above, the same rate of increase applies from Period 2 to 3[164]. Where rates are concerned, this 20th-century trend for SAfE *BE supposed to* is closer to the trend reported for BrE from 1961 to 1991/2 by Mair and Leech (2006:328), which involves a significant increase by 113,6% (with a log likelihood score of 9.2, being above the value where $p < 0.01$), than the trend for AmE, where *BE supposed to* only rises non-significantly by 6,3% over this period of time, but as far as significance goes, SAfE *BE supposed to* is more similar to the trend for AmE.

The rise of SAfE *BE supposed to* from Period 2 to 3 is supported by the fiction/narrative and non-fiction register. *BE supposed to* rises from 3,2 to 8,5 instances per 100k words (one to three raw instances) between these periods (+165,6%), and from 2,7 to 3,1 normalised instances (one to two raw instances) (+14,8%). In Period 2 the single raw instance found in both fiction/narrative and non-fiction respectively convey a deontic (past) (137) and epistemic (138) meaning.

(137) She paid for their food and went to visit them every Sunday, and the cousin **was supposed to** see that they went to school regularly and did not wander about the Location[165] after dark during the week... (1950s; Fiction/narration)

(138) Why then is the argument so consistently trotted out that because the Native **is supposed to**[166] possess interest in a small plot of land in some distant territory something should be deducted from the wages he would otherwise receive? (1930s; Non-fiction)

By Period 3 two out of the three raw instances of *BE supposed to* in fiction/narrative convey deontic meanings (as in [139]) and one expresses an epistemic meaning. One

[164] The generally very low raw frequencies of *BE supposed to* cause these increases to be statistically non-significant, despite their relatively high percentages as rates of increase.

[165] Refers to a designated area of a town or city where the black African population lived under the segregational system.

[166] It is not unlikely that *is supposed to* means 'is postulated to' in this context.

deontic and one epistemic (similar to [138]) meaning were found to be conveyed by the two raw instances of this quasi-modal in the non-fiction of Period 3.

(139) "Are you out of your mind? Since when do you put the television on in my house when you **are supposed to** be cleaning my house?" (1990s; Fiction/narrative [W2F-014])

BE supposed to, as a medium strength auxiliary, is reported by Collins (2009a:80) to have "semantic affinities to *should* and *ought to*", in that it ever so slightly conveys more deontic meanings (50,5% of the time) than epistemic meanings (41,4% of the time) in the large native varieties included in his study (the remainder of instances conveying dynamic and indeterminate meanings). The results for SAfE *BE supposed to* once again seem to reflect the situation in Collins' (2009a) data to the extent that such small numbers can provide evidence.

Quasi-modal *(had) better*, also maintaining an overall low frequency, decreases from Period 0 to 1 (from 3,1 to 2,3 normalised instances [-25,8%]) and Period 1 to 2 (from 2,3 to 1,4 normalised instances [-39,1%]), before rising again in Period 3 to a similar normalised frequency than in Period 1 (2,6 normalised instances [+85,7%]). This means that *(had) better* decreases from Period 0 to 3 (-16,1%). This overall decrease is non-significant, and since the normalised frequencies as well as the raw frequencies of this quasi-modal (respectively 6, 4, 2 and 4 in consecutive periods) are so low, this decline does not bear an influence on the general trend of increase for quasi-modals as found in Figure 6. Indeed Collins (2009a:77) also lists *(had) better* as a "minor item". The AmE and BrE trends reported by Mair and Leech (2006:328) for the 20th century both move in the opposite direction than the trend of SAfE: their report is one of (non-significantly) declining numbers by -17,1% in AmE and by -26% in BrE from 1961 to 1991/2. The increase of SAfE *(had) better* from Period 2 to Period 3, where it returns to and slightly exceeds its previous frequency in Period 1 (see Table 21.2 in Appendix 2) is motivated by the letters register, which is indeed the only register containing any occurrences of *(had) better* by the 1990s. In letters the frequency of this quasi-modal rises from 2,4 normalised instances in Period 2 to 27,7 such instances in Period 3 (a sizable increase rate of 1 054,2%).

The two raw instances of *(had) better* that occur in Period 2, one from the

(social) letters register (140) and one from speech representation in the fiction/narrative register (141), both convey the deontic meaning, grounded in the future and part of a direct address with a 2nd person subject.

(140) You will get first-class tuition in Latin, so you **had better** bluff yourself that you like it. (1950s; Social letters)

(141) Then Jarvis said, "I don't want to lose your company, but if you want to go home dry, you'**d better** be starting. This'll be no ordinary storm." (1940s; Fiction/narrative)

All the instances of *(had) better* in Period 3, as noted above, occur in letters, moreover social letters, and also convey a deontic meaning that is grounded in the future, but not all part of a direct address. Three out of the four subjects of the instances of this quasi-modal in the social letters of Period 3 have 1st person subjects (as in [142]) and one had a 3rd person subject (143). In (143) *(had) better* is used without the auxiliary *had*, which is a generally frequent form of use for this quasi-modal (Leech *et al*., 2009:95).

(142) With November drawing to an end we realised we **had better** get our Christmas cards so we went to Rosebank. (1990s; Social letters [W1B-001])

(143) With any luck my stall will work out – it **better** because I'm opening it in one week! (1990s; Fiction/narrative [W1B-003])

Collins (2009a:77) regards *(had) better* as "essentially monosemous" in conveying the deontic (directive) meaning with a subjective source, which is "best described as 'advisability'" (cf. Jacobsson, 1980:52). As seen in the uses above, all but one case of SAfE *(had) better* in Periods 2 and 3 convey exactly this meaning. In example (141) *(had) better* does not express an advisability or directive meaning, even though it has a subjective source (the speaker), but the deontic expression directs toward something impersonal grounded in external circumstances, rather than a personal addressee. Collins (2009a:79) further mentions that *(had) better* mostly occurs in speech (cf. Leech *et al*., 2009:94). It is therefore not surprising that the only

instances of *(had) better* in SAfE's Periods 2 and 3 occur in social letters, which is a more speech-like, informal and interpersonal register (as noted in § 2.2.2.1), and in speech representation in fiction/narrative, which speaks for itself.

As also previously mentioned, quasi-modal *BE to* has the highest frequency in Period 0, not just in the obligation/necessity cluster, but also for all quasi-modals (61,3 instances per 100k words). Maintaining this frequency in Period 1 (67,2 normalised instances), only a small increase occurs by 9,6% from Period 0 to 1. Downward trends are however noticeable between Periods 1 and 2, as well as between Periods 2 and 3. With 39 instances per 100k words in Period 2, *BE to* declines by 42% from the 1870s-1900s to the 1910s-1950s – a decline that is statistically significant with a log likelihood score of 7.48 (above the critical value where $p < 0.01$). The downward trend accelerates slightly toward Period 3 where there are 19,6 normalised instances of *BE to* (-49,7%) – this decline is also statistically significant with a log likelihood score of 6.23 ($p < 0.05$). Overall, *BE to* declines by 68% from Period 0 to 3 – a highly significant decline with a log likelihood score of 21.74 ($p < 0.0001$), making *BE to* the only quasi-modal showing a general decline in frequency that is significant.

The rate of 49,7% with which *BE to* declines from Period 2 to 3, as noted above, is closer to that of AmE when compared with the rate of decline in BrE from 1961 to 1991/2 (-17,2%), but still the rate in SAfE is somewhat higher than the rate of decline that Mair and Leech (2006:328) reports for AmE from 1961 to 1991/2 (-40,1%). The decline of *BE to* in both BrE and AmE over the 20th century is statistically significant (above the critical value where $p < 0.01$ in BrE and above the critical value where $p < 0.0001$ in AmE). Leech *et al*. (2009:95) mention that *BE to* is "particularly infrequent in speech", which is especially true for contemporary AmE.

The overall significant decline of SAfE *BE to* from Period 0 to 3 is supported to the greatest extent by the non-fiction register, where this quasi-modal decreases from a considerable 158 normalised instances in Period 0 to only 15,6 such instances in Period 3 (-90,1%) (see Table 21.2 in Appendix 2), but there are also noteworthy decreases in letters (-80,8%), news (-48%) and fiction/narrative (-43,3%). Non-fiction and letters are indeed the registers that contain the most instances of *BE to* in the 1820s-1860s. Of the 23 raw instances of *BE to* in the non-fiction of Period 0, the majority (11 raw instances or 47,8%) convey epistemic meanings, which basically express future temporality as part of an informative use (as in [144]), 4 convey

dynamic meanings, where theoretical possibility similar to the meaning of dynamic *can* is expressed (as in [145]), 3 convey weaker deontic meanings similar to that of *should* (see example [146]), 1 instance conveys a conditional meaning (147) and the remaining 4 instances convey meanings that were classed as indeterminate, but which are actually formulaic uses that tend towards deontic expressions, as in (148) (cf. Collins, 2009a:83-6).

(144) There **is to** be a new paper in Cape Town, and it **is to** be called the Standard. (1840s; Non-fiction)

(145) Weather is unusually cold at present and business dull. Coffee **is** not **to** be had for love or money and peas are up in the market. (1840s; Non-fiction)

(146) Having corrected ourselves then so far, the question recurs "what **is to** be done to the Eastward within the Orange River Sovereignty?" (1820s; Non-fiction)

(147) But if they **are to** do us really good we must let them come in our midst... (1820s; Non-fiction)

(148) It **is to** be hoped that our farmers exert themselves to the utmost... (1840s; Non-fiction)

In Period 0 only 7 (20%) of the 35 raw instances of *BE to* in letters occur in social letters and the remaining 28 (80%) are from business letters. Among these raw instances nearly equal amounts of epistemic and deontic meanings are expressed. 16 instances convey deontic meanings, as in (149), 15 instances convey epistemic meanings of temporality, as in (150), and only 3 convey dynamic meanings (similar to [145]), whereas one indeterminate or formulaic meaning similar to the one conveyed in (148) was found.

(149) I beg most respectfully to be informed where I **am to** apply for the same or what channel it is mostly likely to have been sent. (1820s; Business letters)

(150) I should not have troubled you but I have been referred by that Lady to your Office and as I **am to** sail for New Orleans in a few days I sincerely trust you will take it into consideration as early as possible... (1820s; Business letters)

There is no clear pattern in the proportions of *BE to* in the registers over the entire time span, i.e. there is some continual fluctuation as to which register contains the most instances relative to other registers across the periods. Over the four periods, a shift in weight in favour of *BE to* in the news register is however observable, after some fluctuations in proportional frequency with the other registers, to the effect that this register contains the most instances of *BE to* in Period 3. Next in frequency in Period 3 is the fiction/narrative register, but its overall pattern is also one of fluctuation: this register did not continually contain the second most instances of *BE to* in preceding periods. Out of the 11 raw instances of *BE to* in the news of Period 3, 10 (90,9%) convey the temporal, epistemic meaning (151) and only 1 (9,1%) conveys the conditional meaning (152). All 7 raw instances of *BE to* in the fiction/narrative register are found in narration, where 3 of these instances convey epistemic meanings (similar to the use in [151]), 2 deontic meanings, as well as 1 dynamic (153) and 1 conditional meaning (similar to the use in [152]).

(151) What the public had not been told is that Sasol colliery **is to** open a huge open-cast coal mine on the banks of the Vaal. (1990s; News [W2E-002])

(152) Warning that mindsets of the past had to go if advertising **was to** succeed in South Africa, Mpanza quoted from a recent Unisa study which found that the personal disposable income of blacks would soon overtake that of whites by a significant margin. (1990s; News [W2C-008])

(153) How do they know that the food they particularly like, **is to** be found in a particular place? (1990s; Fiction/narrative [W2F-012])

The fact that *BE to* occurs mostly in business letters, non-fiction, news and the narration parts of the fiction/narrative register points to its restriction to more formal contexts that are the least representative of speech, which corresponds with the finding of Leech *et al*. (2009:95) mentioned above that *BE to* is infrequent in speech.

Collins (2009a:83) found that *BE to* in AmE, BrE and AusE combined is basically used with equal frequency to convey epistemic (temporal) (32,4%) and conditional (32,2%) meanings, but is only slightly less popular in the deontic meaning (23,8%) (the rest of its meanings are dynamic [expressing theoretical possibility] or indeterminate). In the combination of four registers analysed semantically for Periods

0 and 3, the epistemic meaning is indeed most frequently conveyed, as reported above, except in the letters of Period 0, where deontic meanings were more frequent by one raw instance, which is not a big difference, but epistemic meanings occur especially frequently in the news of Period 3. This piece of the results suggests that SAfE *BE to* is not much different to the other native varieties concerning its main meaning. There is however also a difference: SAfE *BE to* does not seem to convey as many conditional meanings as in AmE, BrE and AusE.

For the deontic uses of *BE to* in the other native varieties, Collins (2009a:83) found that they are often "strongly subjective, comparable to deontic *must*" and can report aggressive commands, but are more commonly objective, which is closer to the meanings of *HAVE to*. *BE to* is also prone to weakening in certain contexts, e.g. when it takes a passive complement (e.g. *...the strength of the police authorities is something which is to be applauded* [example from Collins, 2009a:84]), which is closer to the meaning of deontic *should*. Because *BE to* has the potential to convey all the different nuances of deontic meanings traditionally expressed by *must, should* and *HAVE to*, it can be said these rival auxiliaries have taken over much of the deontic uses in contemporary times, which is helped by the fact that *BE to* is somewhat restricted to written language.

Based on frequency, this appears to be the case for deontic meanings in SAfE too, seeing that *BE to* is a much higher-frequency item in the obligation and necessity cluster in Period 0 (with a higher frequency than *HAVE to*) than it is by Period 3 (with a lower frequency than *NEED to*). This can be deemed a similar case to that of *ought* and *should* (or *must* in the case of SAfE – see § 4.2.1.4.2), where the combination of morphosyntactical markedness[167] and the semantic overlap with rival forms (cf. § 2.4.1.2) are not "conducive to the survival of any linguistic item" (Harris, 1987:184), and in this case that linguistic item appears to be *BE to*. The same situation may be present with *will* in its expression of temporal meanings, where *will* remains stable in SAfE, if rising slightly over the 20th century, whereas *BE to* drops dramatically in frequency, which suggests that *will* is simply used to convey temporal meanings to avoid the level of formality that is attached to *BE to.* Furthermore, the decline of *BE to*

[167] Leech *et al*. (2009:95) in their discussion of *BE to* as part of the modal idioms (cf. Quirk *et al*., 1985:137), note that it is an anomalous item that have traits of semi-auxiliaries (because of its restriction to finite forms of *be* in its construction), but then again it "shows itself atypical of semi-modals in declining sharply in frequency" (Leech *et al*., 2009:95).

in SAfE can also be linked to the dynamic meaning of possibility. Seeing that *can* in SAfE has such a high frequency by Period 3 and is productive in its dynamic possibility meanings (see § 4.2.1.2.1), this suggests that these meanings (as in [145]) are, once again, simply expressed by the less formal option *can*. The fact that *BE to,* a quasi-modal, declines over the 20th century, is therefore not counter-evidence of colloquialisation progressing over this era, but precisely because *BE to* is associated with formal contexts and written language it has been replaced by more unmarked options in SAfE, which in itself is suggestive of the colloquialisation process, seeing that a move away from more formal styles is detected.

The much lower frequencies of the quasi-modals discussed in this section, *(HAVE) got to, BE supposed to, (had) better* and *BE to* by Period 3, when compared to their frequencies in written AmE and BrE by 1991/2 (Mair & Leech, 2006), suggest a lesser degree of productivity in SAfE and hence some stronger competition from other options to express deontic meanings in the obligation/necessity cluster. The case of SAfE *must*, being the most productive modal in this semantic cluster, once again, comes to mind here (see § 4.3.2). A summary of the diachronic results for SAfE according to semantic clusters will highlight which trends compare or contrast with the trends in other native varieties.

4.2.1.5 Summary: diachronic changes

This section will summarise the results of the diachronic study conducted for SAfE, firstly in terms of how the SAfE trends for individual modals and quasi-modals compare with the trends of AmE and BrE in terms of both frequency and semantics, according to general patterns and patterns within semantic clusters, and secondly in terms of how the overall trends for the modals and quasi-modals of SAfE compare with the overall trends in these varieties. Both similarities and differences will be noted here.

It is evident that the quantitative trends of the individual modals and quasi-modals of SAfE over the 20th century do not correspond more closely to either those of AmE or BrE, and even if correspondence is present, there are some subtle differences in the extent of agreement. In some cases the SAfE trend from Period 2 to

3 generally corresponds to the 20th-century trends of AmE[168], as in the case of increasing *BE going to* (although it rises less in SAfE), the stability of *HAVE to* (remaining stable in both SAfE and AmE), the rise of *NEED to* (although it rises with less significance in SAfE), the decline of *need* (declining more than in AmE) and the non-significant increase of *BE supposed to*, but sometimes the SAfE trends agree with both the trends of AmE and BrE, e.g. the decline of *ought* and *BE to*, as well as the declines of *might, shall* and *would* (but these modals decline more in SAfE). In some cases the SAfE trends for the 20th century do not correspond with those of either of these varieties, as in the case of more stable auxiliaries (when compared to either significant changes in other varieties or to frequencies retaining the 'middle ground' between those of the other varieties) including *will*, *WANT to*, *(HAVE) got to, should* and especially *must* (for which no frequency changes are significant in SAfE), significantly rising *can* (which is more stable in BrE and rises less in AmE), significantly declining *could* (which rises in both BrE and AmE), slightly increasing *may* (which declines in both other varieties) and slightly declining *BE able to* (which increases slightly in other varieties), as well as non-significantly increasing *(had) better* (whereas this quasi-modal declines non-significantly in both AmE and BrE). Hence, out of the twenty auxiliaries studied in this thesis, the trends of five correspond more closely with that of AmE than for BrE for the 20th century (four of which are quasi-modals with the only modal being *need*) and the trends of another five correspond with the trends of both AmE and BrE (this time four modals with the only quasi-modal being *BE to*). One half of the auxiliaries, indeed one half each of the modals and quasi-modals, adhere to the 20th-century trends for the large native varieties and especially so for AmE. The remaining ten auxiliaries do not entirely agree with the trends of either these varieties, five of which have more stable patterns or a pattern between those of AmE and BrE in SAfE (three modals and two quasi-modals), whereas opposite trends for five auxiliaries in SAfE are present when compared with these other varieties (once again three modals and two quasi-modals).

To conclude on the discussion according to frequency changes over the 20th century within semantic clusters, the only two modals that really increase in SAfE are both found in the permission, possibility and ability cluster, namely *may* and

[168] This corresponds to the idea of a trend towards Americanisation, especially relating to the plausibility of American influence on the decline of modals in BrE (see Leech *et al.*, 2009:43-4; cf. Trüb, 2008; cf. § 2.3.3.2).

especially *can*. The other increasing auxiliary verbs are mostly quasi-modals from the obligation and necessity cluster, *NEED to, BE supposed to* and *(had) better*, but *BE going to* of the prediction and volition cluster also increases. The two auxiliaries that increase with a degree of stability setting in over the 20th century are modal *will* and quasi-modal *WANT to*, both from the prediction and volition cluster. It is noticeable that the modals that increase or show stability with some increase are all present forms of modals, viz. *can, may* and *will.* The traditional past forms of these modals all decline over the 20th century, namely *could* and *might* from the permission, possibility and ability cluster (along with quasi-modal *BE able to* from the same cluster) and *would* from the prediction and volition cluster (along with semantically corresponding modal *shall*). The two other modals that decline prominently are *need* and *ought* along with the quasi-modal *BE to*, which all from the obligation and necessity group. The remaining four auxiliaries in the obligation and necessity group all decline to some extent, but in fact rather attain a degree of stability from Period 2 to 3 along with the different degrees of decline. Among these are the two modals *must* and *should*, as well as the quasi-modals *HAVE to* and *(HAVE) got to*.

By discussing the 20th-century trends of the modal and quasi-modal verbs of SAfE in this manner, this thesis has attempted to moved beyond the previous emphasis on anomaly in (especially grammatical) SAfE studies, i.e. the tradition to document only those features that are peculiar when compared with the parent variety (BrE) (see § 1.1), which intends provide a more comprehensive perspective on SAfE grammar and expand its potential for full description. It is of course true that within this more comprehensive approach there are certain striking frequency trends that attract attention in the discussion of the diachronic results. In order to clarify certain noteworthy frequency patterns among the individual modals and quasi-modals, most notably the highly significant increase of *can,* the stability of *will*, and the relative stability of *must* achieved with its non-significant decline from Period 2 to 3 after increasing in previous periods, certain semantic trade-offs were noted, as for e.g. *can* and *could,* and *can* and *BE able to*, as well as other semantic tendencies for the auxiliaries within certain registers, as e.g. the tendency for *can* to convey more dynamic possibility meanings in especially non-fiction. The stable trend of *will* over the 20th century was linked to the balancing effect of various frequency fluctuations on a quantitative level, and especially to the continuity of dynamic and epistemic meanings on a qualitative level. The frequency interrelation between *must, should* and

HAVE to were found to be particularly noteworthy and was postulated as a possible means of clarification for the unique behaviour of *must* in SAfE. These modals and one quasi-modal in SAfE will receive further and in-depth attention in Section 4.3, in order to explore semantic motivations behind their quantitative inconsistencies in comparison to other varieties.

The overall trends for modals and quasi-modals in SAfE (see Figure 6) are not radically different from the overall trends in AmE and BrE involving rising quasi-modal frequency and declining modal frequency, but the extent to which these trends occur is different in SAfE: the modals decrease more and the quasi-modals increase less than in the two other native varieties over the course of the 20th century. To the degree that the rise of quasi-modals is indicative of the colloquialisation process taking place across the 20th century, the smaller increase of SAfE quasi-modals from Period 2 to 3 suggest that this process is at work to a lesser extent than in other varieties, but this was not taken as conclusive evidence, seeing that the overall decline of auxiliaries, as seen in Figure 5, is not consistent with the process of colloquialisation (§ 4.2.1.1) (cf. Mair & Leech, 2006; Rossouw & Van Rooy, 2012). The overall trends for the registers revealed that modal frequencies across letters, news, fiction/narrative and non-fiction, as well as quasi-modal frequencies in letters and news, are converging in the 1990s, i.e. a narrowing stylistic gap between registers is detected, but no direct link with the process of colloquialisation could be established. Some indicators of the colloquialisation process surfaced in the discussion on individual modals.

Motivation for the larger degree of decline in modals in SAfE over the 20th century was e.g. found in the larger decline of modal *would* from Period 2 to 3 than in other varieties over the 20th century, where in its decrease in hypothetical uses in letters, i.e. a lesser degree of formality (by means of politeness and tentativeness) links with the overall concept of colloquialisation (§ 4.2.1.3.1). Other modals that decline more in SAfE than in other varieties are *might*, *shall* and *need,* as well as the decline of *could*, whereas this modal increases in other varieties. The decline of e.g. *shall* was found to be motivated by the shift of epistemic meanings in news and dynamic meanings in letters to quasi-modal *BE going to*, and the rise of *BE going to* as an effect of this shift was linked with colloquialisation (§ 4.2.1.3.2). The decline of modals *need* and *ought*, as well as quasi-modal *BE to* as a by-product of their restriction to more formal contexts also point toward colloquialisation in terms of a

move away from formal styles (§ 4.2.1.4.2 and § 4.2.1.4.3). These trends for individual modals and quasi-modals are indicators that the colloquialisation process is at work in SAfE on the level of individual auxiliaries to some extent at least. Apart from colloquialisation, the other reason reported for the general trends in BrE and AmE, especially the decline of *must*, is democratisation (see § 2.4.2.1). The fact that *must* does not significantly decline over the 20th century and rises over the 19th century in SAfE, as noted in § 4.2.1.4.1, does not ordain that the democratisation process is less present in SAfE during the 20th century, but that there might be something else at work in SAfE *must*, as will be discussed in § 4.3.2. Furthermore, it was also noted that the register distribution of *must* in Period 2 and 3 gained some convergence and hence stability in its 20th century trend, as did *HAVE to* to an extent (§ 4.2.1.4.1).

This being the diachronic face of SAfE modality in comparison to other large native varieties of English especially over the 20th century, the synchronic picture needs to be sketched in order to present a chronologically complete account of modality in SAfE. Even though no conclusive evidence for colloquialisation is found in the diachronic results, the progression (if at all) of this process needs further exploration with the addition of spoken data, which currently is only available for contemporary SAfE.

4.2.2 Synchronic results

4.2.2.1 Introduction

As mentioned in the previous section, no overall, conclusive proof of colloquialisation in SAfE is found in the diachronic results, except in the cases of some of the individual modals and quasi-modals in the 20th century. The first goal of this section is therefore to establish whether colloquialisation is at work on a larger scale when the spoken data for contemporary SAfE are added to the discussion. The second goal of this section is to explore the synchronic frequencies of the individual modals and quasi-modals in speech and writing. This especially involves a comparison of the

frequency of *must* in speech and writing relevant to its rivals in the obligation/necessity cluster, namely *should* and *HAVE to*, in order to compare the findings with the overall results of Collins (2009a) and the specific results for speech and writing in Mair and Leech (2006) and Leech *et al.* (2009). The third goal is to present the results for the contemporary Afrikaans data (both written and spoken) (cf. § 3.2.2.3), and compare it with the contemporary data for SAfE to determine whether there is evidence of a possible quantitative similarity between Afrikaans and SAfE. As an extension of the latter analysis, the frequencies of the senses of Afrikaans *moet/moes* and SAfE *must* in the spoken data will be compared as part of the overall semantic discussion in § 4.3.2.2.3, in order to establish whether there is evidence of such a similarity on a qualitative level as well, especially regarding uses with 2nd-person subjects. The overall contemporary modal and quasi-modal frequencies of SAfE are portrayed in Figure 13, as generated from the synchronic data in Tables 24.1 and 24.2 in Appendix 2.

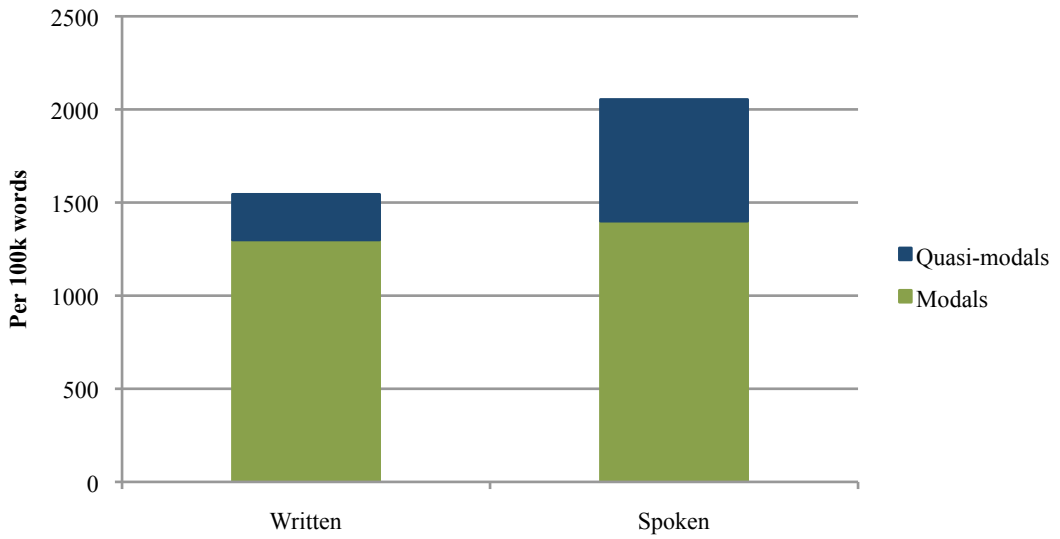

Figure 13

Modal and quasi-modal frequencies in contemporary SAfE for written and spoken registers

Figure 13 shows that the modals maintain a somewhat higher frequency in spoken than written SAfE (having 100,9 modals per 100 000 words more in speech), with a speech/writing ratio of 1,08:1, but that the quasi-modals are much more frequent in speech than in writing (by 408,4 per 100 000 words), with a

speech/writing ratio of 2,67:1, as is indeed expected (see e.g. Huddleston, 2002; Biber *et al.,* 1999). Furthermore, when calculating ratios the other way around, the modal/quasi-modal ratio is 5,30:1 in written and or 2,14:1 in spoken SAfE. A comparison to the findings of Collins (2009a:159) reveals that the speech to writing ratio for quasi-modals in SAfE (2,67:1), as mentioned above, is most similar to that of AusE (2,59:1), which more or less maintains the middle ground between the ratios of AmE (4,33:1) and BrE (2,01:1). A comparison of the proportional use of quasi-modals in speech and writing across the varieties can be seen in stacked column form in Figure 14, for which Collins' (2009a:159) above ratio counts were used as data, together with the counts for SAfE.

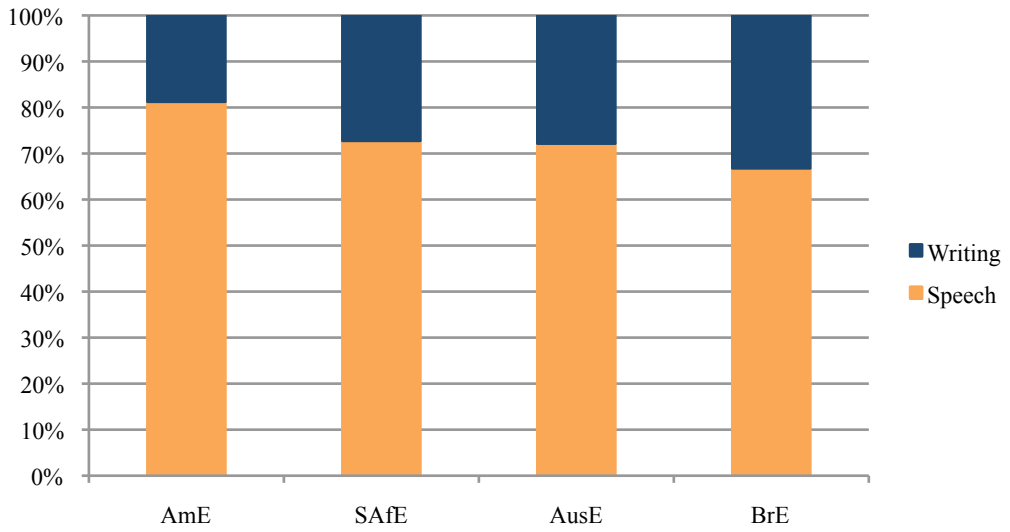

Figure 14

Speech/writing ratios for quasi-modals across native varieties of English

It was noted in § 2.4.2.1 that colloquialisation is implied by the combination of Collins' (2009a:9) findings for contemporary BrE, AmE and AusE that the quasi-modals are more frequent in the spoken register in all these varieties, with the findings of e.g. Mair and Leech (2006) and Leech *et al*. (2009) that the quasi-modals increase in both speech and writing in all these varieties across the 20th century, which seems to be indicative of the increasing popularity of a speech-like style of language involving a shift in preference from modals to quasi-modals. This is based on "the tendency for innovations to spread rapidly in informal, spoken genres before becoming more broadly established in the language" (Collins, 2009a:159) and the

consequential 'pull' that spoken language would have on written language to follow in this direction. However, finding evidence of colloquialisation in a variety of English is, of course, not merely as simple as this. To the extent at least where a more marked "stylistic gulf between semi-modals and modals" (Collins, 2009a:9), or in other words, to the extent that the synchronic quasi-modal differential between speech and writing predicts language change in other English varieties can be taken as evidence of the workings of colloquialisation (cf. van der Auwera *et al.*, 2012), Figure 14 suggests that this process is in a more advanced stage in AmE than in other native varieties, i.e. the variety is more innovative, while the process is making more conservative progress in BrE in the 20th century. Collins (2009a:159) notes that AusE, in "occupying a middle position", appears to "distance itself from both the innovativeness of the Americans and conservatism of the British".

Figure 14 illustrates something of the point of progression of colloquialisation in SAfE in comparison with its progress in other large native varieties. On the account of the proportional difference between quasi-modal usage in speech and writing, SAfE is in a more advanced stage of colloquialisation than BrE, but the progression of this process is rather more conservative than AmE in the 20th century. This placement is comparable with the intermediate progression of AusE, as noted above, even if SAfE appears to be in a slightly more advanced stage of colloquialisation than AusE. The addition of spoken data to the discussion on general results for SAfE therefore does yield some indication at least of colloquialisation having progressed more than in AusE and BrE, and less than in AmE, over the 20th century, given the frequency of quasi-modals in spoken language in Figures 13 and 14, but seeing that the overall diachronic trend for quasi-modals in Figure 6 is that of a less prominent increase than in both AmE and BrE (cf. § 4.2.1.1), the evidence is not strong. Still, even the slight increase of quasi-modals in SAfE can be indicative of some degree of colloquialisation. The overall synchronic and diachronic patterns pertaining to quasi-modals are therefore not as consistent with each other like they are for other varieties in this respect, and consequently, it remains inconclusive whether colloquialisation as a widespread process is indeed at work in SAfE, and if it is, to what extent exactly.

On a more specific level, the case for synchronic trends indicating possible colloquialisation based on the assumption that diachronic predictions can be made by considering synchronic findings (cf. § 2.2.1.1), would be strengthened if quasi-modals were found to replace modals on an individual level, especially in spoken SAfE. The

contemporary results for the separate modals and quasi-modals of SAfE in speech and writing will accordingly be presented and discussed in the next section, together with a comparison to other native varieties.

4.2.2.2 Modals and quasi-modals in speech and writing

This section will discuss the quantitative results for the individual modals and quasi-modals across speech and writing in SAfE. The discussion will firstly focus on overall frequencies in terms of semantic clusters for comparison with the findings of Collins (2009a) (cf. 2.4.2.3) and Collins (2005), and secondly on the frequencies of the modals and quasi-modals in speech and writing where frequencies will be compared to those of Mair and Leech (2006) and Leech *et al.* (2009), as well as Biber *et al.* (1999). Tables 6.1 and 6.2 present the synchronic findings from the spoken and written[169] components of ICE-SA.

Table 6.1

Normalised synchronic frequencies of modals in written and spoken SAfE

Register	*can*	*could*	*may*	*might*	*must*	*shall*	*should*	*need*	*ought*	*will*	*would*	TOTAL Modals
Written	251,7	131,7	129,8	29,3	113,5	8,5	100,4	1,3	0,7	347,6	184,5	1299
Spoken	418,4	143,6	46,2	39,6	130,3	6,1	104	2,5	2,5	254,2	252,7	1399,9
TOTAL	372,8	140,3	69,1	36,8	125,7	6,8	103	2,1	2	279,7	234	1372,3

Table 6.2

Normalised synchronic frequencies of quasi-modals written and spoken in SAfE

Register	*BE able to*	*HAVE to*	*(HAVE) got to*	*NEED to*	*BE supposed to*	*BE going to*	*WANT to*	*(had) better*	*BE to*	TOTAL Quasi-modals
Written	35,9	83,5	1,3	26,7	3,9	25,4	46,3	2,6	19,6	245,2
Spoken	39,8	158,8	61,7	37,1	10,8	188,3	146,3	4,7	6,1	653,6
TOTAL	38,7	138,2	45,2	34,3	8,9	143,7	118,9	4,1	9,8	541,8

[169] The same selection from the written data of ICE-SA was used as in the diachronic analysis (cf. § 3.2.2.2). The repetition of the figures for the written data is only given here as a comparative base for the newly added spoken figures.

It was noted in § 2.4.2.3 that in Collins' (2009a) ICE-data the modals and quasi-modals from most to least frequent in AmE, BrE and AusE combined are: *can, will, would, could, HAVE to, should, may, BE going to, might, WANT to, must, BE able to, NEED to, (HAVE) got to, BE to, shall, BE supposed to, ought, (had) better* and *need.* In SAfE the order of overall frequency is *can*, *will*, *would, BE going to, could, HAVE to, must, WANT to, should, may, (HAVE) got to, BE able to, might, NEED to, BE to, BE supposed to, shall, (had) better, need* and *ought.* Here the most notable differences in frequency rank are that SAfE *BE going to, must, WANT to and (HAVE) got to* have higher rankings than in the other varieties, whereas SAfE *might* ranks much lower on the list than in these varieties. The rest of the auxiliaries have similar ranking orders in SAfE and in the other varieties, with a few small variations that are not drastic (e.g. the order of the three least frequent auxiliaries is different for SAfE and the other varieties, but still these remain the least frequent auxiliaries, viz. *(had) better, need and ought*). The most noteworthy frequency differences between the overall frequencies of the modals and quasi-modals of SAfE (the totals in Table 6.1 and 6.2) and the three large native varieties combined (as in Collins, 2009a) will be highlighted below in terms of semantic clusters.

Just like in AmE, BrE and AusE, (Collins, 2009a:92) SAfE *can*, *could* and *may* overall rank as the modals with the highest, second highest and third highest frequencies in the possibility, permission and ability cluster (cf. § 2.4.2.3). The high rank of *can* in SAfE corresponds with the diachronic findings for this modal regarding its highly significant rise over the 20th century (cf. § 4.2.1.2.1), but *can* is however not the highest-frequency modal in SAfE by Period 3, with *will* occupying this position (cf. § 4.2.1.5). The overall dominance of *can* in the synchronic data is motivated by its much higher frequency in speech than in writing, amplifying the overall frequency gap between *can* and all the other auxiliaries, along with the lower spoken frequency of *will*. Where *might* is next in frequency for the other native varieties, followed by *BE able to*, in SAfE *BE able to* is slightly more frequent than *might*. In the possibility, permission and ability cluster the biggest anomaly is therefore the low frequency of *might*, which corresponds to the diachronic finding that *might* declines more in SAfE than in other varieties, where this was found to occur due to a semantic trade-off with *may* regarding the expression of (epistemic) future possibility meanings (cf. § 4.2.1.2.2). Seeing that *may* maintains its synchronic frequency rank in SAfE to a similar degree as in other varieties, this does not contradict the diachronic finding.

In the prediction and volition cluster, the order of high-ranking *will* and *would* and low-ranking *shall* in SAfE is the same as for other varieties, but the two quasi-modals in this cluster *BE going to* and *WANT to* do not share this similarity of frequency order (Collins, 2009a:126). Despite *would* declining sharply over the 20th century in SAfE (cf. § 4.2.1.3.1), this modal maintains its high rank among the modals due to its high spoken frequency (however overall slightly less frequent than *will*), and reciprocally, despite the stable trend for *will* over the 20th century, it remains a high-frequency modal because of its higher frequency in writing. As noted above, this very fact has caused *will* to give way to *can* as the higher-frequency modal in contemporary SAfE with the addition of spoken data. Overall SAfE *BE going to* outranks *HAVE to* and *should* from the obligation and necessity cluster, as well as *may* from the possibility, permission and ability cluster, whereas the opposite is true for the other varieties. As noted in § 4.2.1.3.3, both the trends of *BE going to* and *WANT to* in SAfE maintain the middle position in-between the trends of AmE and BrE, which, it itself, does not support their higher frequencies in the synchronic data, where the key to their high frequencies in contemporary SAfE is indeed in the spoken data.

In the obligation and necessity cluster the quasi-modal *HAVE to* is the most frequent auxiliary in both SAfE and the other native varieties. One big difference, however, is that *must* is overall more frequent than *should* in SAfE, which is the direct opposite of the situation in the other varieties combined (Collins, 2009a:34). In SAfE *must* is indeed overall more frequent than *WANT to, should, may* and *might*, maintaining its place as a mid-frequency modal in SAfE, and only being outranked by *could* in this regard, whereas *must* is less frequent than all these above-mentioned auxiliaries in the other varieties. The frequency relationship between *HAVE to*, *must* and *should* in SAfE is therefore slightly different in the diachronic and synchronic data, seeing that *HAVE to* is spurred by its high spoken frequency (as is the typical tendency for quasi-modals) in the synchronic data to become the highest-ranking item among its rivals. The differences between the spoken and written frequencies of *must* and *should* are not as drastic as for *HAVE to*, especially in the case of *should*, which allows *HAVE to* to maintain its higher frequency in relation to these modals. In SAfE quasi-modal *(HAVE) got to* is more frequent than in the other varieties, even more so than *NEED to*, which is the higher-ranking item in AmE, BrE and AusE combined (Collins, 2009a:34). The remaining five auxiliaries in this semantic cluster have quite similar ranking orders in SAfE and the other varieties. Both *BE to* and *BE supposed to*

are respectively the fifth and fourth least frequent items in this cluster for SAfE and the other native varieties, but, as noted above, the order of the three last items differs in SAfE, where *ought* is the least frequent auxiliary of all, and in the other varieties, where *ought* is more frequent than *(had) better* and *need* is the least frequent auxiliary of all. A discussion of the spoken and written frequencies of the modals and quasi-modals will shed some more light on the overall tendencies reported above.

Table 6.2 shows that all the quasi-modals except *BE to* are more frequent in speech than in writing. Indeed this quasi-modal also shows a preference for the written register in the Brown Corpora of both AmE and BrE (Mair & Leech, 2006:328; Leech *et al*., 2009:100-2). For SAfE this is unsurprising given the significant decline of *BE to* noted in the diachronic results, due to its association with formal contexts (cf. § 4.2.1.4.3), which was linked to the colloquialisation process. The ratio for the writing:speech frequency of contemporary *BE to* in SAfE is 3,2:1, which is smaller than the ratios for contemporary AmE (6,3:1) and BrE (10,2:1) in the favour of writing (the ratios were calculated from the data of Mair & Leech [2006:328] for Frown and F-LOB, and Leech *et al*. [2009:100-2] for LCSAE [Longman Corpus of Spoken American English] and the spoken component of the BNC [British National Corpus]). As far as these ratios go, SAfE *BE to* can be said to be less restricted to use in more formal contexts than in other varieties, which can indicate a slightly lesser degree of colloquialisation.

For the modals, Table 6.1 shows that only *may*, *shall* and *will* have higher written than spoken frequencies. The preference of mid-frequency modal *may* in written SAfE agrees with the tendency for this modal to occur the least in the conversation register of the LSWE Corpus (Longman Spoken and Written English Corpus) for AmE and BrE combined (Biber *et al*. (1999:5; 489), and the most in the academic writing register, followed by the news register. It was indeed noted in § 4.2.1.2.2 that in Period 3 SAfE *may* has the highest frequency in the non-fiction register (which includes academic writing [see § 3.2.2.2]), followed by the news register, where it mostly expresses epistemic possibility. This, in turn, correlates with the finding of Biber *et al*. (1999:491) that *may* mostly conveys the meaning of extrinsic possibility, which is just a different term denoting the (logical) epistemic category, as discussed in § 2.4.1.1. The ratio of writing:speech for SAfE *may* is 2,8:1, which is similar to the ratio of 2,2:1 for AmE, whereas this ratio is more pronounced in BrE, viz. 7,3:1. Low-frequency modal *shall* on the other hand has a writing to

speech ratio of 1,4:1 in SAfE and 3:1 in AmE, whereas the higher frequency of *shall* in spoken BrE result in a ratio of 1:2. In both these cases, the ratios of SAfE *may* and *shall* for writing relative to speech are closer to that of AmE. Like the case of *BE to*, the restriction of *shall* as a perceived archaic or formalised expression to the written register – what Leech *et al*. (2009:81) call "distributional fragmentation" – particularly in AmE, points to a degree of colloquialisation, as a move away from such formalised uses is made. As SAfE *shall* does not have such a pronounced proportional difference between its spoken and written frequencies as in AmE, it can be argued that a lesser degree of colloquialisation occurs in SAfE in the case of this modal.

There is however a larger degree of evidence for colloquialisation in SAfE in the frequency relationship between *will* (and to some extent *shall*) and *BE going to* in written and spoken registers, when considering the higher written frequency of *will* and the higher spoken frequency of *BE going to* as mentioned earlier. The ratios for writing:speech in *will* and *BE going to* are respectively 1,4:1 (which is exactly the same ratio as for *shall*) and 1:7,4. Colloquialisation can therefore be noted to the degree that a quantitative trade-off between a quasi-modal with a higher frequency in speech and a modal with a higher frequency in writing can be regarded as evidence toward more informal styles in speech. Support for this in the diachronic results of the prediction and volition cluster is not as strong for *will* and *BE going to* as it is indeed for *shall* and *BE going to*, where it was found that a semantic shift of epistemic meanings in news and dynamic meanings in letters from *shall* to *BE going to* occurs over the 20th century (cf. § 4.2.1.3.2). The fact that *shall* and *will* has a semantic trade-off between epistemic meanings in news and non-fiction over the 20th century suggests to some extent at least that although both *will* and *BE going to* express more epistemic meanings in news, the higher frequency of epistemic meanings of *will* in non-fiction influences its higher written frequency in the synchronic data, i.e. it is associated with expressions of future prediction in more formal contexts – more so than *BE going to*, and by implication the move away from such formal uses in speech (in the case of *will*) points toward a degree of colloquialisation.

As a last remark on the prediction and volition cluster in SAfE, and in further consideration of Bowerman's (2004b:477) claim for the supposed usage of *won't* as a directive softener in requests in SAfE, the spoken tokens of *won't* were also analysed for this usage, seeing that no such uses were found in the diachronic data (see §

4.2.1.3.1). It was found that only two such uses of this contracted, negated modal form occur in the 105 tokens in the spoken component of ICE-SA, which are given below:

(154) Uh when my wife goes out walking and I'm here she will say the carpet needs vacuum cleaning or **won't** you do the dishes... (Direct conversation [S1A-030])

(155) Uh excuse me, **won't** you bring me my bag please... (Direct conversation [S1A-082])

In example (154), the request is not a direct one, but it is rather used in the narrative of a past situation or the projection of a hypothetical situation were the wife made or would make a request by means of *won't*. A direct request is however found in (155), where *won't* is indeed used as a directive softener. This in a way supports Bowerman's (2004) claim qualitatively, but the frequency of these instances, being two, and actually only one such direct request in a corpus of over half a million words, which is compounded when considering that these are also the only instances of this usage in the entire data set of this thesis (diachronic and synchronic), there is no real quantitative support for this claim.

The next few paragraphs will focus on both the overall frequencies, as well as the spoken and written frequencies, of the three highest-frequency auxiliaries in the obligation and necessity cluster (*must*, *should* and *HAVE to*), for which the diachronic trends were noted as striking in § 4.2.1.4.1. The data used to generate Figures 15 and 16 can be seen in Tables 27 and 28 in Appendix 4. The overall AmE, BrE and AusE data of Collins (2009a:34) are compared to the matching ICE data for SAfE in Figure 15[170], with the addition of NZE data from Collins (2005:269).

[170] For comparative purposes in Figure 15 and Figure 16, the data for SAfE was re-normalised per million words.

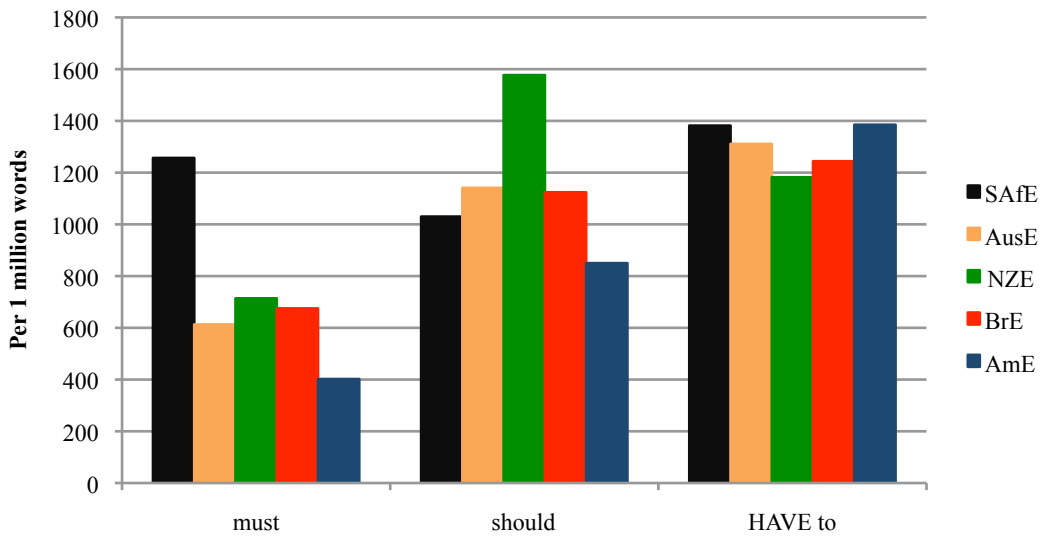

Figure 15

Combined spoken and written frequencies of *must, should* and *HAVE to* in ICE-corpora across native English varieties

The precedence of *must* in SAfE when compared to the other native varieties is plain here. This modal is 150% and 76% more common than in AusE and NZE respectively, 86,2% more common than in BrE and a substantial 212,6% more than in AmE – clearly at the expense of *should* (calculated from the data in Table 27 in Appendix 4). The average frequency of *must* in the four other native varieties is 601 normalised instances; this calculates to SAfE *must* being on average 109,1% more frequent than in other native varieties – i.e. just over twice as frequent as in these varieties. This is certainly a noteworthy finding for SAfE. As mentioned in § 2.4.2.3, *should* basically enjoys double the frequency of *must* in the other native varieties in the ICE-corpora, but for SAfE *should* occurs 22% less than *must*, which means that this semantic cluster is evidently leaning in the opposite direction in modal preference than BrE, AmE and even fellow SHEs AusE and NZE, albeit not to the same extent. Only *should* in AmE has a lower frequency than it has in SAfE, and in NZE this modal clearly stands out as the highest-frequency item in this modal trio.

It is interesting to note that *should* is not only more frequent than *must* in the native varieties of English, i.e. BrE, AmE, AusE and NZE (representing Kachru's inner circle [cf. § 2.2.1.2.1]), but this is also true across the board for other World Englishes, that include the non-native (outer circle) varieties Philippine English (PhilE), Singapore English (SingE), Hong Kong English (HKE), Indian English

(IndE)[171] and Kenyan English (KenE) – all represented by ICE-corpora in Collins (2009b). The frequencies reported for these varieties for *must* against *should* (Collins, 2009b:286) calculate to a combined ratio of 1:2 for the inner circle varieties and 1:1,9 for the outer circle varieties, so, in effect, *should* is overall twice as frequent as *must* in nine of the larger World Englishes. This renders the case of SAfE *must* being more frequent than *should* even more noteworthy.

For all the other varieties except NZE, quasi-modal *HAVE to* is the most frequent of these three auxiliaries, with AmE having the highest frequency here, which is only very slightly higher than the frequency of *HAVE to* in SAfE. This renders *HAVE to* the item with the most stable high frequency among the five native varieties in Figure 15, in terms of the higher degree of similarity between the normalised frequencies. As noted earlier, the addition of spoken data to the discussion is responsible for the slight dominance of *HAVE to* over *must* and *should* in contemporary SAfE, whereas this quasi-modal had a lower frequency than its two rivals in the written data for the 1990s (see Figures 10 and 11). It is therefore useful to compare these three auxiliaries in terms of written and spoken frequency, which will be done below. Despite this, the major difference in frequency distribution between *must* and *HAVE to* in SAfE compared with the other varieties becomes clear when percentages of difference are calculated. AmE has the largest difference in frequency between these auxiliaries, with *HAVE to* being 244.5% more frequent than *must*. Furthermore *HAVE to* is more frequent that *must* by 113,9% in AusE, by 84,3% in BrE, and by 65,5% in NZE. In SAfE, however, *HAVE to* is only more frequent than *must* by 9,9%, rendering the proportional difference much smaller than for the five other varieties.

In order to investigate how the findings of SAfE compare with those from other corpora than ICE, some of the data used in the diachronic section for other varieties were added to the discussion here, but comparisons are still drawn to the ICE-data from Collins (2009a) as well. When comparing the data of Leech *et al*. (2009:100-2) for contemporary spoken AmE and BrE, respectively from the LCSAE (Longman Corpus of Spoken American English) and the conversational part of the BNC (British National Corpus) to the data of Mair and Leech (2006:327) for contemporary written AmE and BrE, respectively from Frown and F-LOB, the

[171] The data of Balasubramanian (2009:110) confirm this tendency for *should* to be twice as frequent than *must* for Indian English.

following pattern emerges from the data in Table 28 in Appendix 4, as illustrated in Figure 16.

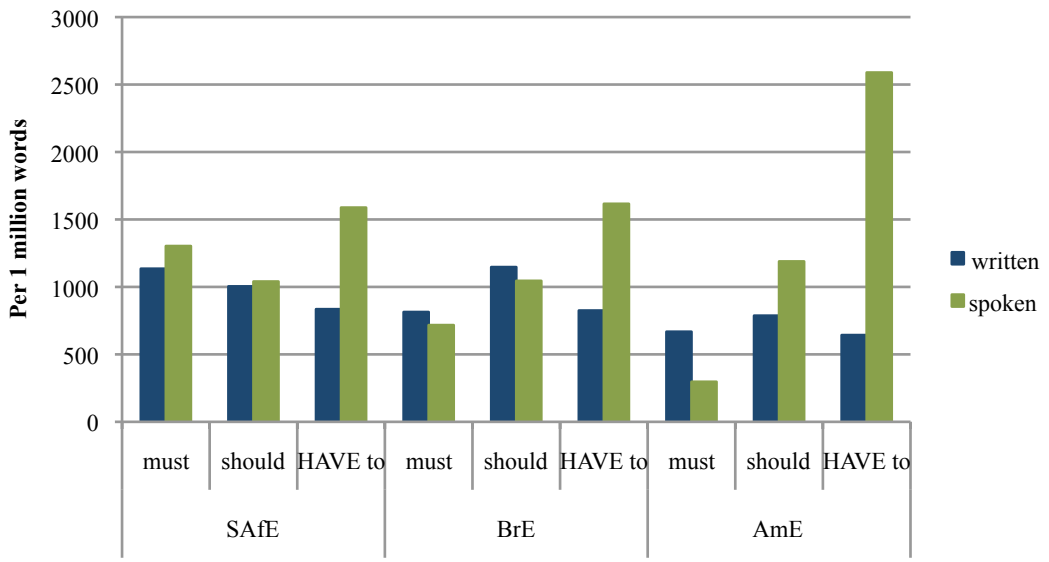

Figure 16

The spoken and written frequencies of *must, should* and *HAVE to* in SAfE, BrE and AmE

From the data in Figure 16 the following patterns for AmE, BrE and SAfE is apparent. AmE *must* is less frequent in speech (297 instances [pmw]) than in writing (668 instances [pmw]), and BrE *must* is also slightly less frequent in speech (717 instances [pmw]) than in writing (814 instances [pmw]), but to a lesser extent than in AmE. As percentages, these figures calculate as AmE *must* being 124,9% more frequent in writing, whereas BrE *must* is 13,5% more frequent in writing. Similarly, in the ICE-corpora Collins (2009a:44) found written *must* on average to be more frequent in not only BrE an AmE, but in AusE as well, where *must* is on average 65,1% more frequent in writing than in speech for all the varieties combined. Support for this tendency can also be found in the data of Biber *et al*. (1999:489) for the LGSWE corpus, where *must* is reported to be more frequent in fiction and academic writing than in conversation for AmE and BrE combined. Indeed Leech *et al*. (2009:101) note that the quasi-modals *HAVE to, NEED to* and *(HAVE) got to* are more frequent than *must* in spoken AmE. As Tables 6.1 and 6.2, as well as Figure 16 show, only quasi-modal *HAVE to* has a higher spoken frequency than *must* in SAfE, so the ranking order for the entire obligation and necessity cluster in terms of spoken

frequencies in SAfE is: *HAVE to*, *must*, *should*, *(have) got to*, *need to*, *be supposed to*, *be to*, *(had) better*, and, in joint last place, *need* and *ought*, which means that *HAVE to*, *must* and *should* remain the biggest role players in this semantic cluster, just like in written SAfE (cf. § 4.2.1.4.1). This is supported by the fact that, despite its higher written frequency than in the other two varieties, SAfE *must* is itself more frequent in speech than in writing, as Table 6.1 also reveals. The difference is however not very great: spoken *must* (1303 instances [pmw]) in SAfE is 14,8% more frequent than written *must* (1135 instances [pmw]), which is close to the percentage with which written *must* is more frequent in BrE – i.e. the pattern for SAfE and BrE is the direct opposite. When compared with the spoken frequencies for *must* in the other varieties, SAfE *must* in speech is 338,7% more frequent than in AmE and 81,7% more frequent than in BrE.

Only the BrE written frequency of *should* (1 147 instances [pmw]) is higher than its spoken frequency (1 045 instances [pmw]) by a slight 9,8%, whereas for both AmE and SAfE the spoken frequency of *should* is higher. In AmE spoken *should* (1 189 instances [pmw]) is 51,1% more frequent than written *should* (787 instances [pmw]), but in SAfE *should* in speech (1 040 instances [pmw]) is only 3,6% more frequent than in writing (1 004 instances [pmw]). The results of the ICE-data of Collins (2009a:52) are somewhat different, where *should* is reported to be slightly more frequent in writing in BrE, AmE and AusE, but for all three of these varieties the difference in speech and writing frequencies are not extensive, with written *should* being on average 22,1% more frequent than spoken *should* for the varieties combined. In the data of Biber *et al.* (1999:489) for BrE and AmE combined, there is no difference in the frequencies of *should* between conversation, fiction and academic writing, but its frequency is slightly lower in news. All and all, the tendency for *should* to be used more frequently in one register or another across the various corpora for the other native varieties is less clear than for *must*.

It is unsurprising that *HAVE to*, being a quasi-modal, enjoys such prominence in the spoken register for all three the varieties in Figure 16. Yet, the extent to which this occurs is different in each variety. *HAVE to* in spoken AmE (2 589 instances [pmw]) is the most prominent by 302,6% over written AmE (643 instances [pmw]), whereas spoken *HAVE to* in BrE (1 616 instances [pmw]) is 95,9% more frequent than in writing (825 instances [pmw]). In SAfE *HAVE to* in speech (1588 instances [pmw]) is 90,2% more frequent than in writing (835 instances [pmw]), which is comparable

with the difference in BrE. The preference of *HAVE to* in speech over writing corresponds to the findings of Collins (2009a:67) for the ICE-data that this quasi-modal has higher frequencies in speech for AmE, BrE and AusE, with a combined average of 1 729 for speech and 699 for writing (with speech being 147,4% more frequent). Strong evidence for this register preference of *HAVE to* is also found in the data of Biber *et al*. (1999:489), where the bulk of the instances of *HAVE to* for AmE and BrE combined occur in conversation.

From the above discussion and from Figure 16 it is clear that the spoken and written patterns for *must, should* and *HAVE to* are much more stable in SAfE than in other varieties, in that all three these auxiliaries have higher frequencies in speech – to extent that, on a purely quantitative level, no shift in weight from one auxiliary to another in speech or writing is apparent. On the other hand, such a shift is very clear in AmE, where *must* appears to be more restricted to writing, with the obvious beneficiaries being both *should* and *HAVE to* with their higher frequencies in speech. AmE is indeed the variety in this case with the most pronounced preference for a particular auxiliary in a particular register, with larger frequency gaps between spoken and written frequencies for all three the auxiliaries under discussion. This tendency has been linked to democratisation in previous discussions on the decline of *must* in AmE, so in terms of the synchronic data, there is support for this process with the apparent avoidance of *must* in speech. With both *must* and *should* having higher written frequencies in BrE, *HAVE to* is the obvious choice among these auxiliaries to be used in spoken BrE. The shift in register preference is however not as conspicuous as for AmE, with the written and spoken frequencies of BrE *must* and *should* being closer together than in AmE. Nevertheless, as far as democratisation can be detected based on the tendency away from *must* in speech in favour of *HAVE to*, some evidence for this process is present for BrE. Despite the fact that such a tendency is not clear from the SAfE data, the issue of whether SAfE experiences democratisation will be discussed in § 4.3.

In the above discussion it was noted that SAfE *must* is neither lower in overall frequency than *should,* nor is it restricted to writing, and nor does it display a clear tendency of a register preference for *must* or its rivals, which is in all three cases very different from the tendencies in other native varieties. The question now needs to be addressed why *must* behaves so oddly in SAfE. As the first, quantitative point of investigation, a comparison with the tendencies in Afrikaans is provided below.

4.2.2.3 Results for contemporary Afrikaans

The influence of Afrikaans on SAfE has already been established on a social level in § 2.3.3.4, where the factors of migration, contact and identity were shown to shape the development of SAfE, with contact surfacing as an especially prominent factor. On a sociopolitical level, the complexity of the relationship between the Afrikaners and the English-speaking population in South Africa was pointed out a factor unique to the country. To recapitulate, the extensive social and linguistic contact between the Afrikaners and English-speaking South Africans involves the following, as consolidated from the discussion in § 2.3.3.4 (see Hooper, 1951:81-3; Lanham, 1978:157-8; Steyn, 1980:130-1; 178; Mesthrie & West, 1995:115; Branford, 1996:36-9; Giliomee and Mbenga, 2007:206; 433-437; Schneider, 2007:177; Bekker, 2012:136):

(a) The nature of collective and individual relationships with fluctuating surges of antagonism and symbiosis, e.g. war, peace, intermarriage, urbanisation or close contact in smaller rural environments, etc.

(b) Eras of bilingualism on government and individual level during the formation of SAfE.

(c) Mutual social and linguistic influence, most relevantly involving Afrikaans/Dutch input into Cape English after the 1820 settlement, and especially into early SAfE during the mineral revolution of the 1880s.

(d) The maintenance of separate identities, and only a very small extent of language shift among some Afrikaners.

(e) The unique, forced contact situation under apartheid laws based on European descent, i.e. a 'racial affinitive relationship'.

(f) Contact induced by the military situation during the last half of the 20th century.

(g) A degree of hindrance of the Founder Effect by the continual influx of new British settlers (three separate *tabula rasa* contexts), i.e. SAfE more being a product of its environment than of its original heritage.

(h) Scholastic English, particularly in white communities, being commonly taught by Afrikaans-speaking teachers over the 20th century.

It was also mentioned in § 1.1.1 and § 2.4.3.1 that there is some evidence in the literature that the influence of Afrikaans on SAfE has moved beyond lexical means and into the domain of grammar and semantics (cf. Jeffery and Van Rooy, 2004; Schneider, 2007:184), but that that this influence needs further investigation. This section will investigate the possibility of quantitative support for the influence of

Afrikaans modal and quasi-modal frequencies on those of SAfE, especially regarding *moet/moes* and *must*, in the hope of providing such evidence on a quantitative level. The possibility of such influence is reinforced by the fact that the modal systems of Afrikaans and English are related, e.g. the cognates *moet/moes* and *must* share etymological and hence grammatical, broad semantic and phonological characteristics (cf. § 2.4.3.1; cf. Abraham, 2001; Conradie, 1976; Traugott, 2006). It was mentioned that *moet* is the central and unmarked/neutral modal verb in Afrikaans regarding commands or requests (Ponelis, 1979:251), which, in English, is often done without a modal in plain imperative form. Being more 'neutral', the root meaning of Afrikaans *moet* can convey both stronger and weaker deontic meanings, but is often weaker, even often dynamic (the past form *moes* mostly conveys weaker, i.e. median modal meanings). It was also noted that *should* is not semantically equivalent to its structural cognate, *sou*, but that its meaning is closer to the Afrikaans quasi-modal construction *behoort te*, which also expresses a similar meaning to *ought*, whereas quasi-modal *HAVE to* has no Afrikaans cognate. Furthermore it was shown that *moet*/*moes* have the most semantic equivalents in English, including both modals and quasi-modals (e.g. *HAVE to*). The modal options in Afrikaans are therefore more limited than in English, especially regarding the expression of obligation and necessity.

Table 7 presents the frequencies of the Afrikaans equivalents of some English modals and quasi-modals, according to semantic and not always structural correspondence (Ponelis 1979:246-247), as also discussed in § 2.4.3.1 and indicated in Table 1. The discussion will highlight some general, noteworthy frequencies, before turning its attention to the modal cognate pairs in the obligation and necessity cluster. Afrikaans examples will be reserved for the semantic discussion presented for *moet/moes* in § 4.3.2.2.3.

In terms of overall frequencies, the most frequent modal in Afrikaans is *kan*, the cognate form of *can*, with a higher frequency in speech than in writing, which is similar to the case of SAfE *can* in the synchronic data (see Table 6.1). Second in overall frequency is the present/past form pair *moet/moes*, being even more frequent than *sal*, which is cognate with *will/shall*. *Sal* has a higher frequency in writing, which corresponds to the register distribution of *will* and *shall* in SAfE, as reported for the ICE-SA data in Table 6.1.

Table 7[172]

Afrikaans modal and quasi-modal frequencies per 100k words with English semantic equivalents

English semantic equivalent	Afrikaans auxiliary	Written	Spoken	AVERAGE
should/ought	***behoort te***	103,7	2,9	53,3
BE going to	***gaan***	157,3	393	275,2
need/NEED to	***hoef te***	9	15,3	12,2
can	***kan***	609,2	746,8	678
could	***kon***	85,5	65,3	75,4
may/might	***mag***	37,6	43,5	40,6
must/NEED to/HAVE to/(HAVE) got to	***moet/moes***	431,6	514,9	473,3
will/shall	***sal***	403,2	326,3	364,8
would	***sou***	88,7	139,2	114
WANT to/WANTED to	***wil/wou***	148,8	301,7	225,3
AVERAGE		207,5	254,9	231,2

In view of the many semantic equivalents of *moet/moes* in English, along with its property as the unmarked option for expressing requests or commands with a range of possible degrees of root meanings, the high frequency of this modal pair is hardly surprising, considering the amount of work it has to do in Afrikaans. *Moet/moes* is more frequent by 19,3% in spoken Afrikaans than in written Afrikaans. This agrees with the pattern of register distribution for SAfE *must*, where it is 14,8% more frequent in speech than in writing, even if Afrikaans spoken *moet/moes* is slightly more frequent over written *moet/moes* than SAfE spoken *must* is over written *must*. On the whole, the very high overall frequency of *moet/moes* in Afrikaans and especially its high frequency in speech are quantitative patterns that correspond with those of SAfE *must*. This kind of overall and register-based correspondence supports the idea of a quantitative relationship and hence a degree of influence of *moet/moes* on *must*. This idea is reinforced by the much higher overall frequency of *moet/moes* (473,3) than *must* (125,7), where the high frequency of *moet/moes* can be regarded as

[172] The data presented in Table 7 was kindly extracted and normalised by Bertus van Rooy, who has copyrighted access to the Afrikaans corpora. My thanks to him in this regard.

having a 'dragging' effect on the frequency of *must*. It must however be remembered that the *moet/moes* pair not only expresses the broad meaning of *must* alone, but also that of *HAVE to, (HAVE) got to* and *NEED to*.

Apart from high-frequency Afrikaans modals *kan, moet/moes* and *sal*, the rest of the auxiliaries in order of overall frequency are *wil/wou, sou, kon, behoort te, mag* and *hoef te*. It is interesting to note the overall order of frequency of those auxiliaries in Afrikaans that are equivalent to the modals and quasi-modals of obligation and necessity in English, viz. *moet/moes*, *behoort te* and *hoef te.* When comparing the frequencies of these three Afrikaans auxiliaries to those of the whole group of semantic equivalent modals and quasi-modals in SAfE in each case (three groups in all), the following picture, presented in Figure17, arises across speech and writing.

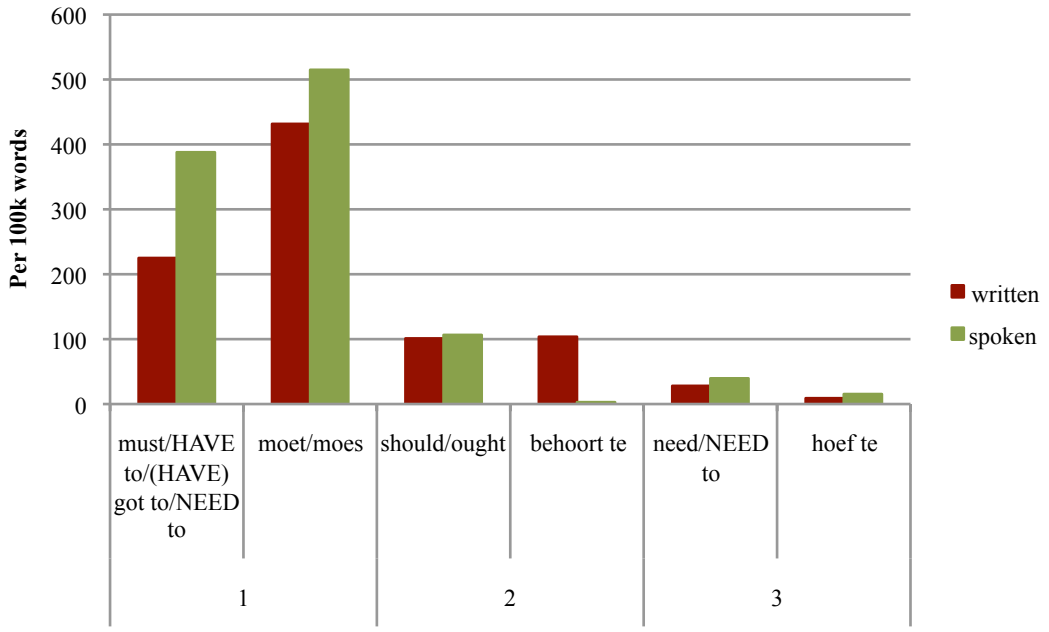

Figure 17

The spoken and written frequencies of semantic equivalent modals and quasi-modals in SAfE and Afrikaans in the obligation and necessity cluster

On a purely quantitative level, the possibility of the above-mentioned 'dragging' effect or 'pull' of the frequency of *moet/moes* on the frequency of its semantic equivalents in the first group, in terms of overall frequency and register distribution, is evident when the data is presented in clustered column form. This effect is further supported by the much lower frequencies of the Afrikaans and SAfE semantic

equivalents in the other two groups. Such quantitative influence can be argued to be compounded in the case of *must*, due to its additional grammatical and phonological correspondence with *moet/moes*, which the other semantic equivalents in the first group generally lack, being quasi-modals. The very fact that the three other semantic equivalents of *moet/moes* in the first group, namely *HAVE to, (HAVE) got to* and *NEED to*, are quasi-modals, and perhaps especially the presence of *HAVE to* and *(HAVE) got to* (see Table 6.2), is responsible for the high frequency of this group in spoken language. When the discussion is restricted to only *must*, the above-mentioned percentage by which it is more frequent in speech is more like the pattern for *moet/moes* in Figure 17.

By this quantitative account, contact with Afrikaans emerges as a plausible factor in the uniqueness of SAfE *must*. It would indeed serve as further evidence if SAfE *must* had taken on a meaning similar to that of Afrikaans *moet/moes*, i.e. neutral/weaker (median degree) obligation meanings. If indeed *must* could be found to increasingly express such meanings, evidence for Afrikaans influence would certainly be warranted on a qualitative level as well, seeing that social influence and quantitative influence have already been argued for.

4.2.2.4 Summary: synchronic findings

The first goal of this section was to establish, with the addition of spoken data, whether colloquialisation occurs in SAfE. As far as a higher synchronic frequency of quasi-modals in speech than in writing combined with a diachronic increase in written quasi-modals provides evidence of the growing popularity of speech-like styles, a degree of colloquialisation appears to be at work in SAfE over the 20th century, but since the diachronic increase is not as prominent as in other native varieties, the extent of this is limited. Based on quasi-modal ratios in speech and writing alone, however, SAfE appears to be more advanced than BrE and AusE in the colloquialisation process, but more conservative than AmE.

The second goal of the synchronic discussion was to explore the contemporary frequencies of the individual modals and quasi-modals in speech and writing, especially *must, should* and *HAVE to*, and compare the findings with the results for other native varieties. In the possibility, permission and ability cluster the high spoken

frequency of *can* and the overall low frequency of *might* were noteworthy, and in the prediction and volition/cluster the high written frequency of *will* and high spoken frequency of *BE going to* was noted to be indicative of a quantitative trade-off in favour of a quasi-modal in speech, which pointed to a degree of colloquialisation on an individual modal level. The most noteworthy findings however arose for *must, should* and *HAVE to* within the obligation and necessity cluster, apart from the restriction of formalised *BE to* to written contexts, which also indicated a degree of colloquialisation. The high overall frequency of *must* over its rival *should* was seen very noteworthy when compared to the reversed situation in all four other native varieties in Figure 15. SAfE *must* was indeed found to also be overall more frequent than most mid-frequency modals and quasi-modals that enjoyed higher frequencies in other varieties. Overall SAfE *HAVE to* only outranks *must* slightly because of its high frequency in speech, and indeed *HAVE to* is the only auxiliary in this semantic cluster to outrank *must* in its high spoken frequency, which is very noteworthy considering its apparent restriction to written contexts in AmE and BrE. Additionally, *should* was found to be only slightly more frequent in speech. Because *must*, *should* and *HAVE to* all have higher spoken frequencies in SAfE, no shift in weight from one auxiliary to another in speech or writing is apparent, whereas such a shift, which is linked to democratisation, is very clear in AmE, where the restriction of *must* to writing benefits *should* and *HAVE to* in the spoken register, and in BrE, where the restriction of *must* and *should* to writing benefits *HAVE to* in speech. Democratisation in SAfE will be explored in the section on semantic results.

Seeing that Afrikaans influence on SAfE had already been established on a sociohistorical level in Chapter 2, the third goal was to compare the results of the Afrikaans data with the SAfE data to determine whether there is evidence of a possible quantitative influence of Afrikaans on this variety, especially on *must*, given its high usage preference in SAfE. With the modal cognate pair of *must*, namely *moet/moes*, ranking second in overall frequency, and also second in specifically spoken frequency, among all the modals and quasi-modals with cognates in English, evidence of such a quantitative influence was suggested. This was considered as being reinforced by the very low overall frequency of *behoort te*, the semantic equivalent of *should/ought*. Whether qualitative evidence of such an influence is present will be explored in § 4.3.

4.2.3 Conclusion: general results

The general results section of this chapter focussed on both diachronic and synchronic findings on the levels of overall and individual frequencies for modals and quasi-modals. On a diachronic level, comprehensive discussions of frequency trends, with different degrees of stability, decelerations or accelerations, were conducted in terms of each of the three semantic clusters, three frequency groups and four registers, in clarification of the general and specific results presented in Figures 5 and 6, and Tables 5.1 and 5.2. Similarities and differences with the overall and individual trends of AmE and BrE were pointed out. Overall, similar trends for modals and quasi-modals occur in SAfE than in BrE and AmE over the 20th century were reported, regarding the decline of modals and the rise of quasi-modals, but the modals in SAfE decrease more and the quasi-modals increase less than in the other native varieties. The most noteworthy difference on an individual modal level involved the 'ecology' of the obligation and necessity group, where *must* was shown not to decline significantly over the 20th century, becoming the highest-frequency modal in its semantic cluster by the 1990s, whereas this modal is in serious decline in other native varieties, and is exceeded in frequency by both *should* and *HAVE to*. The synchronic distributional pattern for *must* was also found to be noteworthy, where it was reported to have a higher frequency in spoken than written language, and an overall higher frequency than *should*, whereas *must* in other native varieties is overall less frequent than *should* and *HAVE to*, and moving to a degree of restriction in written contexts. Quantitative support for the influence of the Afrikaans modal pair *moet/moes*, which is cognate with *must*, was provided, in anticipation of qualitative evidence for this influence, which will shed more light on the nature and progression of the democratisation process in SAfE.

Apart from democratisation, some evidence of the colloquialisation process was found in 20th-century SAfE, especially regarding the higher synchronic frequency of quasi-modals in speech and in the trends of some individual modals and quasi-modals, e.g. *shall* and *BE going to* (diachronically) and *will* and *BE going to* (synchronically). Indication of this process was further found in the trend toward distributional fragmentation in the diachronic and synchronic trends of certain individual auxiliaries, where perceived archaic or more formalised options, e.g. *BE to* and *shall*, were shown to decline in frequency and to become restricted to formal,

written contexts, like non-fiction and news. This links with the trend toward monosemy, for which some evidence was found in certain modals, e.g. in *may* and *can*, which are respectively increasing in their epistemic and dynamic possibility meanings over the 20th century, as well as *shall*'s, which is becoming restricted to dynamic volition meanings. Along the same lines, some low-frequency modals were found to experiencing paradigmatic atrophy, just as they are in AmE and BrE: *shall* becomes virtually restricted to first-person subjects and *need* is restricted to use in negative polarity. In the other native varieties the trend toward monosemy is also noted for *must*, *should* and *have to*, which is indeed postulated as the motor behind the changing frequency relations of these auxiliaries over the 20th century. The next section will focus on the macro- and microsemantics of these three prominent role players in the obligation and necessity cluster in SAfE, in order to gain a full sense of the implications of the frequency trends described in the general results section.

4.3 SEMANTIC RESULTS OF THE MAJOR CONTENDERS IN THE OBLIGATION AND NECESSITY CLUSTER

This section presents and discusses the results of the diachronic macrosemantic and microsemantic analyses done on all the tokens of *must, should* and *HAVE to*, as fully described in terms of the parameters for analysis in § 3.4.1. Each of these three auxiliaries are treated individually and the discussions are divided into two sections, the first of which considers macrosemantics, i.e. deontic, dynamic and epistemic frequencies (plus the quasi-subjunctive and preterite uses for *should*), and the second of which discussing microsemantics in terms of both modal strength, i.e. the high and median degrees or formulaic uses of obligation, and the relationship between the subjective and objective source of obligation and the degree of obligation (strength/force).

4.3.1 Introduction

The macrosemantic results for *must, should* and *HAVE to* presented in this section will be compared to the semantic findings of Collins (2009a:33-89) for the obligation and necessity cluster. To recapitulate from Chapter 2, it was noted in § 2.4.2.3 that for contemporary AmE, BrE and AusE combined *must* mostly conveys deontic (57,3%) and epistemic (32%) meanings, followed by dynamic (6,3%) and indeterminate meanings (3,6%). No clear, direct trend toward monosemy is therefore detected for *must*, but rather a restriction to both epistemic and strong deontic meanings, i.e. it expresses dual meanings – a more restricted kind of polysemy in a sense (§ 2.4.2.2). *Should*, on the other hand conveys deontic meanings more often than *must* (69%) and consequently convey epistemic meanings in fewer cases than *must* (11,8%) (which corresponds to the tendency for *must* to be used more frequently in epistemic contexts [cf. Smith, 2003:242; Leech *et al.,* 2009:114-116]), whereas *should* also conveys meanings that are not semantic options for *must*, viz. the (putative) quasi-subjunctive meaning (6,6%) and the preterite meaning (where *should* = *would*) (2,7%) (indeterminate meanings cover 10%) (cf. Leech *et al*., 2009:86). The fact that deontic *should* (2 148 tokens [pmw]) is more than double the frequency of deontic *must* (969 tokens [pmw]) (Collins, 2009:35;45) was emphasised in § 2.4.2.3. Most of the deontic meanings of *must* are found in BrE, whereas most of the deontic meanings of *should* occur in AusE. Quasi-modal *HAVE to* on the other hand conveys even more deontic meanings (79,1%) than *should* in the three varieties combined (2 999 tokens [pmw]), but also expresses dynamic meanings (22,4%) and very few epistemic meanings (0,7%) (0,8% is indeterminate). The findings of Collins (2009a) therefore support the trend that the epistemic 'weight' in the obligation and necessity cluster is shifting toward *must* and, consequently, (root) deontic meanings have become more likely in *HAVE to* (Leech *et al.*, 2009:87-9). Whether the same trend is noticeable in the ecology of the obligation and necessity cluster in SAfE remains to be seen, as the sections below will explore. The raw and normalised results of the semantic analyses for this section can be found in Appendix 5.

4.3.2 *Must*

From both the general quantitative diachronic and synchronic findings concerning the obligation/necessity cluster in SAfE, the case of *must* was found to be remarkable among other native Englishes, in that it is not in statistically significant decline and enjoys a higher frequency in speech than in writing. This section seeks to explore the semantics of SAfE *must* diachronically, in order to clarify why this might be the case. The macrosemantic results for *must* will receive attention firstly, where the deontic, dynamic and epistemic frequencies will be discussed in 4.3.2.1, followed by the discussion on the microsemantics of the deontic category, in terms of high- or median-degree obligation or formulaic uses in § 4.3.2.2.

4.3.2.1 Macrosemantics of *must*

The normalised diachronic macrosemantic trends for *must* in SAfE that arise from the data in Table 29.2 in Appendix 5 are presented in Figure 18.

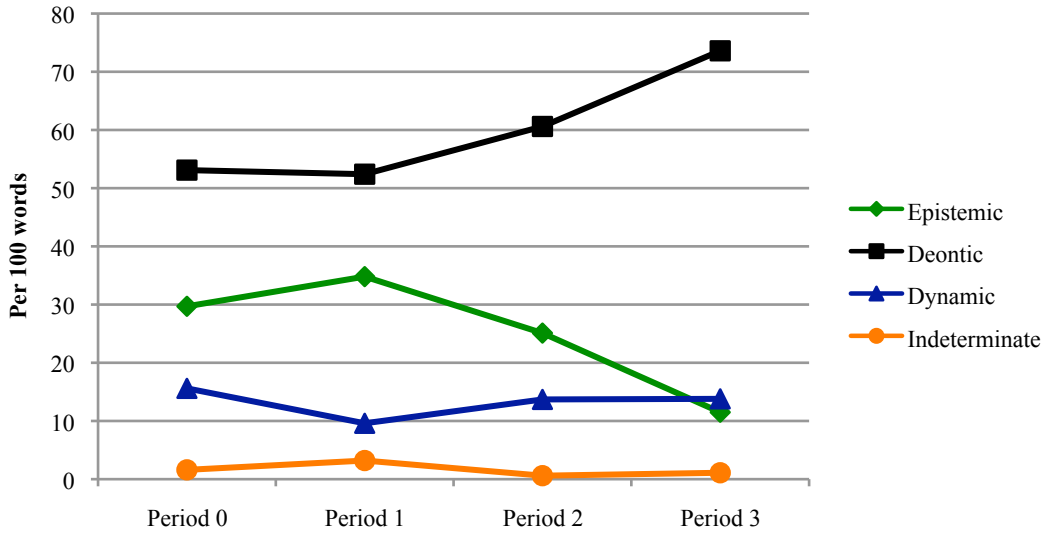

Figure 18

Macrosemantic frequency changes of *must* in SAfE over time

From the above figure the frequency of deontic *must* is striking. Not only does it maintain the highest frequency over all periods, but its upward trends between Period

1 and 2, as well as between Period 2 and 3, are also noteworthy, together with the decrease visible for epistemic meanings from Period 1 all the way to Period 3. It was noted in § 2.4.2.2 that in 20th-century AmE and BrE *must* has become restricted to strong deontic meaning and has declined as a result of this restriction, and apart from this has also been moving toward expressing more epistemic meanings (Smith, 2003:242; Leech *et al.*, 2009:114-116; Millar, 2009:203). Since the frequency of epistemic *must* in SAfE is in decline, the same move toward the expression of dual meanings is clearly not present as for the other native varieties. The fact that the deontic meaning rises in SAfE is not evidence on its own that it is becoming restricted to strong deontic uses, i.e. although Figure 18 suggests a trend toward a macro level of monosemy for SAfE *must*, the fact that a microsemantic trend toward a particular degree of obligation was mentioned for the other native varieties shows that this needs consideration in SAfE. The discussion on the degrees of obligative strength of *must* will clarify which microsemantic trend is discernible in this respect.

For AmE Millar (2009:203) indeed reports that *must* moves away from root meanings in TIME over the period spanning the 1920s to 1960s and into the 2000s to make place for epistemic meanings, even though the root meaning of *must* remains the most frequent overall. Where the 1920s to 1960s can be regarded as more or less representative of my Period 2 and the 2000s of the period just after my Period 3, the fact that the rise in deontic meanings and the drop in epistemic meanings in SAfE occur even since Period 1 and continues to Period 3 shows that opposite trends are present for *must* in SAfE and AmE in TIME over the same time span. Since root meanings include the dynamic category of meaning on top of the deontic, the fact that the dynamic frequency of *must* in SAfE remains stable from Period 2 to 3 (13,7% and 13,8% respectively [see Table 29.2 Appendix 5]) also does not agree with the trend in TIME for root meanings to decline. Even more interesting in this respect is the slight dominance of dynamic meanings over epistemic meanings by Period 3.

The opposing semantic trends for deontic and epistemic meanings of *must* (a rising and declining respectively) illustrated in Figure 18 appear to be the main reason behind the relative stability of mid-frequency modal *must* across the 20th century compared to other native varieties, as noted in § 4.2.1.4.1 and § 4.2.1.5. Deontic meanings increase by 15,6% from Period 1 to 2 and accelerate as they increase by a larger 21,5% from Period 2 to 3. On the other hand epistemic meanings decline by 27,9% from Period 1 to 2 and also accelerate as they decrease by 54,2% from Period

2 to 3. Despite the degree of stability the diachronic frequency of *must* gains between Period 2 and 3, it does decline non-significantly, as reported in § 4.2.1.4.1. The slight decline of *must* over the 20th century can therefore be attributed to the influence of the decline of epistemic meanings during this period. This is motivated firstly by the fact that epistemic meanings decline more than deontic meanings increase (albeit not so much as to hamper the degree of stability gained by the opposing trend), secondly by the acceleration of the declining epistemic trend in the 20th century, and thirdly by the stability of dynamic meanings during this era, in that they do not affect the influence of the above trends.

By Period 3, the semantic ecology of *must* comprises 73,6% deontic, 13,8% dynamic, 11,5% epistemic and 1,1% indeterminate meanings. When compared to the findings of Collins (2009a:34) that *must* in AmE, BrE and AusE combined is divided between 57,3% deontic, 32,8% epistemic, 6,3% dynamic and 3,6% indeterminate meanings, the much lower percentage of epistemic meanings, as well as much higher percentages of deontic and dynamic uses in SAfE are apparent.

Figure 19 shows the change in macrosemantic frequency within each of the four registers over the 20th century, i.e. from Period 2 (1910s-1950s) to Period 3 (1990s). It is generated from the data in Table 30.2 in Appendix 5.

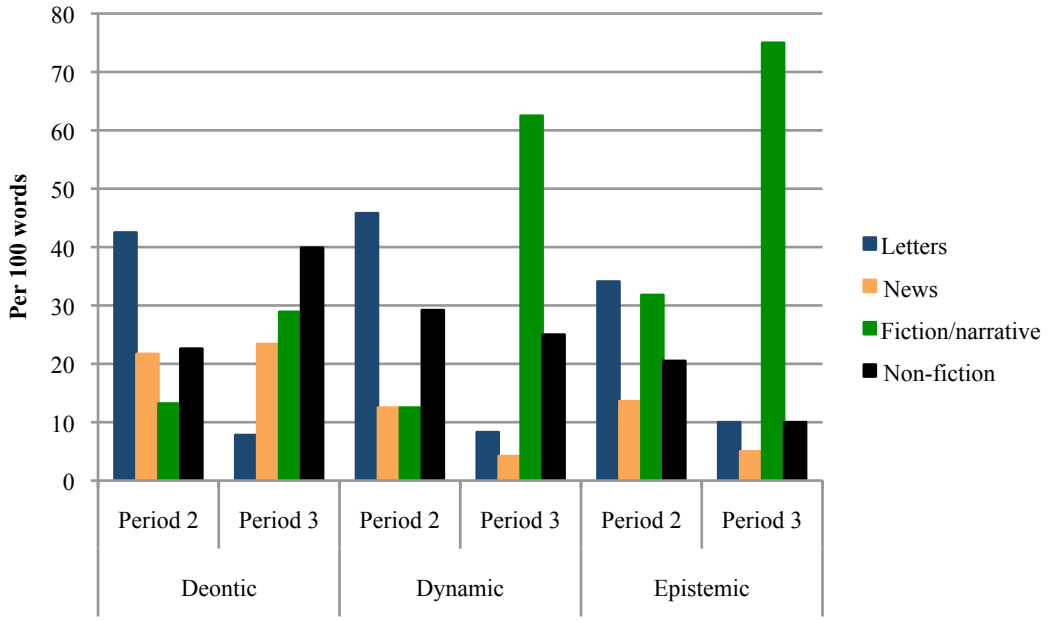

Figure 19

Macrosemantic frequency changes of SAfE *must* per register over the 20th century

The above figure illustrates that deontic *must* decreases in the letter register during the course of the 20th century (-81,6%), but it increases in all three remaining registers: (+7,8%) in news, (+76,5) in non-fiction and by (+118,9%) in fiction/narrative. The fact that deontic *must* increases in three out of the four registers clearly supports the general rising trend of the deontic meaning for *must* from Period 2 to 3, as shown in Figure 18, which fiction/narrative apparently being the main carrier of the trend. While dynamic *must* also decreases in letters (-81,9%), just like deontic *must*, dynamic *must* declines in news (-66,4%) and in non-fiction (-14,4%), unlike deontic *must*. Dynamic *must* however also increases in fiction/narrative (+400%), which is again similar to deontic *must*. The general stability of dynamic *must* from Period 2 to 3, as seen in Figure 18, is supported by the counteractive dynamic trend in fiction/narrative, which increases so much that it balances out the trends of decline in the other three registers. Similar to the tendencies of dynamic *must*, epistemic *must* also decreases in letters (-70,7%), in news (-63,2%) and in non-fiction (-51,2%), but increases in fiction/narrative (+135,8%). Here epistemic *must* in fiction/narrative increases much less than it does in the dynamic meaning, failing to balance out the trends of decline for epistemic *must* in the other three registers. This indeed supports the trend of general decline from Period 2 to 3, regarding epistemic *must*, as also seen in Table 18.

The semantic frequency changes of *must* over the four registers therefore show a very stable pattern for letters and fiction/narrative, since these registers respectively decrease and increase in all macromeanings, but there appears to be a trade-off between news and non-fiction at the expense of the dynamic and epistemic meanings, and in favour of the deontic meaning.

The general 20th-century increase of *must* in fiction/narrative that was noted in § 4.2.1.4.1, which was found to be the only register where *must* increased from Period 2 to 3, is therefore supported by the increase found in all three macromeanings in the fiction/narrative register over the same period of time, especially that of epistemic and dynamic *must*. On the other hand, the general 20th-century decline of *must* in letters (see § 4.2.1.4.1), which displayed the greatest decline of all registers, is also supported by the overall decline found in all three macromeanings of *must* in the letters register over the same period of time. Generally, *must* in both news and non-fiction was shown in § 4.2.1.4.1 to decline to a lesser extent than in letters, which is most likely supported by the trade-off described above, involving that these registers

both increase in the deontic meaning, and both decrease in the dynamic and epistemic meanings. The balance is therefore swayed in favour of a decline for these registers, but the increase in deontic meanings assures that the decline is not great in both cases. The next subsections describe the use of dynamic and epistemic *must*, before attention is drawn to deontic *must* and its micromeanings.

4.3.2.1.1 Dynamic *must*

As Figure 18 shows, the highest percentage of dynamic meanings is observed (15,6%) during Period 0, where it is most often used as part of narration in the fiction/narrative register (156) and also in non-fiction (157) to express a circumstantially (externally) derived need, as well as part of a description of hypothetical circumstances in news (158) (cf. Collins, 2009a:40-1). The slight drop of dynamic meanings in Period 1 (-38,5%), before rising again in Period 2 (+42,7%) and remaining stable in Period 3 (+0,0007%) (as noted above), occurs when fewer uses in narration in the fiction/narrative register are present that express a circumstantially derived need as in (159). The dynamic meanings in the following examples are not sharply distinct from deontic *must*, where it is slightly ambiguous with a less salient obligation meaning that can be "understood to be imposed by a deontic source" (Collins, 2009a:41), where they either have a non-specific subject (as in [156]) or a specific subject (as in [157] and the second use of *must* in [159]), or have a habitual future reading (as in [158]). Despite this degree of gradience that is a characteristic of modal semantics (§ 2.4.1.1), these examples were classed as dynamic, since the more salient reading involved necessity rather than obligation.

(156) The country looks magnificent; white resplendent quartz fields as far as the eye can reach, relieved by dark granite masses, interspersed with baobab trees 80 feet high, and occasionally enlivened by the appearance of ...the elephant, the buffalo, the rhinosceros, the giraffe, and the eland... the whole presents a scene of grandeur that **must** be seen to be fully realised. (1860s; Fiction/narrative)

(157) ...it may be remarked, that there seems to be something in a virgin soil unfavourable to the support of human life; and it seems to be with men as with

vegetables - they **must** suffer, after being transplanted, before they can take root. (1820s; Non-fiction)

(158) ...if there be a time more favourable than another, at which a reasonable being would feel disposed to look back with reflection on the past, it **must** be at the commencement of a new season, when another year has been spared to his existence, another period added to that account which he must sooner or later be called upon to render. (1820s; News)

(159) So far matters seemed promising enough ; but one point was certain, that everything I might want **must** be taken from Cape Town, as nothing whatever but oxen could be bought where the Missionaries were. Servants, waggons, and things of every kind I **must** take with me, for the ship would land me on the desert sands... (1880s; Fiction/narrative)

Dynamic meanings in Period 3 also often express externally derived needs (160) similar to the uses exemplified above, but there are almost just as many uses where *must* expresses "an internal need in the subject referent" (Collins, 2009a:40), as in (161), (162) and (163), a meaning that is expressed relatively rarely in AmE, BrE and AusE (2009a:41). Once again a degree of ambiguity with the deontic meaning is present, especially in (161) and (162), but the necessity reading is more salient. This suggests that the more frequent expression of not only external necessity, but also internal necessity, motivates the more frequent occurrence of dynamic *must* in SAfE than in other native varieties. Figure 19 illustrated that by Period 3 dynamic *must* is used most frequently in the fiction/narrative register, followed by the non-fiction register. Hence, examples (161) and (162) are respectively from narration (from a first-person narrator) and speech representation in the fiction/narrative register, whereas (163) is from non-fiction.

(160) ...on the border of South Africa and Mozambique, farmers **must** work sandy soils that are poor in nutrients and do not hold water for long. (1990s; Non-fiction [W2D-016])

(161) I wanted to rush out of the window and pick it up. Was it the photo? – I **must** know. The conductor bent an [sic] picked it up. (1990s; Fiction/narrative [W2F-013])

(162) "I will find your child. Then we will deal with these shadow Chinese, but I **must** move fast." (1990s; Fiction/narrative [W2F-009])

(163) ...the male spins a special mating thread on to which he **must** lure the female with more plucking and tweaking movements (1990s; Non-fiction [W2B-023])

4.3.2.1.2 Epistemic *must*

The rise of epistemic *must* from Period 0 to 1 (+17,2%) appears to go unhampered by the stability in the frequency of deontic meanings across the same time span. In Period 1 epistemic *must* expresses objective, logical certainty ("a logical necessity based on what is known" [Collins, 2009a:38]), as in (164), but more often conveys a speaker's subjective inference of what is certain, (often accompanied by harmonic expressions such as *certainly* and *really* or hedges such as *I think* [cf. Coates, 1983]), as in (165), (166) and (167). The following examples illustrate the use of epistemic *must* in Period 1.

(164) Surprise and gratitude **must** have produced a touch of effusiveness which jarred on him; for, to the eager exclamation and thanks, he made no answer – just moved on... (1900s; Fiction/narrative)

(165) ...showing him a beautiful cast of an encrinite two or three inches long, the most perfect representation of a rusty screw that I ever saw. "That **must** certainly be a screw of the ark!" said he... (1870s; Fiction/narrative)

(166) ...I have since you last heard of me been pursuing my career in what I really think **must** be the most barren and desolate place in the world – Namaqualand[173]. (1870s; Social letters)

[173] A region surrounding the Orange River, spanning part of the Northern Cape and Namibia (see Figure 1 in § 2.3.3.1).

(167) But before I could speak, he said: 'It **must** be very beautiful down there. Can you always see the mountain?' (1900s; Fiction/narrative)

These meanings become less frequent for *must* in Period 2 (-27,9%), as well as into Period 3 (-54,2%), but, as will be seen in § 4.3.3.1.1, epistemic *should* rises over the same periods. The implication of this apparent trade-off between *must* and *should* regarding the epistemic meaning is that of an opposite trend for SAfE than the trend for other native varieties, where *must* is reported to proportionally express more and *should* less epistemic meanings over the course of the 20th century (cf. § 2.4.2.2; Leech *et al.*, 2009:87-9; see also Collins, 2009a:34;45). The next section focuses on deontic *must* and its micromeanings.

4.3.2.2 Microsemantics of deontic *must*

The increase in deontic meanings for *must* from Period 1 through to 3, after its stability from Period 0 to 1, is supported by microsemantic changes relating to the strength of obligation. The stacked column form used for Figure 20 illustrates the proportional differences in deontic frequency between high and median degrees of obligation, as well as formulaic uses. Where examples are discussed, those of the six parameters that the instance of either high or median obligation adheres to are marked by means of the letter 'P' plus the number of the parameter, as in 'P1' to signify 'parameter 1' (as also stipulated in § 3.4.1.1). This will both indicate the application of the parameters in microsemantic interpretations and serve as a means to evaluate their validity and productivity for such analyses.

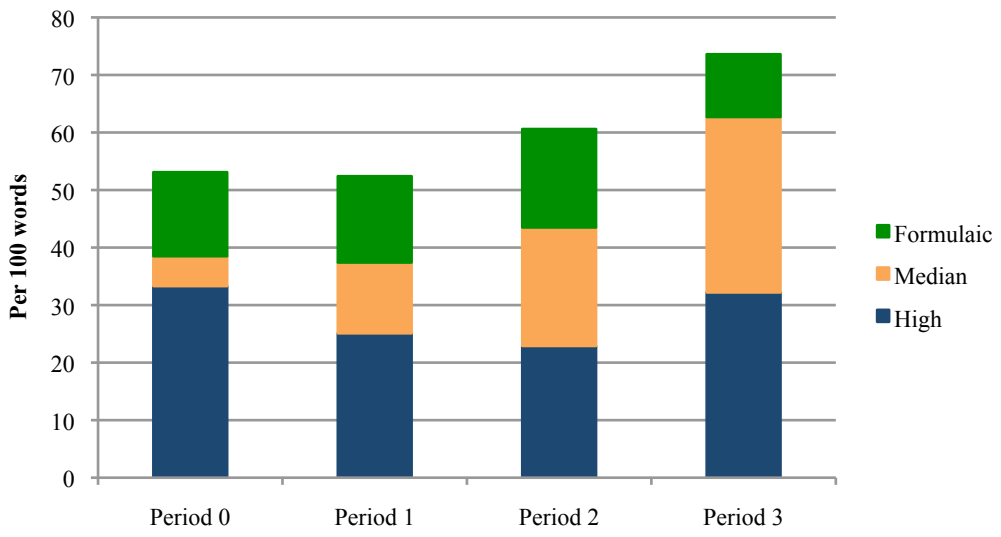

Figure 20

Microsemantic frequency changes in SAfE deontic *must* over time

The rising trend of deontic *must*, noted in the previous section, is seen in Figure 20 to be mainly influenced by the gradual increase of median degrees of obligation over the course of the SAfE data, but fluctuations between the frequencies of strong obligation and formulaic uses also influence the deontic trend. The changing relationship between the frequencies of high and median obligation uses will firstly be considered as the main microsemantic finding of this thesis, where after formulaic uses will be discussed.

4.3.2.2.1 *Must*: high and median degrees of obligation

The data in Table 33.2 in Appendix 5 show that, after the high degree use is the dominant meaning of *must* in Period 0 (with 33,3 normalised instances) over the median degree use (5,2 normalised instances), with a ratio of 6,4:1 high relative to median uses, this ratio narrows considerably to 2:1 when the high degree uses decrease to 25,1 normalised instances and the median degree uses increase to 12,3 normalised instances in Period 1. The relative proportions of high and median obligation uses in Period 2 (respectively 22,9 and 20,6 normalised instances) are basically equal in Period 2 (with a ratio of 1,1:1 high to median uses) and maintain these proportions into Period 3 with the exact same ratio of 1,1:1 (with 32,2 and 30,5 normalised instances for high and median uses respectively), and hence acquiring

stability with very similar frequencies among these uses over the course of the 20th century.

If *must* in other native varieties was found to become restricted to strong (high degree) obligative uses in the 20th century, then *must* in SAfE did not follow the same trend, even if the high degree uses increased from Period 2 to 3, but because the median uses in SAfE also increased with equal degree over the same periods, *must* is not more restricted in the microsemantics of its deontic use, but quite the opposite. Thus, one part of Bowerman's (2004b:477) first anecdotal claim concerning modality (noted in § 1.1.3), namely that *must* has less social impact, i.e. less illocutionary force, in SAfE than in other Englishes, is supported by the empirical findings drawn from the real data in this study. It largely remains to be seen in § 4.3.3 whether the second part of his claim, that *must* is often used as a substitute for polite *should*, can also be supported, but investigations into this aspect will also be conducted in the current section.

In Period 0 the more frequent expression of a high degree of obligation for *must* is unsurprising, seeing that by this time of input (1820s-1860s) proto-SAfE was made up of the various dialects of 19th-century British English, for which the standard meaning of *must* at least was either that of strong deontic obligation or epistemic necessity (cf. § 2.4.1.1; Funk, 1898). The instances of a high degree of obligative *must* were further analysed for a subjective or objective source, the results of which are presented in Table 8. Indeterminate cases denote those instances of deontic *must* that are "ambivalently subjective/objective, with uncertainty as to the deontic source" (Collins, 2009a:36).

It was noted in § 2.4.1.1 that the prototypical or core examples of *must* (in the terminology of respectively prototype theory and fuzzy set theory) have a high degree of obligation with a subjective source, but that these are not the more frequent uses in English (cf. Coates, 1983:33); Collins (2009a:37-8) for instance reports that for the native varieties in his study objective deontic *must* tends to be used more frequently than subjective deontic *must*, but that objective uses more often coincide with weaker (median) obligation. In another place Collins suggests that this tendency is not always clear: "there is no necessary connections between subjectivity and the use of *must*" (2009a:35) – hence there is not much clarity as to the exact tendency of native Englishes, although the traditional view is that of the former report of Collins (cf. § 2.4.1.1; Depraetere and Verhulst, 2008:3).

Table 8

Percentages of high-degree deontic uses with subjective, objective or indeterminate sources of obligation in *must*

Period	Subjective	Objective	Indeterminate
0	64	36	0
1	38,3	61,7	0
2	35	57,5	7,5
3	30,4	66,1	3,5
Average	41,9	55,3	5,5

Although Collins' analysis found objective uses to be more frequent and often used in combination with weaker meanings, this of course does not imply that the weaker meaning is more frequent for *must*, seeing that other studies, as noted above, have reported stronger degree *must* to become more frequent. As far as SAfE goes, Table 8 shows that *must* with a high degree of obligation is more often paired with an objective source from Period 1 onward, but the difference between the average scores for subjective and objective high-degree uses over all periods is not big. The ratio for subjective:objective uses with high-degree *must* basically reverses from 1,8:1 in Period 0 to 1:1,6 in Period 1, before stability is gained in Period 2 with the same ratio as in Period 0, namely 1:1,6, and the increase of objective meanings in Period 3 result in a final ratio of 1:2,2.

The occurrences of high-degree obligative *must* with a subjective source are more frequent in **Period 0**, i.e. more prototypical examples are present. Examples (168), (169), (170) and (171) exemplify the prototypical use of *must* during the period of input.

(168) Kruger insisted that I **must** instantly depart, and assured me that he would keep my false friends back. ... Kruger pointed to an opening between two mountains as the course I **must** take, and directed his brother to lead us through the river. (1860s; Fiction/narrative)

(169) Moshesh is himself, we think, too shrewd a man to make war from the sheer love of disturbing the peace of the country. That we believe, but we fear he is comparatively impotent in restraining the various Kraals under him; and we

are forced to repeat that we **must** not in considering this question, lose sight, for a moment, of the maxim, that a Chief is responsible for the conduct of his subjects. (1850s; News)

(170) At the same time after making every concession, and offering every shade of explanation of which we are really capable, with our present amount of information, we have come to this deliberate conclusion, that we **MUST** have PEACE amongst the tribes to the Eastward,–that that peace is at present on an insecure footing-that a terrible responsibility rests on the heads of those who would allure over the real circumstances of the country, and that unless Peace, universal throughout the Sovereignty, be maintained, at whatever cost or trouble, this magnificent country will not be developed in the manner, nor to the extent, contemplated by Sir Harry Smith, when he proclaimed the Sovereignty of the Queen of England over it; and we assert without the fear of just contradiction, that Christianity and civilization will make little or no progress north of the Orange River, unless PEACE be placed on a permanent basis HOW momentous then, the issues of the present negotiations! (1850s; News)

(171) Discussions have arisen periodically in the City Council, the past year or two... In times of more than ordinary drought their tone was serious. "Something **must** be done," was the prevailing sentiment... (1860s; News)

In example (168) *Kruger*, the fictional subjective source of the obligation, refers to the historical figure President Paul Kruger, whom the narrator here – by means of reported speech – depicts giving direct orders in the midst of war to one of his officers. Some of the parameters for a high degree of obligation present in this example include that there is an indication of status and authority in the source, and that the first instance of *must* is used with the agentive verb *insisted* (P1), which conveys a high degree of intensity, and the adjunct *instantly* (P2), which conveys a sense of urgency. On the other hand, expressions similar to high-intensity adjuncts (P2) are found in (169), e.g. *we are forced to repeat that* and *for a moment*. In example (170) *must* is capitalised for emphasis in a much embedded sentence in a news article, which, along with the expression *deliberate conclusion*, and the emotional tone (P5) projected by capitalising other words such as *peace* and ending

the sentence with an exclamation, renders it an expression of a high degree of obligation. The reference to the serious tone with which the obligation in (171) was conveyed by the members of the city council (subjective sources), as well as the verb *prevailed*, also indicate a high degree of obligation.

The frequency of subjective and objective sources in instances of median obligation is given in Table 9.

Table 9

Percentages of median-degree deontic uses with subjective, objective or indeterminate sources of obligation in *must*

Period	Subjective	Objective	Indeterminate
0	30	70	0
1	60,9	39,1	0
2	50	50	0
3	35,9	54,7	9,4
Average	44,2	53,5	2,4

The average scores of subjective and objective source pairings with high- and median-degree obligation over all periods, as seen in Tables 8 and 9, are not unalike, as are the percentages for these pairings by Period 3, which corresponds to Huddleston and Pullum's (2002) suggestion that English *must* tends to have both subjective and objective meanings. This suggests that there is no clear tendency for high-degree uses of *must* to more frequently have a subjective source of obligation or for median-degree *must* to have an objective source of obligation in SAfE, as is generally suggested for native varieties of English by Collins (2009a:38). Thus, for SAfE no clear link between the subjective or objective source with a high or median degree of obligative force is present, except during the period of input, which hosts more prototypical uses of high-degree *must,* and reciprocally so, median uses strongly tend toward the objective end of the continuum during this era. The traditional view therefore only holds for Period 0. The less-frequent objective use with high obligation *must* is illustrated in (172).

(172) The men, the women, and the children used to eat in separate groups; but woe to the wight that was missing at supper time! No food would be saved or put away for the use of the late comer, who, when arriving after time, could never hope to make up his loss, but **must** go supperless to his couch (mat), and learn to be a more careful observer of punctuality for the future. (1830s; Fiction/narrative)

In example (172) the source of obligation is seated in a consequence based on an external influence derived from circumstances, but the obligation has a high degree of obligative force, precisely because the compulsion or obligation is derived from the consequences of non-compliance (P5). Non-compliance with the punishment will however also have a high degree of gravity and is not really possible in the context (P5). Here some gradience is noticeable between the interpretations of an objective source and a subjective source of obligation, since it is implied that the consequences of being late were previously laid down by a subjective agent, but since the directive or performative quality of such an issue of punishment is absent, the source of obligation tends toward the objective end of the continuum. The interpretation of a high degree of obligation is supported by the dramatic phrasing in *woe to the wight that was missing at supper time* indicating imminent, unavoidable punishment, as well as by the indication of no hope of escaping punishment in the relative clause *who... could never hope to make up his loss*, here reinforced by an adjunct of high usuality and negative polarity, viz. *never* (P2) (§ 3.4.1.1).

In Period 0 the more frequent objective source pairings with uses of *must* with median degrees of obligative strength are exemplified in (173), where the source of obligation is unspecific and hence objective. In this example the use of the passive construction *be warned* (P1), the statement that the call for warning is based on an opinion (P6), the allusion that it is not the intention of the writer to encourage *enmity* (P5), and the following sentence's use of the phrase *tolerable guess*, all contribute to interpret this example as conveying a median degree of obligation. The lack of an indication of any consequences of non-compliance reinforces this analysis (P5).

(173) Personal enmity, as Horne Tooke said, 'is a motive fit only for the Devil;" at the same time, our opinion is, that men **must** be warned as well as edified.

Such is the attachment of every man to his favourite pursuit, that a tolerable guess may be formed of what a man is, by attending to what he first reads in a paper. (1820s; News)

As noted above, the pairings of median degree *must* with subjective sources are much less frequent in the period spanning the 1820s to 1860s. In (174) subjective *must* adheres to many of the parameters for a high degree of obligation, such as a second-person, animate subject (P1), negative polarity (P4), future time indication (P4), the agentive verb *send* (P1), active voice (P1), interpersonal content (P6), etc., but a pragmatic analysis of this example reveals that the clause containing *must* is meant to be read with irony, i.e. in the form of a joke from one friend to another (P5), reinforced e.g. by the direct addresses found in *little woman* indicating a quasi-reprimanding tone, and the affectionate *dear* (P5), as well as by the ironic use of an adverb conveying a sense of intensity, viz. *too* (P2). The situation is regarded as friendly, reinforced by the greeting *with a large bundle of love*, and is therefore not face threatening; it is furthermore unmotivated by social norms, and the non-compliance of the obligation has no indication of gravity (P5) – subjective *must* hence conveys a median degree of obligation here.

(174) My dear Mrs Stewart,
... I am just as glad I got the house done for I had intended to have done it before you came back at any rate. I could not bear the idea of you coming off a journey with a house to clean. So you **must** not send me too big a scolding little woman, for I would do a good deal more than that for you, dear, and I quite enjoyed polishing the dear old house. Tell me about little woman in your next and give her a kiss from me. Now goodbye with a large bundle of love,
I remain,
Yours very affectionately,
Jane E. Waterston. (1860s; Social letters)

When ratios for subjective:objective uses with median-degree *must* are considered, the ratios for Period 0 and 1 also reverse just like for high-degree uses, but in the opposite direction; from 1:2,3 in Period 0 to 1,6:1 in Period 1 in the favour

of subjective uses. It is interesting to note that this ratio for median obligation in Period 1 is exactly the reversed ratio for high obligation in the same period, which is 1:1,6 as mentioned above. Even though the ratios for subjective:objective uses are exactly 1:1 in Period 2 with the slight decrease in subjective uses, the ratio almost returns to what it was in Period 1 in the last period, but this time in the favour of objective meanings, namely 1:1,5 for Period 3. By the 1990s the proportional difference in favour of the occurrence of objective uses with high obligation *must* is consequently more pronounced than for median obligation *must*. Based on this, it can be concluded that both high and median obligation *must* in contemporary SAfE are more often linked with an objective source, but more so for high obligation *must* (with twice as much objective as subjective uses) than for median obligation *must*. This is indeed an unprototypical use of *must* (cf. § 2.4.1.1). The connection in SAfE between *must* and the objective source, especially in high-degree obligation contexts, is largely absent in other varieties, in which *HAVE to* is the more typical choice in this regard (see § 2.4.1.1; cf. Collins, 2009a:60).

As stated above, **Period 1** is the only period in which median obligation is more often paired with a subjective source. Subjective uses with median degrees of strength in Period 1 frequently occur in social letters, as can be seen in example (39) in Chapter 3, as well as in (175) and (176).

(175) B will see you all right, if Percy suffers loss, and for the rest we **must** thank God if we come out at all well from the fearful mess into which Jameson plunged us all. (1890s; Social letters)

(176) My Havelock
I am to be married on Saturday morning. ... I shall just wear my old grey dress that I wear every day. I am better, much better in health. Destroy or keep Cron's letters, don't trouble to return them. Some day you and Edith **must** come out and see us. ... Havelock dear, I love you so. Give my love to your little wife. My doctors say there's no reason I shouldn't have a child now. The womb is quite in its right place. It was doubled back on itself, you know.
Your little old Olive. (1890s; Social letters)

In (175) subjective, median obligation *must* is used in a social letter as part of a discussion of a hypothetical situation (P6), induced by the conditional marker *if*. Some formulaic expressions conveying resignation in uncontrollable, unpromising circumstances reinforce the analysis for median obligation, including *for the rest* and *at all* (P2), as well as the allusion to the *fearful mess* and the notion of thanking God if indeed the circumstances end well. In a sense, *must* is used as part of a formula, so some gradience with formulaic uses are present here, but it is quite a different formula in comparison with the uses in those instances that were classed as formulaic in § 4.3.2.2.2. In (176) the case for median obligation is motivated by the friendly communication of happy news from one friend to another (P5) – more particularly from Olive Schreiner to her friend Havelock Ellis, who were neither married, nor lovers, by all accounts, but who were very intimate friends[174]. The intimacy of the friendship is obvious from the very familiar and affectionate opening and closing, as well as addresses such as *dear* and declarations of affection as in *I love you so*. The fact that Olive chooses to discuss such intimate details about her *womb [being] in its right place* and it being *doubled back on itself* with a male friend is indeed indicative of a very close (platonic) friendship quite surprising in view of a supposed conservative 19th century context. By this account, there is no indication of a difference in social status between the correspondents (P1 and P5), despite gender differences. Furthermore the parameters present in this example is very similar to (39) in § 3.4.1.1, in that it also fits some parameters for a high degree of obligation, such as the interpersonal nature of the letter and an indication of future temporality (P1, P4 and P6), but these are overruled by parameter 5 as described above, as well as by the adverb *some day* (contemporary 'someday') (P2), which point to a simple wish for the future (as part of the act of socialising) regarding a hypothetical time and situation (P6).

As noted earlier, subjective uses with a high degree of obligation in Period 1 are fewer than in Period 0. Such uses are illustrated in (177), (178) and (179).

[174] As noted in http://www.oliveschreiner.org/vre?view=personae&entry=21

(177) My dear Laddie
When the first shot is fired and we know that all is over, we can "trek los"[175]. You however, and the men with you **must** sit fast and quiet, to the end: even after all the country is in flames. You **must** hold fast and give no opportunity to the enemy displacing you. I had a wire from Smuts the other day, begging me to come and join them at the front: what I may or may not do is not yet clear to me. (1890s; Social letters)

(178) We often have big, wet clouds around us and then I can't see the ships or the houses. The clouds come over us, and the wind blows. Then Mother says I **must** stay in the house. (1900s; Fiction/narrative)

(179) Messengers came in the name of the chief: 'Masonda does not like the blanket, he wishes you to salute him with a bag of gunpowder, and a box of caps'. I explained to them that I was not a trader; I was a messenger of peace and love and had nothing to do with powder. At sunset the chief had come himself, he **must** have powder, I **must** give it, or else I am not his friend. I repeated what I said in the morning—and to prove my friendly disposition, I offered a new American axe, which he disdainfully refused. He sulkily sat at our fire while we went to the tent to our tea. It was dark. Soon Aarone came in and whispered, 'Sir, we are surrounded by Masonda's men, and they still come down the mountain.' We went out to the chief who was devouring a piece of our meat. I tried in a pleasant way to induce him to go to his mountain home. 'No, he **must** first get a dog—then he **must** have two—then he **must** have his choice.' (1870s; Non-fiction)

In example (177) the broad textual analysis is that of interpersonal communication (P6) written in the context of the first year of the Anglo-Boer War (1899-1902) (cf. § 2.3.3.4.1), which, apart from the date of the letter, is also indicated by mentions of *the first shot* and the country going up *in flames*, as well as by allusions to *the enemy*, *the front* and *Smuts* (Jan Christiaan Smuts), one of the leaders of the Boer forces. The time frame in which Olive Schreiner writes to her brother (hence no indication of status – P1) is obviously characterised by political and social tension in South Africa.

[175] An Afrikaans expression literally translated as 'pull loose', meaning 'to begin' or 'to run'.

Although parameters for high-degree obligation are not ubiquitous, e.g. the situation is not face threatening (P5), the weight of the parameters tends toward an analysis for high-degree obligation, mostly because of the context, where the wartime context intensifies a sister's concerned and emotional pleas (P5). For the first instance of *must* in this example, an analysis for a high degree of obligation is reinforced by the adverbs *however* and *even* (P2), the latter of which denotes an exceeding level of counterexpectancy (cf. § 3.4.1.1) and which also implies such a high degree of gravity of non-compliance (P5) that might even be fatal in a wartime situation. For the second *must* the analysis for high-degree obligation is supported by the negative polarity (P4) of the coordinated clause containing the elided *must* before the main verb *give*, as in *give no opportunity*. In example (178) the subjective source of the obligation is a mother and the addressee is her child; indications of e.g. status (P1), a face-threatening act (P5) and interpersonal content (P6) are therefore present. The use of the adverb *often* (P2) connotes habitual obligation and a high degree of gravity of non-compliance (both P5). Gravity of non-compliance is also a part of parameter 5 that renders example (179) an indication of high obligation, seeing that the interpersonal negotiations (P6) are between a native South African chief (high status [P1]) and his apparent English rival (reinforced by the verb *salute*), who, if he does not comply with the chief's wishes or conditions, might face the loss of his good will, as suggested by *or else I am not his friend*, or even impending battle, as suggested by *we are surrounded by Masonda's men*. A face-threatening, tense situation is apparent here (P5), which is notable from the fact that the chief *[did] not like* or *refused* the previous gifts or bribes for safe passage offered to him previously, to the point that there is now added pressure to please him or to *prove* friendship to him. Adjuncts that contribute to the analysis include the adverbs *disdainfully* and *sulkily*, as well as time adverbs *first* and *then* connoting a demanding tone, and the modal adjunct of polarity *no* (P2).

The more frequent objective uses with high-degree obligation in Period 1 are often found in contexts where legal rights (as in [180]), politics (as in [181]) or the like are discussed, which is also true for other native varieties (Collins, 2009a:35). Here the force of obligation can be quite abstract and even function on some level of human involvement regarding an "unspecified consideration of what may be considered morally desirable" or denote an "official position" (P6) (Collins, 2009a:35-6).

(180) Anyone steadily contemplating this legislation will be struck with its distinctly socialistic character. It deprived a large number of citizens of their property, on the pleas that the needs of the many **must** supersede the rights of the few, and it cancelled legal contracts, upon the underlying principle that the land **must** not be absolutely the property of its owner though he might have paid a high price for it, but that tenants have a right to require the use of it on terms which will enable them to thrive (1880s; News)

(181) To give the Franchise would be to give without a blow the means of getting all that they would rather fight to refuse: so the only hope was Armed Bluff, ...and that rebellion **must** succeed at once and so Dutch South Africa find itself quietly taking the fait accompli (which was not to be the Union Jack, always remember...) (1890s; Social letters)

In (180) the high-degree obligation of the first instance of *must* is strengthened by the reference to *pleas*, which indicate some emotional involvement (P5), and in the second instance of *must* is made stronger by the fact that it denotes a *principle* and a *right to require* (which is indicative of social norms [P5]), as well as by the negative polarity of *must* (P4) and the adverbs *though* and *absolutely* (the latter being a mood adjunct denoting total intensity) (P2). In (181), which centres on matters surrounding the Anglo-Boer War, the mentions of *the only hope* (where *only* denotes a limiting level of counterexpectancy) (P2) and *rebellion* indicate the gravity of the situation. Furthermore, the fixed expression *at once* denotes urgency and the adverb *always* conveys a high degree of usuality (P2) – both indicators of high-degree obligation. The Latin expression *fait accompli* (literally 'accomplished fact') also contributes to the analysis, seeing that it expresses a sense of factuality and a degree of resignation to inevitability. In the following example the second instance of *must* also has a high degree of obligation and an objective source, after the first *must* expresses an epistemic meaning.

(182) Let not the reader suppose that any knowledge of the laws of Nature or of Hygiene was required to settle this question. The disease **must** be produced by supernatural causes alone. It is probable that the doctor may have obtained

from his assistants a sketch of the family circumstances of the patient, for he proceeded to elicit by guesses in the same way, that "long ago" an old woman had been defrauded of a cow, or that some other fatal mistake had been made in connection with a cow and an old woman, the result of which naturally was that some spirit (either of the forefathers generally, or of the old woman in particular) could not be appeased but by the death of the unfortunate boy who had in some way inherited a share in the transaction, or by the sacrifice of an ox. This is the inevitable conclusion: an ox **must** be sacrificed. (1880s; Fiction/narrative)

In (182) a narration of events surrounding a traditional healer or sangoma's assessment of a situation is given. After epistemic *must* in the first instance conveys the logical evaluation of the sangoma, reinforced by the use of *naturally* (which the narrator conveys in a somewhat mocking tone), high-degree deontic *must* in the second instance expresses the conclusion made by the sangoma based on that evaluation. The sangoma is, of course, a person of high status in the particular society (P1), and this, along with the adjective *inevitable*, which indicates little possibility of non-compliance (P5), renders the obligation a high degree of obligative force. The objective analysis of *must* is here motivated by the fact that the use in this example is similar to the uses in (180) and (181), seeing that the focus is on cultural rites, which in some ways correspond with legal proceedings (P6). In some way the sangoma's conclusion is based on what is logical under the social circumstances – that which epistemic *must* expressed beforehand. This corresponds to what Depreatere and Verhulst (2008:5) term a discourse-external source (§ 2.4.1.1).

The less frequent uses of objective, median-degree *must* in Period 1 is illustrated in (183) and (184). In both examples mainly the parameter of broad textual analysis renders the obligation a median degree.

(183) We wonder if the Native tribes inhabiting territories adjacent to these "true Colonists" hold that their liberties are safe? Or, if it comes to that, we should like to ask Englishmen living amongst those "true Colonists" whether they consider their liberties safe? ... We should like to ask them whether, in such cases, they **must** not carefully and cautiously consider

who among the number of such "true Colonists" might possibly feel himself aggrieved by any slip of the tongue or pen on the part of an English resident who may thoughtlessly use his old-fashioned English liberty of speech. (1870s; News)

(184) So, the man fitted to be the national leader of a great heterogeneous people requires certain qualities not asked for in the leaders, even the great leaders, of a homogenous race. ... The man who should help to guide us toward the path of true Union and a beneficent organisation, **must** be more than a great party leader, the keen diplomatist, far seeing politician, or even the renowned soldier. (1900s; News)

Must in (183), despite being negated, is basically interchangeable with *should*, in that it expresses a hypothetical meaning as part of a text with a philosophical tone (P6). The median degree of obligation is reinforced by phrases and words conveying a sense of tentativeness, e.g. *We wonder if*, *if it comes to that* and *possibly* (a weaker mood adverbial) (P2). The hypothetical meaning of *must* in (183) is further supported by polite expressions using modals, such as *should like to ask* (P5), where the modals themselves, *should* and *might*, are both traditionally the past forms of other modals and in Contemporary English indeed more often convey hypothetical meanings (cf. § 2.4.1.2; Halliday, 1970:340; Coates, 1983; Bybee, 1995:511; Larreya, 2003:21). In (184) hypothetical projections are also made by means of *must* in terms of the ideal nature of a possible national leader, and can also be substituted by *should* to the same effect as in (183).

As noted previously, and as seen in Figure 20 and Tables 8 and 9, both the slightly decreasing trend of high-obligation *must* and the increasing trend of median-obligation *must* into **Period 2** result in near equal ratios between these two uses, and are supported by the equal distribution of subjective and objective uses with median obligation (plus the occurrence of some indeterminate uses), as well as by the stability gained due to the objective use remaining the more frequent option with high obligation. The latter use is illustrated in (185). Most of the examples given below are from letters, moreover social letters, being the register with the highest occurrence of deontic uses in Period 2, as seen in Figure 19.

(185) And what exactly is love? You may love your girl friend as well as you father, but whoever heard of someone wanting to fondle his father! ... And this business of loving your neighbour: does it mean that you **must** not dislike him, even if you feel you would like to? (1950s; Social letters)

In the above instance *must* has a high degree of obligation, seeing that it is used as part of an allusion to a Bible passage, Matthew 22:37-9[176], where Jesus interprets the greatest commandment (with reference to e.g. Exodus 20: 1-17 and Deuteronomy 5: 1-22) (Bible, 1911). By implication, the obligation of *loving your neighbour* comes from God, which indicates status and authority (P1) and the commandment can be seen as being reflective of an imperative (P3), but because the obligation can be interpreted as predetermined by a religious law or in this case a commandment, indicates a tendency toward a discourse-external and hence objective source of obligation on the subjective-objective continuum. The high degree of obligation is further indicated by the adjunct *even* (with an exceeding level of counterexpectancy) and the negative polarity expressed by *not* (P4). Example (186) illustrates the objective source co-occurring with median-degree obligation.

(186) The Afrikaner, therefore, **must** be taught that he had every reason to be proud of his language, his literature, his religion, his way of life; and the Englishman, in his turn, **must** be constrained to recognise in the Afrikaner a person not inferior but different, and in his own way every bit as good. (1940s; Non-fiction)

The two instances of *must* in (186) are similar to the use in (184), where they express hypothetical meanings (as part of a non-fiction text) similar to that of *should*, and the argumentative nature of this text (corresponding to philosophical content [P6]), as indicated in the adverb *therefore* and the phrase *in his turn* (P2) support the analysis for median obligation. This further indicated by the unspecified third-person subjects

[176] "Jesus said unto him, Thou shalt love the Lord thy God with all thy heart, and with all thy soul, and with all thy mind. This is the first and great commandment. And the second is like unto it, Thou shalt love thy neighbour as thyself. On these two commandments hang all the law and the prophets" (Bible, 1911).

the Afrikaner and *the Englishman*, which add a layer of generalisation to the obligation, as well as the use of the passive voice in both instances (P1).

Median-degree uses like these have equal propensity to co-occur with a subjective source in Period 2, as seen in (187), (188), (189) and (190) (see Table 9).

(187) I came away feeling that we in S.A. **must** solve our own prison problems with our own intelligence. (1940s; Social letters)

(188) I think I **must** get a new suitcase & put into it such things as army coat, rain coat, old hat, jerseys, old flannels, black shoes (which are too painful to wear), & leave the case at the hotel. It will lighten my other luggage. (1940s; Social letters)

(189) She was standing in the yard again, her bundle in her hand. I came out smiling; I felt better for her. Good-bye, Janet, I said. And tell your mother I hope she'll be better. And you **must** come and tell me how she is, eh? There was no answer... ...and then she began to cry... ... But what's the matter, my girl, I said. What's wrong? You **must**n't cry. ... Here, I said. Here – take this, and gave her my handkerchief. (1950s; Fiction/narrative)

(190) My dear Sarie,
Forgive me for taking so long in writing to thank you for your very very welcome Xmas parcel... ... I am hoping that when the weather gets warmer I'll be a nice friend for my friends to have – & a better correspondent! But shall I? You **must** try, my dear Sarie to have faith in men so that I may have faith in myself! ...
Bless you...
Pauline (1950s; Social letters)

In (187) and (188) the subjective source is internal to the discourse, in this case the first person. The obligation in (187) is median mainly because of pragmatic factors, e.g. there is no gravity in non-compliance with the obligation (P5), the obligation is unmotivated by social norms or emotions, and textual factors, e.g. the content is hypothetical and based on personal opinion (P6). This is reinforced by the use of the non-finite verb *feeling*, which the clause containing *must* is complement to, to indicate

a tentative degree of conviction or stance (similar to other expressions, as in P2). In (188) the tentative degree of conviction is summed up in the verb *think*, here used as a hedge (P5), which, just like in (187) takes the complement clause with *must*, to express an internal train of thought. A small degree of overlap with the dynamic meaning can be read into especially (188), but the fact that *must* does not so much convey an internal need of the subject referent or a need arising from circumstances, but rather a projection of a future action to be taken in order to solve a problem, in this case to *lighten ...the luggage* (i.e. a conclusion is made), is more indicative of a deontic meaning, but at the very weakest end of the strength continuum. Despite the interpersonal content in (189), the collocation with imperative clauses, as in *tell your mother...* and *take this* (P3), an implied difference in social status (an older woman speaking to a younger girl) (P1 and P5), as well as a somewhat emotional situation (P5), some pragmatic factors once again weaken the obligation (of both instances of *must*) to the extent that they overrule the fact that there are some parameters present for a high degree of obligation. The main factor in this regard is that the obligation is not of a face-threatening nature (P5): this is clearly a situation where the subjective source of the obligation cares about the addressee and her feelings, as is apparent in the fact that she comforts her (and offers her a handkerchief when she cries) – this moment of comfort is indeed the second instance where *must* is used. Another factor here is that there is no apparent gravity of non-compliance (P5), and the presence of some expressions that convey a sense of polite concern (P5), including *eh?* and *my girl*. In example (190) an even less face-threatening, friendly relationship (P5) is evident between the relevant correspondents, as revealed by the opening and closing of the social letter, e.g. *Bless you*, and especially by the affectionate, first-name address *My dear Sarie*, and the informality of the correspondence is clear from the spelling choice in *Xmas.* The first part of the letter is apologetic, which indicates that there is no indication of status (P1 and P5) in the English writer (here Pauline Smith) over the Afrikaans addressee, *Sarie*[177]. Furthermore, *must* is here used as part of what can be interpreted as an ironic or humorous sentence, which is suggested by the final exclamation mark, meaning that the obligation expressed here, being part of a joke

[177] 'Sarie' is a common female Afrikaans first name for the era (a derived, diminutive [affectionate] form of the Afrikaans equivalent of 'Sarah', viz. 'Sara' or 'Saar')

(P6), will have no consequences if not complied with (P5). Considering all these factors, a median degree of obligation is evident.

The less frequent cases were high-degree *must* has a subjective source in Period 2 are exemplified in (191).

(191) I spoke to your house master about a number of things. He said, "You **must** conform but you need not agree." ... As part of the school and as a member of your own house you **must** toe the line in conformity with the rest of the boys, even though you disagree in principle with some of the rules and regulations. (1950s; Social letters)

Here the obligation expressed by the *house master* (contemporary 'housemaster'), is relayed by a third party, a father, who addresses his son, so even though the source is technically external to the discourse, he is still an actual authority figure whose direct words are being reported to the intended receiver of the obligation, rendering it subjective. The housemaster is of higher social status than the son (P1), the situation is hence governed by social norms and refers to habitual obligation (P5), by which the son is compelled to comply with *rules and regulations* (P6), rendering this an instance of high-degree obligation. Collocating with the first instance of *must*, the deontic use of modal *need* to expresses what is allowed contrasts with what is not allowed (reinforced by the coordinator *but*), and hence contributes, by stating the contrary, to intensify the force of the obligation. In the second instance of *must* in (191), the high degree of obligation is reinforced by the expression *even though* (with *even* conveying an exceeding level of counterexpectancy) (P2), which has much the same effect as *but you need not agree* on the first instance of *must*.

In the final period of discussion, high- and median degree *must* share a near-equal level of frequency, as noted previously (see Figure 20). Also, both high- and median degree *must* are more frequently paired with the objective source of obligation by **Period 3**, but more so for high-degree *must.* Examples (192) and (193) illustrate the objective, high-degree use in this period. All of the examples from Period 3 are from non-fiction, fiction/narrative and news, being those registers that show an increase in the occurrence of deontic *must* over the 20th century, as seen in Figure 19.

(192) ...from the provisions of subsection (2), all official languages **must** enjoy parity of esteem and **must** be treated equitably. (1990s; Non-fiction [W2D-003])

(193) A will is a formal document and you **must** make sure every page is signed or at least initialled by you, and the final page signed by you... (1990s; Non-fiction [W2D-004])

Both (192) and (193) are found the 'administrative or regulatory writing' genre of ICE-SA and hence express regulatory meaning, here specifically in legal contexts (P6), and, as seen in previous examples, this is the context in which most of the objective uses of *must* occur (cf. Collins, 2009a:35). There is therefore little possibility of non-compliance (P5). Nouns like *subsection* and *provisions* in (192), and *will* and *formal document* in (193) confirm the regulatory nature of these texts. Apart from the textual and pragmatic factors, the analysis for high-degree obligation is strengthened by adjective *every* and the fixed phrase *at least*, despite the passive voice in (193).

Example (194) contains median-degree *must* with an objective source. The direct quotation in (194) contains an opinion (P6), where *must*, as part of the (philosophical) argument concerning an unspecified, hypothetical situation (reinforced by *When*, *just* and *but*) (P6), could just as easily be replaced with *should.*

(194) "When black people are invited to sit on boards they **must** not just function as a token black but **must** shake and move things," he said. (1990s; News [W2C-010])

The less common subjective uses with high- and median-degree obligation in Period 3 are exemplified respectively in (195), as well as (196), (197) and (198). Firstly (195) illustrates the subjective, high-degree use.

(195) Yet many whites have come to accept the inevitable. "They are mad!" says florist Wendy Marshall of the school authorities. "It was going to happen some or other time. They **must** accept it!" And while many white residents share the view that the end of segregation was the inevitability rather than a

moral imperative, the parents encamped at the school have long since lost the support of apartheid's one-time moral apologist, the Nederduitse Gereformeerde Kerk[178] just across town. (1990s; News [W2C-004])

In (195) the subjective source of the obligation is Wendy Marshall. Even though there are quite a few parameters for median-degree obligation present in (195), this instance of *must* was analysed as having a high degree of obligation, based on the emotional motivation for the utterance (P5), reinforced by the collocating exclamation *They are mad!* and the fact that the sentence containing *must* is presented in the same way, viz. *They must accept it!*. To an extent, it can be argued that the obligation is imposed by circumstances, e.g. an objective source, since the acceptance of the obligation is based on what is *inevitable* rather than what is the *moral imperative*, which in itself, can be taken as an indicator of a high degree of obligation, but since the person performing the speech act is internal to the discourse and directly quoted, the case tends toward the subjective end of the continuum. The gradient nature of micromeanings in deontic *must* is therefore plain here, as this case could probably be classed as unprototypical or peripheral. Secondly, (196), (197) and (198) illustrate the subjective, median degree use of *must* in Period 3, excepting the one epistemic use in (196).

(196) "This **must** be the boy. ... Miss Ellie said he was quite big." ... "She said we **must** fetch you; bring you to the house. She says you know me. I don't know where she gets that from, because we never did meet. (1990s; Fiction/narrative [W2F-013])

(197) "Duifie, you **must** read what he says after he came back from America. One of those dom[179] sisters of his married a whatchamacallit... you know... am... ambassador." ... "Are all his letters to you like this?" "But not always such big words." "No, I mean so philosophical. What do you write to him about?"

[178] The Dutch Reformed Church ('Nederduitse' would directly translate as 'Dutch-German'). This is a traditionally Afrikaans denomination in South Africa, with a reformed (as opposed to charismatic) spirituality, essentially based on Protestant doctrine. Despite the fact tact the Dutch Reformed Church initially (during the 1960s-70s) supported and promoted segregational legislation (being closely affiliated with the National Party government), it later in the 1980s played a major role in pointing out Apartheid's flaws and injustice (under such figures as e.g. J.A. Heyns, David Bosch and Nico Smith), and in the eventual abolishment of its laws (Hofmeyr, 2002:194;208-9;216). Today, this church advocates equality and engages with many moral issues and controversies in South Africa.

[179] An Afrikaans word meaning 'unintelligent' or 'stupid'.

"Ag[180], Duif, jy weet[181]. What is happening on the farm, and the drought and what **must** I do with all the little kaffertjies[182] which are being born on the farm all the time." (1990s; Fiction/narrative [W2F-005])

(198) "But don't be disheartened at the small number of black women who participate in races. You **must** run and encourage your friends to run as well." (1990s; Fiction/narrative [W2F-014])

In example (196) the second instance of *must*, was analysed as conveying a median degree of obligation, since the subjective source of the obligation, even if external to the discourse, in all probability did not give the speaker an order to fetch the child, but rather made a request, although it cannot be said for sure. There is however a degree of certainty to be gained in the matter when considering the context of the utterance. The clause *She said we must fetch you*, is offered as a means to inform, and hence reassure the child, so the intention is not so much to relay that the speaker is under some obligation, but rather to give a clarification of matters that will gain the child's trust, as they *never did meet*. Even if the subjective source is discourse external, her direct request is relayed here and *She*, the subject of *must*, is specified as *Miss Ellie*, much like in (191), rendering it subjective. The uses in (191) and (196) illustrate that the distinction between discourse-internal and -external sources does not always apply directly to only either subjective or objective cases, especially in reported speech, although this is a useful distinction in general (see § 2.4.1.1; cf. Depreatere and Verhulst, 2008:5). It is not clear from the context whether the speaker is of lower status than the subjective source, and this makes it difficult to directly apply some of the parameters to this instance (although there might be some indication of higher status in the use of *Miss* to refer to the source of the obligation), but as far as the overall intention of the utterance containing *must* goes, it can hardly be described as face-threatening in the discourse (P5). This example is therefore not a prototypical instance of subjective median-degree *must*, but it does tend toward this end of the scale. In example (197) a friendly, informal situation is implied (P5) by the use of the

[180] The Afrikaans loan expression 'ag' is equivalent to English 'oh'.

[181] An Afrikaans loan expression that translates as 'you know'.

[182] The plural, diminutive form of an offensive term denoting a native African, derived from the Arabic *kāfir*, meaning 'infidel'. The term is notorious in South Africa as a symbol of oppression and racism.

affectionate addresses *Duifie* and *Duif* (Afrikaans words meaning respectively meaning 'little dove' and 'dove'), and the very informal expression *whatchamacallit* (probably a reduction of 'what do you call it?'). The first instance of *must* in (197) would have the same meaning if it were replaced by *should*, since it conveys a median degree of force on the lower end of the scale – apparently to inspire curiosity in the context of socialising (P6). The second instance of *must* in example (197) is a relayed question the speaker asked earlier in a letter, and could just as easily be replaced with *should* or *BE supposed to* in asking for advice. In cases such as these there is evidence for Bowerman's (2004a) claim for SAfE *must* often being replaceable with *should*. This is also true for example (198), where the addressee is encouraged by means of *must*, and is under little or no obligation to comply. This interpretation is reinforced by the consolation in the imperative form, *Don't be disheartened* (so in this case P3 is counterproductive), and by the fact that the addressee is motivated to *encourage* others *as well*, where *as well* implies that the addressee herself was encouraged in the first place. Thus, this case does not contain a face-threatening act (P5) and in some ways expresses a wish of the subjective source (P6), with no apparent gravity in the event of non-compliance (P5).

Apart from the distinction between high and median degrees of *must* and their slight inclination toward objective sources, another kind of deontic meaning is discernible, namely the formulaic use, which more frequently co-occurs with subjective sources.

4.3.2.2.2 Formulaic *must*

For the formulaic deontic uses of *must*, which are more than the median obligation uses in Periods 0 and 1 (with 14,6 and 15 normalised formulaic instances in these periods respectively), a slight increase from Period 1 to 2 (to 17,1 normalised instances, which is less than for median uses by this time) can be detected before a decline occurs into Period 3 (to 10,9 normalised instances – a lower frequency than in Period 0). Collins (2009a:38) notes that for formulaic uses of *must*, which are part of the "conventional politeness formulae", the degree of modality is very low, i.e. it has little or no social force, is not at all face-threatening, and is actually "readily omissible" as a result. It can be said that the higher frequency of formulaic *must* in the

periods preceding Period 3 may point toward the colloquialisation of SAfE, seeing that it is part of the "conventional politeness formulae", as described above, that can here be seen to have become less over the 20th century, and hence a move toward less formalised styles can be detected. The drop in deontic *must* in the letters register from Period 2 to 3, as seen in Figure 19, indeed supports this, since formulaic uses in letters have higher frequencies in Periods 0, 1 and 2 than in Period 3, which has more formulaic uses in non-fiction – this in turn supports the rise of deontic *must* in non-fiction over the 20th century, as also seen in Figure 19.

Table 10 gives the percentages for the co-occurrence of formulaic uses with the two deontic sources, as well as indeterminate cases.

Table 10

Percentages of formulaic deontic uses with subjective, objective or indeterminate sources in *must*

Period	Subjective	Objective	Indeterminate
0	57,2	32,1	10,7
1	67,9	25	7,1
2	76,7	10	13,3
3	52,6	26,3	21,1
Average	63,6	23,4	13,1

From Table 10 it is clear that subjective uses are on average much more frequent than objective uses. Far more indeterminate uses are also present than for high- and median-degree obligation, precisely because the degree of modality is so low, which makes it more difficult to determine a source of any kind. The ratios for subjective relative to objective uses change from 1,8:1 in Period 1 to 2,7:1, and from 7,7:1 back to a smaller 2:1, resembling the ratios of Period 0 and 1 to a degree.

Formulaic *must* with an objective source is exemplified below for each of the four periods in chronological order. Such instances are mostly found in news and non-fiction (despite [201] being a 1945 letter extract by a Natal resident), and the impersonal pronoun *it* is often the subject of the clause. The passive voice is also very productive in these uses.

(199) Two methods have been devised for ascertaining the age of fishes, both of which (it **must** be confessed,) are more ingenious than certain... (1820s; News)

(200) It **must** be further and especially observed that Natal is nearly 200 miles from the nearest of the Cape towns, and that the communication between the two colonies is by sea. (1870s; News)

(201) It **must** not be forgotten that the Native generally is still a savage at heart, understanding unfortunately only brute force, full of Native cunning, mistaking decency for weakness. (1940s; Non-fiction)

(202) It **must** be stressed that entrepreneurs who succeed possess not only a creative and innovative flair, attitudes and behaviours, but also solid general management skills, business acumen and the ability to use networks effectively. (1990s; Non-fiction [W2A-014])

The next four examples illustrate formulaic uses of *must* with subjective sources over the four periods.

(203) I **must** now entreat Sir that these statements may not be construed into complaints, for I do assure you my brother speaks of the assistance rendered him and all the settlers by Gov't with gratitude and respect... (1820s; Business letters)

(204) I **must** close now, dear sister. Write to me as soon as you can. (1870s; Social letters)

(205) You **must** appreciate that lack of loyalty does not mean disloyalty. (1950s; Social letters)

(206) One could read anything into such a long look, and I **must** admit I was a bit afraid of her. (1990s; Fiction/narrative [W2F-013])

These kinds of uses were mostly found to collocate with first- and second-person subjects in letters, but also in fiction/narrative. These being the registers with a relatively interactive nature (either by means of written correspondence or the

interaction of fictional characters), the overall propensity for subjective sources is nothing out of the ordinary as part of that nature.

In summary, the ecology of SAfE *must* in Leech *et al.*'s terms (2009) by the late 20th century markedly tends toward deontic uses (macro-level monosemy), and moreover, deontic uses that are, firstly, less face-threatening, in that uses with a median degree of force are just as frequent as those with a high degree of force (micro-level polysemy), but in both cases objective sources are more frequent than subjective sources (i.e. no clear link between one force and one source is detected), and secondly toward uses that are less formalised. For the second finding it was already noted above that the decline of politeness formulae in formulaic *must* over the 20th century is indicative of the colloquialisation process, and for the first finding it was noted that the increase of less face-threatening uses in median *must* is indicative of democratisation, despite the rise of high degree uses. The two prominent processes at work in 20th-century (native) English (cf. e.g. Mair, 2006), namely colloquialisation and democratisation, are therefore also at work in SAfE in the case of *must*, but democratisation is here indicated not by the decline of *must* due to its restriction to stronger uses, but contrastingly by the rise of *must* and its expansion into median-degree obligative contexts. The next subsection will explore the influence of Afrikaans on this trend, being so very different from the trends of other native Englishes.

4.3.2.2.3 Deontic *must* in spoken SAfE and deontic *moet/moes* in spoken Afrikaans

Seeing that Bowerman (2004b:477) described *must* as having "less illocutionary force" in SAfE than in other varieties (and given his example of such a use, which is one of direct address with a second-person subject) (cf. § 1.1.3), the implication here is that this weaker or median-degree use would arise in interactive contexts, typically in spoken conversation where a second person is directly addressed. Because the diachronic corpus only comprises written texts, with some registers in a way representing interactive language (as e.g. social letters and speech representation in fiction/narrative), deontic *must* with second-person subjects in the spoken component

of ICE-SA was additionally analysed for strength of obligation in order to fully assess this claim. Furthermore, to assess whether there is correspondence between the deontic use of SAfE *must* with second-person subjects and deontic Afrikaans *moet/moes* with second-person subjects, an analysis of this fellow West-Germanic cognate modal and its past form was conducted for the spoken corpus of contemporary Afrikaans as well. On account of the relation between English and Afrikaans, the same parameters were applied to the Afrikaans analyses as to the SAfE analyses. The results of these analyses are presented in Tables 11 and 12. The analysis of the synchronic, spoken SAfE and Afrikaans data reveals that Afrikaans *moet/moes* with second-person subjects are 18% more frequent than SAfE *must* with second-person subjects, as seen in Table 36 in Appendix 6. Only these second-person subjects were analysed for the deontic meaning and its micromeanings, for the reasons explained above.

Firstly, Table 11 presents the results for the macrosemantic analysis, according to deontic, dynamic and epistemic meanings.

Table 11

Macrosemantic frequencies per 100 words for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register

	Deontic	**Dynamic**	**Epistemic**
SAfE *must*	83,7	4,8	11,5
Afrikaans *moet/moes*	87,4	11,2	1,4

From the above table it can be seen that, among all the instances of *must* and *moet/moes* with second-person subjects, the vast majority of these instances have deontic meanings in both SAfE and Afrikaans. On this macrosemantic level, SAfE and Afrikaans appear to have very similar qualitative tendencies in the spoken register, which links with the finding regarding their quantitative similarity, as found in § 4.2.2.3. This contrasts with the finding of Collins (2009a:37) that deontic *must* in

AmE, BrE and AusE is least often used with second-person subjects[183]. Afrikaans *moet/moes* only has a slightly higher percentage of deontic uses over SAfE *must* (by 3,7%), and *moet/moes* is more often used in dynamic contexts than *must* (by 6,4%), but the reverse is true for epistemic contexts, where *must* is used more often than *moet/moes* (by 10,1%). Table 12 presents the results of the microsemantic analyses for these deontic uses, according to high- or median-degree obligation, as well as formulaic uses. Because only *must* and *moet/moes* with second-person subjects were analysed, the part related to subject and personal pronouns in high-degree obligation in parameter 1 naturally applied to all cases, but the other parameters usually overruled this part of parameter 1 in the case of median obligation.

Table 12

Microsemantic deontic frequencies per 100 words for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register

	High	**Median**	**Formulaic**
SAfE *must*	18,4	77	4,6
Afrikaans *moet/moes*	24,8	70,4	4,8

The microsemantic frequencies of SAfE *must* and Afrikaans *moet/moes* can be seen from Table 12 to also follow very similar synchronic tendencies, which is indeed qualitative confirmation that SAfE *must* follows an Afrikaans-like pattern in its semantics, especially regarding median-degree uses. The high frequency of median obligation *must* in spoken SAfE (with second-person subjects) by the end of the 20th century, supports the above-mentioned finding for *must* in written SAfE to increase in its frequency of median uses over the 19th and 20th centuries (§ 4.3.2.2.1). The tendency for median-degree *must* to be frequently used is therefore not only a written trend, but moreover a spoken trend in interactive contexts, where a second person is mainly directly addressed, which confirms Bowerman's (2004b) claim and the implications of his 'example' noted above. With a high percentage of 77% out of all deontic uses in this context, median-degree *must* is 58,6% more frequent than high-

[183] Collins' data was however analysed in a different way, in that all deontic instances were analysed according to first-, second- or third-person subjects, so this comparison should only be seen in this light.

degree *must*. By comparison, median-degree *moet/moes* is 45,6% more frequent than high-degree *moet/moes*. Formulaic usage frequencies for both *must* and *moet/moes* are near equal. On the whole, the microsemantic patterns of SAfE and Afrikaans are quite alike. The role of language contact and identity in this case will be discussed and interpreted in Chapter 5, but first a general exemplification of uses will be given

High-degree SAfE *must* with a second-person subject is frequently used in medical contexts in the spoken part of the ICE-SA corpus, adding a formal, regulatory layer of meaning to the obligation. Such uses are exemplified in (207), as well as in the first instance of *must* in (208).

(207) B: ...normally you would say go home and stay in bed for a week and um you know you'd get over the flu with some aspirins to keep his temperature down.

A: That's right.

B: Now I can remember Morris getting flu, and and the doctor said to him look you **must** really stay in bed for a week, and after three days Morris got up and there was a great new job going on ...and it was practically snowing in that night, and you know, Morris nearly ended up with double pneumonia in the end it was that <unclear> the doctor said to him you know it's your fault entirely because I told you to stay in bed for a week, don't give me this <unclear> now. (1990s; S1A-037)

(208) D: There we are. ... That's the breast tablet. ... It's Stilabanacol, OK? You **must** take one three times a day. You're only going to start on the twenty-ninth of the seventh, when you get chemo.

G: On the twenty-ninth?

D: Yeah. From the twenty-ninth of the seventh to the eleventh of the eighth. One three times a day every day for two weeks.

G: Oh so I **must** take this on the twenty-ninth. (S1A-083)

From the contextual information in the corpus, it is clear that the interlocutors in (207) are two elderly female friends. Speaker B relates a story to Speaker A, in which a high-degree of obligation is expressed by means of *must* in reported speech. In the

situation in which the obligation is expressed, the source of obligation is *the doctor* and the receiver is *Morris*, here referred to in the second person (P1), which indicates the authority of the source (P1). Furthermore, *must* is used here as part of interpersonal content (P6) in conveying instruction. The obligation expressed here requires actualisation with a high degree of gravity of con-compliance (P5), because of the authority of the source and the topic of the reported speech, which is medical and in some sense therefore regulatory (P6), as noted above. This is supported by the very fact that the obligation was not complied with, with severe, near fatal, consequences. The analysis for high-degree obligation is further strengthened by the serious address of the doctor in *look*, as well his reproach in the phrases *it's your fault entirely*, *I told you* and *don't give me this*, which all contribute to signal a face-threatening act (P5). The adjuncts *really* (a mood adjunct conveying a high degree of counterexpectancy) and *entirely* further strengthen the obligation (P2). The medical theme is also discernible in (208), where the conversation takes place between a pharmacist (Speaker D) of unknown gender and a female patient (Speaker G). A degree of authority is therefore discernible (P1). In the first instance of *must* the customer is directly addressed in the second person (P1), showing the interpersonal content (P6), and in the second instance the customer repeats the obligation to signal her understanding and acceptance thereof, showing willingness to comply, since non-compliance would have severe consequences (P5). This analysis is strengthened by the mentions of *chemo* and *breast*, which suggest that the medication is for cancer. In the first instance of *must* direct instruction is given regarding medication dosage. This is reinforced by the elided imperative expressed in *[Take] [o]ne three times a day...* (P3). The analysis for high-degree obligation is further reinforced by a limiting level of counterexpectancy expressed by the adjunct *only* (P2).

Much the same situation is present in Afrikaans *moet/moes*, where regulatory contexts mostly contain the high-degree uses with second-person subjects. These are exemplified in (209) by means of *moes* and in (210) by means of *moet*. My English translation of the Afrikaans is given directly below each example.

(209) Weet julle weet julle hoe't ons gaan klasdraf? Ons 'n man **moes** 'n langbroek aangehad, jy mag glad nie 'n kortbroek aangehad het nie, **moes** 'n langbroek gewees het, jy **moes** toe skoene aangehad het. Jy mag glad nie 'n t-hemp aangehad het nie, jy **moes** 'n ding aangehad het met 'n kraag, enigiets wat

soos ’n t-hemp gelyk het sonder ’n kraag het daai hulle’t jou summier weggejaag... (2000s; GAG9)

‘Do you know do you know how we would go to class? We a man **had to** (must) wear trousers you may not have worn shorts, **had to** (must) be trousers, you **had to** (must) wear closed shoes. You may not have worn a t-shirt, you **had to** (must) wear a thing with a collar, anything that looked like a t-shirt without a collar that they’d forthwith chase you away...’

(210) … so hy hoef nie, o, as hy nooit getroud was nie hoef hy nie, maar so jy **moet** onderhoud betaal as jy getroud was met haar, hmm, net aan die vrou, maar jy **moet** altyd vir jou die kind onderhoud betaal… (2000s; GAG5)

‘... so he need not, oh, if he were never married he need not, but so you **must** pay child support if you were married to her, hmm, only for the woman, but you **must** always pay child support for your the child...’

In section 2.4.3.1 it was mentioned that Ponelis (1979:246) deems Afrikaans preterite *moes* to sometimes have a lower degree of obligation than its present form *moet*, seeing that the preterite often conveys hypothetical meanings. In example (209), however, *moes* is used purely as an expression of past obligation, and the degree of the obligation is high, since it denotes social rules and regulations (P6), much like in some of the cases mentioned above, and also in § 4.3.2.2.1. This example is indeed the only instance of deontic *moes* with a second-person subject with a high degree of obligation, but interestingly, only one instance of *moes* in this context with a median-degree of obligation was also found in the spoken Afrikaans corpus. There is therefore no conclusive evidence for the preference of *moes* in expressing median obligation in Afrikaans regarding second-person subjects. In example (209) the second-person singular pronoun *jy*, as found in the latter three instances of *moes* in this extract (after the first instance uses the third-person), is actually used here as a generic second person (P1), where the Afrikaans expressions *jy moes* can be best translated to the English *you had to* or *one had to*, since *must* has no grammatical past form in Contemporary English (cf. § 2.4.3.1), but, nevertheless, a high degree of obligation remains present. The obligation is therefore habitual (P5), despite being grounded in the past and having a slightly hypothetical layer, which further renders

the obligation a higher degree. This analysis is reinforced by the use of the modal *mag*, the cognate of English *may*, (in the translated form 'may + have + past participle') here denoting past permission that is motivated by social norms (P5) (in the past contexts perhaps more aptly translatable as 'could'), and by the use of the phrase *summier weggejaag*, where *summier* translates to 'forthwith', or in its adjective form in legal proceedings to a meaning denoting 'summary jurisdiction, proceedings or judgement' (Pharos Afrikaans-Engels & English-Afrikaans Woordeboek/Dictionary, 2005:567), and *weggejaag* to 'chased away'. This indeed reinforces the idea of rules of conduct or social laws, which is in itself face threatening (P5), and that the gravity of non-compliance would be devastating socially (P5). In (210) the regulatory nature of the obligation expressed by the present form *moet* in both cases is clear from the legislative term *onderhoud*, which translates as 'child support' (P6). This is supported by the use of *as* ('if'), indicating the condition on which the obligation is contingent. The degree of the gravity of con-compliance is therefore very high, seeing that it would entail breaking the law – essentially committing a crime (P5). The median-degree modal use in both instances of negated form *hoef nie*, which translates as 'need not', contrasts with the high degree of strength in positive *must*, which is a reversed case of parameter 4. The contrast is also indicated by the coordinator *maar* ('but'). Furthermore, the adjunct *altyd* ('always') conveys a high degree of usuality (P2), which strengthens the analysis for high-degree obligation by adding a habitual layer of meaning (P5). The second-person subject (P1) in this context is also generic.

Seeing that the high-degree uses of both *must* and *moet/moes* in the spoken corpora most frequently occur in regulatory contexts of one sort or another, it can be argued that, because the sources of obligation are almost equally divided between the subjective (as in [207] and [208]) and objective (as in [209] and [210]), in some ways even the uses with subjective, discourse-internal sources in (207) and (208) have an implied objective source, seeing that the consequences of non-compliance are not social, but natural (like, e.g. disease). In (209) and (210), however, the consequences of non-compliance will be social, even if the source of obligation is really objective.

In contrast to these high-degree uses, median-degree *must* in spoken SAfE is exemplified in (211), (212), (213) and (214), as used in informal face-to-face conversation. Each of these examples will be individually analysed.

(211) A: No but I am going through, what I mean, if you were having visitors and you didn't have Selina would you parboil them still?

B: Well you can, but then then you **must** only put them in the olive oil near the time you know you **must** get that <unclear> tender and soft

A: No but I mean if you are dealing with a parboiled <unclear> would you do this about chicken?

B: Yes and no, I mean I still find it better to parboil them and they go in into the hot <unclear> while they are hot I mean... (1990s; S1A-037)

The interlocutors in (211) are two female friends in their 70s – the same as found in (207). The fact that they are friends is indeed annotated in the text as part of the contextual information, but even from the speech itself this is discernable from the reference to *Selina*, which is indicated to be the domestic worker of one of the speakers, showing that they assume common knowledge, and are on familiar terms (P5). Speaker B is giving cooking advice (P6) to Speaker A. The personal pronouns show that they are speaking face to face in an interpersonal context, but despite these isolated indicators of a possible high degree of obligation, the full analysis according to all six parameters evinced a median degree. The obligation expressed here via *must* indeed has a reading on the lower end of the strength continuum, which can be argued to convey a quite neutral instruction, which is grounded in a hypothetical context (P6). This is supported by the use of the subjunctive in *if you were* and the hypothetical expression using the modal *would*, namely *would you... still*. The fact that the friendly advice that is given by means of *must* (P6) (reinforced by *I mean* and *I ...find*) is not accepted by Speaker A in the second turn, because the first question was apparently misunderstood, shows that there is neither a face-threatening situation, nor social gravity in the event of compliance (P5). This is reinforced by the phrase *No, but I mean*. Median obligation as part of the giving of advice is also exemplified in (212).

(212) C: Well it's tough because I can't rush with her you know, she wants to look at this it takes ages to climb out of her car seat and I just didn't feel like rushing. It was hot.

S: You **must** take her to St James now after schools have gone back.

C: Oh ja[184].

S: You know the way she's interested in water that's got nothing in it and you take her to St James and all those little pools there she'd have a ball. (1990s; S1A-025)

In (212) the interlocutors are family members, probably female, since the text is marked to have been recorded at a family lunch party, and they discuss Speaker C's child. Apart from the broad context of the text, the fact that a child is being discussed is evident from the mentions of a *car seat* and *after schools* (places where children receive day care after school hours). This is therefore a friendly situation, where no actualisation of the obligation is required and non-actualisation would not have any consequences (P5). Advice is merely offered as an act of socialising (P6) and the addressee (Speaker C) clearly has a choice whether to accept the advice or not. This is reinforced by the polite reply *oh ja*, which signals that the addressee acknowledges the polite (P5) advice of Speaker S, basically by means of backchanneling[185], without necessarily accepting it. In fact, *must* (with its positive polarity [P4]) in this context is easily replaceable with 'polite' *should*, which would express an equally low degree of social force (according to the traditional view of *should*) – agreeing, once again, with Bowerman's (2004b) aforementioned claim. Thus, no threatening of the addressee's face takes place (P5). At this point, it is interesting to note that the act of giving advice or making suggestions, which is the least face-threatening among the acts listed in § 2.2.1.2.2, was linked to the modal *ought* by Brown and Levinson (1987:65), one of the modals with the lowest degree of social force by traditional standards. Yet, in this case, *must*, the traditional marker of a high degree of social force, is used in exactly such a context. This is supported by the fact that Speaker S offers an explanation in clarification of her suggestion, which is much the same strategy applied in dynamic and deontic uses of *moet* in Afrikaans, as noted in § 2.4.3.1 (cf. Conradie, 1987:178), adding a layer of politeness.

Much the same situation is present in (213).

---

[184] An Afrikaans loanword meaning 'yes'. 'Ja' itself is an entrenched lexical feature of SAfE.

[185] 'Backchanneling' refers "to the reactions given to a speaker by way of feedback". These include "monosyllabic responses (*mhm*), short phrases (*I guess so*), utterance repetitions and sentence completions, as well as non-verbal cues (e.g. nodding, gaze variation)" (Crystal, 2008:48).

(213) C: It's the same with Jan and Annette that that they bought something and then

D: Is the signal strong enough that

C: and Christo can also

A: Well I had to put in a booster. Just a small booster to to get the signal strong, so

D: Is that a sort of amplifier?

A: No no. It's more a <unclear>.

C: You **must** ask Jan because they also had that

A: Willie Willie knew about these things and he just told me buy a booster and he told me where to buy the booster. (1990s; S1A-055)

Not enough direct information is available for the text from which (213) is extracted, but from the context itself it clearly appears to be a conversation between friends – in the case of this extract a conversation between four interlocutors. From this extract, for example, the fact that the speakers constantly interrupt one another, as well as the mentioning of common acquaintances by means of first names, signal familiarity and hence a friendly situation (P5). Furthermore, the topic, in this part of the conversation being a M-Net[186] decoder and a problem with its transmission signal, reveals that Speaker A is sharing this problem with the others, who offer their advice or suggestions in this regard (P6). The aim of the suggestion given by Speaker C by means of *must* is clearly to help solve Speaker A's problem, and in this context, as in (212), is easily replaceable with *should*. The intention is therefore not face threatening (P5). The analysis is strengthened by Speaker C's previous two attempts at offering advice on whom to speak to – people who also encountered similar problems as Speaker A, namely *Jan and Annette*, as well as *Christo*. When Speaker C is interrupted, an attempt is later made again regarding the suggestion to *ask Jan*, in this case by means of *must*. Further factors that render this a case of median obligation are that non-compliance with the suggestion would have no gravity (P5) and the positive polarity of the suggestion itself (P4). Notwithstanding, even in some uses with

[186] A large, commercial South African television broadcasting company, which required a special decoder to access its programmes, and which was popular in especially the 1990s.

negative polarity, *must* in the spoken part of ICE-SA conveys median obligation, based mostly on pragmatic factors, as seen in (214).

(214) A: OK now the question part may be negative or it may be positive and you can ...make any riddle that is going to carry the positive part of the question. Do you have any riddles that you know?

B: Are you asking me?

A: Yes. <laughter>

B: You **must**n't ask me. I don't count you see. <laughter>

A: Yes. (S1B-004)

Example (214) is taken from a conversation between two male, academic colleagues at a class seminar, so there is no apparent difference in status (P1 and P5). They are indicated in the text as the lecturer (Speaker A) and his colleague who is observing the class (Speaker B). During the course of the class, the lecturer asks his colleague a question relating to the course, after which Speaker B makes an observation using *must*, rather than really imposing an obligation. Despite the more formal nature of the class situation, this extract illustrates a lighter, more informal moment, as indicated by the laughter of both Speaker A and B, as well as by Speaker B's surprise and/or uncertainty in *Are you asking me?*. This is therefore not a face-threatening situation (P5) with a low degree of gravity in the event of non-compliance (P6). Indeed Speaker B motivates his observation when he says *I don't count you see*, and even though Speaker A does accept this, he clearly has a choice whether to do so or not, and hence there is a possibility of non-compliance (P5). In a sense this use is quite neutral, and is replaceable with 'you shouldn't ask me' as well as with the imperative instruction without a modal 'don't ask me' – much the same as often happens with Afrikaans *moet* (cf. § 2.4.3.1). Textual and contextual factors, as indicated in the title of this thesis, therefore greatly aid the interpretation of this instance of *must* as conveying a median degree of obligation.

Examples (215) and (216) illustrate the unmarked positive form *moet* and (217) illustrates the entrenched negated form *moenie* (equivalent to 'mustn't') in expressing median obligation. It is interesting to note that out of the ten raw instances of *moenie* in the spoken Afrikaans corpus, all ten convey deontic meanings and nine

of those have second-person subjects. Out of these nine instances eight carry a median degree of obligation and only one carries a high degree of obligation. Negation is evidently not so much a parameter for high-degree obligation in Afrikaans *moet/moes* as it is mentioned to be for e.g. English *must* in § 3.4.1.1, but as seen in (214), other parameters can easily overrule the presence of negative polarity in the interpretation of median obligation in SAfE.

(215) …weet jy ek is lief vir blomme en diere, hm, en natuurtonele ook en toe't ek eendag dit gesien en gedink jy weet dis mos simmetries en jy **moet**, hm, jy **moet** dit perfek doen, hm, jy kan 'n blom skeef maak, maar as jy dit doen **moet** dit hm ja perfek wees (2000s; GAG7)

'...you know I love flowers and animals, hm, and natural scenes also and when I saw it one day and thought you know it's symmetrical and you **must**, hm, you **must** do it perfectly, hm, you can make a flower at an angle, but if you do that it **must** hm yes be perfect'

(216) A: …ek sal eerder diere gaan werk. Tannie, die onderwyseresse wat jou skoolgehou nè, jy daai kleintjies skoolhou, jy **moet** hulle bliksem. Hulle't vir jou hulle't jou mos baie moeilikheid gegee.

B: Ja ja ja, onderwysers se kinders... (lag) (2000s; GAG4)

A: '... I will rather work [with] animals. Aunty, the teachers that taught you, you know, you teach those little ones, you **must** smack them. They gave you a lot of trouble, didn't they?

B: Yes yes yes, teachers' kids...[187] (laughs)'

(217) A: …want sy't gesê sy weet hoe ek voel, hm, want ek is nie lus om te gaan sonder jou nie, dit gaan nie

B: OK, as jy nie lus is nie **moenie**. Lus om te gaan? (lag) (2000s; GAG5)

A: '...because she said she knows how I feel, hm, because I don't really want to go / I am not inclined to go without you, it won't

[187] The Afrikaans syntax in example (216) is quite unclear in some places, and it was attempted to give a direct as possible translation; however some phrases have been freely translated to make some sense of it.

B: 'OK, if you are not inclined to go, **don't**. Want to go? (laughs)

In (215) *jy* refers to a generic second person, rendering the situation in which the obligation is grounded a hypothetical one (P6). Indeed *moet* in this context is almost translatable to *NEED to* and borders upon a dynamic meaning, since the modality arises from an external, objective necessity in some ways, which is reinforced by the dynamic possibility meaning conveyed by *kan*. This is indeed an ambiguous case, but the reading where the speaker is the discourse-internal, subjective source of the obligation is somewhat more salient. This level of gradience between deontic and dynamic readings renders the obligation a very low degree indeed. In (216) a very informal conversation takes place between a younger speaker and an older speaker addressed as *Tannie*[188] ('Aunty'). This is a direct translation, but in Afrikaans *Tannie* is an address of respect to any older woman, so it does not necessarily refer to a relative, an aunt, in this context. It is unclear whether Speaker B is in fact a relative of Speaker A, but nonetheless, the obligation conveyed by *must* is offered by a younger to an older person, so the lower social status (P1) of Speaker A renders it a median degree. A degree of irony is however present (P5), as evident in the laughter of Speaker B in reaction to Speaker A (P6), so the obligation is not face-threatening and no actualisation is required (P5). The informal, friendly nature of the conversation (P5) is further shown by the use of the colloquial, mildly offensive term *bliksem*[189] and the expressions *nè* and *mos*, that do not have any direct equivalent expressions in English, but in this context broadly translate to respectively 'you know' and the tag question 'didn't they?', signalling a level of informality. In example (217) the interlocutors are apparently in a romantic relationship, though the exact nature of the relationship, whether married or not, is not clear from the context. It is not clearly indicated what gender Speaker A and B have, but to some degree it is recoverable from the context that Speaker A is female and Speaker B is male. This is implied by Speaker A's mention *sy't gesê sy weet hoe ek voel* ('she said she knows how I feel', showing that a female third person relates to her current situation. There is therefore no difference in social status between the interlocutors (P1 and P5), and their informal conversation is friendly (P5). Here *moenie* is translatable to *don't*, since it is unnatural

[188] The contemporary diminutive and affectionate form of achaic Afrikaans *tante*, meaning 'aunt'.

[189] This Afrikaans word originally means 'lightning', but is often used as a verb meaning 'to smack/hit' or 'to fall down' (where it is considered offensive), or as a noun as a form of insult.

and indeed not possible to say 'if you are not inclined to go, mustn't' in English. Hence, in this context *moenie* is essentially doing the same job as a normal imperative – once again being fairly neutral in strength, once again confirming Ponelis' (1979:251) notion that *moet* is the unmarked or neutral option to make requests or give commands. This is of course not really a command or a request, but more a suggestion (P6). In fact, a very interesting thing happens here. Speaker B actually releases Speaker A from an obligation (to go somewhere alone despite her not wanting to) by means of *moenie*, rather than really imposing a new one. Speaker B is essentially saying it will be all right if Speaker A chooses not to go. This is reinforced by Speaker B's use of *OK* and by the laughter at the end of the extract, as he jests (P5) with Speaker A when he asks *Lus om te gaan?* ('Want to go?'). The degree of obligation is therefore very low in this example.

In this subsection it was found that both spoken SAfE *must* and spoken Afrikaans *moet/moes* clearly prefer the deontic meaning, and moreover, the median obligation meaning (with near-equal percentages), where they collocate with second-person subjects. This was taken as qualitative evidence that SAfE *must* follows an Afrikaans-like pattern in its semantics, especially in interactive contexts, namely the precedence of the weaker deontic force of *must* over the stronger, as it is for *moet/moes*, which contrasts with the tendency for *must* in other native varieties of English to mostly convey a high degree of obligation, i.e. a stronger force. Since the semantic analyses of SAfE *must* yielded such noteworthy results, *should*, is in need of a macro- and microsemantic description to compare its use with that of *must*.

4.3.3 *Should*

Seeing that part of Bowerman's (2004b) claim involved interchangeability, in other words equivalence or synonymy, between weaker *must* and polite *should* in SAfE, this section will discuss the macro- and microsemantics of *should*, the traditional modal rival of *must*, to investigate whether this is true. Evidence for this is indeed already suggested above, where many instances of median-degree *must* were found to enjoy a degree of synonymy with median-degree *should*. More generally, these discussions will also shed light on the diachronic trend of written *should* over the 20th

century to become less frequent than *must* by the 1990s, as well as the synchronic tendency for *should* to be less frequent than *must* in spoken SAfE as well, which is very different from the trend in other native varieties (see § 4.2.1.4.1 and § 4.2.2.2).

4.3.3.1 Macrosemantics of *should*

As noted above, *should* conveys two special meanings not conveyed by other modals or quasi-modals in the obligation and necessity cluster, namely the quasi-subjunctive and preterite meanings, and does not convey dynamic meanings, like its rivals *must* and *HAVE to* do (cf. Coates, 1983:67; Huddleston & Pullum, 2002:177; 186-7; 202; Leech, 2003:233; Collins, 2009a:45). The major distinction between deontic and epistemic modality is however also present for *should*. The normalised diachronic macrosemantic trends for *should* in SAfE that arise from the data in Table 29.2 in Appendix 5 are presented in Figure 21.

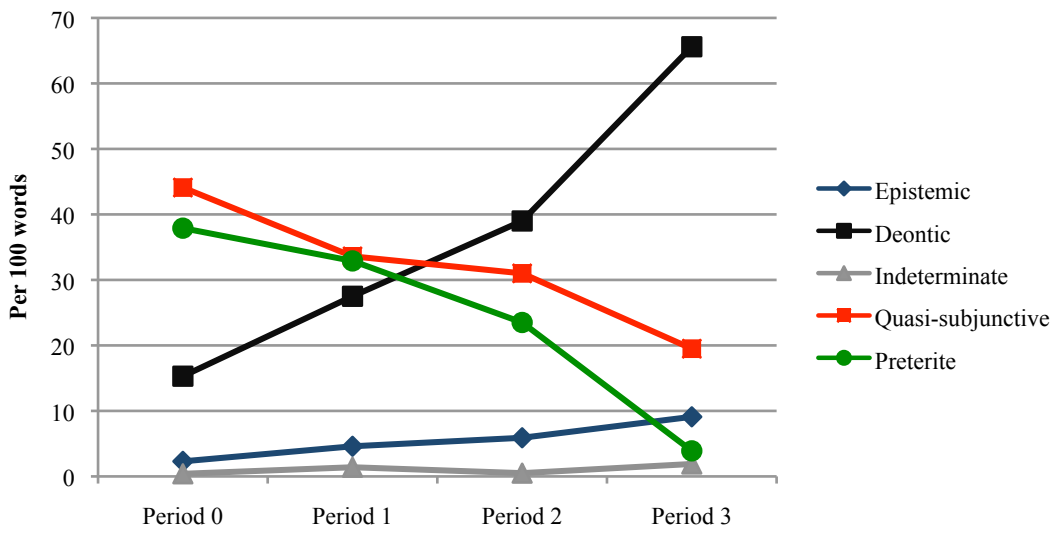

Figure 21

Macrosemantic frequency changes of *should* in SAfE over time

The rising trend of deontic *should* is immediately striking in Figure 21, where it contrasts sharply with the declines of the quasi-subjunctive and preterite meanings. In Period 0 the frequencies of the various meanings are indeed quite different from their arrangements by Period 3. The quasi-subjunctive, after having precedence over the

other meanings in Period 0, is for instance far less frequent than the deontic meaning in Period 3, which is indeed the most frequent meaning from Period 2 onward. The decline of the quasi-subjunctive and especially preterite meaning, particularly from Period 1 onwards, are therefore the main carriers of the general declining diachronic trend of *should* from Period 1 to 3, as reported in § 4.2.1.4.1. The deontic frequency starts out at 15,3 normalised instances in Period 0, which increases to 27,5 instances in Period 1 (+79,7%), being only slightly less frequent than the preterite and quasi-subjunctive meanings in this period (respectively 32,9 and 33,6 normalised instances). Into Period 2, deontic *should* again increases with some deceleration to 39 normalised instances (41,8%), but into Period 3 the increasing deontic trend accelerates when it rises to 65,6 normalised instances (+68,2%). By Period 3, the semantic ecology of *should* comprises 65,6% deontic, 9,1% epistemic, 19,5% quasi-subjunctive, 3,9% preterite and 1,9% indeterminate meanings. When compared to the findings of Collins (2009a:34) that *should* in AmE, BrE and AusE combined is divided between 69% deontic, 11,8% epistemic, 6,6% quasi-subjunctive, 2,7% preterite and 10% indeterminate meanings, the percentages of deontic, epistemic and preterite meanings are quite similar, but the quasi-subjunctive is 12,9% more frequent in SAfE, and indeterminate meanings were found more frequently in the other native varieties.

The root meaning of *should*, which is basically the deontic meaning (since *should* does not express dynamic meanings [Collins, 2009:45]), is reported to become more dominant in AmE and BrE from 1961 to 1991/2 (Leech *et al*., 2009:86), which corresponds to the trend of SAfE *should* from Period 2 to 3, as the deontic meaning takes increasing precedence over the other meanings over these periods. On a macrosemantic level, the same trend toward monosemy for *should* in other native varieties can hence be detected for SAfE *should*, in that root/deontic meanings are becoming more frequent, but it will be shown below that this is not a complete account regarding trends related to monosemy and polysemy. Leech *et al.* (2009:86) further reports that root *should* already enjoys dominance over other meanings by 1961 in AmE, but that it only becomes more dominant in BrE in the 1991s, i.e. it rises more in BrE over the 20th century and remains more stable in AmE. Since deontic SAfE *should* is already dominant over other meanings by Period 2, which represents the early to mid-20th century, the degree of dominance of deontic *should* in SAfE over

the other meanings is more similar to the case in AmE (cf. Millar, 2009:203[190]). Because deontic *should* covers somewhat less semantic space in SAfE (deontic *should* increases from covering 39% in Period 2 to 65,6% in Period 3, as mentioned above) than in AmE over the course of the 20th century (root *should* increases from covering 72% in 1961 to 79% in 1991), with SAfE deontic *should* therefore rising more than AmE, it can be argued that SAfE is more like BrE in this instance. According to this account, then, SAfE deontic *should* seems to maintain a middle position between the trends for AmE and BrE, which indeed matches the overall diachronic trend for *should* reported in § 4.2.1.4.1, where it also maintained the middle ground in comparison with the other varieties. But, the trend for *should* toward monosemy in other native varieties moreover includes a microsemantic aspect. It was reported in § 2.4.2.2 that *should* progressively expresses meanings on the weaker end of the deontic continuum of strength, which was found to contrast with the rise of stronger uses in *must* (cf. Leech *et al*., 2009:86-7). Since the latter trend was found not to correspond to the trend in SAfE, the former will be investigated for SAfE in § 4.3.3.2.1, when the degrees of obligative strength of *should* will be discussed.

Figure 22 shows the change in *should*'s macrosemantic frequency within each of the four registers over the 20th century, i.e. from Period 2 to Period 3. This figure is generated from the data in Table 31.2 in Appendix 5.

Figure 22 shows that deontic *should* rises in news (+264,7%), fiction/narrative (+93,5%) and non-fiction (+32,1%) over the 20th century, which all support the general rising trend of deontic *should* shown in Figure 21 (with news apparently being the major contributor to this trend), but decreases in letters (-85,6%), albeit not sufficiently to hamper the general deontic trend. Epistemic *should* also shows a decline in letters (-48%) and increases in news (+57,1%) and non-fiction (+18,2%), but unlike the deontic meaning, the epistemic meaning disappears altogether in the fiction/narrative register from Period 2 to 3 (-100%). This slows down the rising trend of epistemic *should*, as seen in Figure 21, and induces a slight rather than sharp rise.

[190] Millar (2009:203) however does not distinguish between weaker and stronger degrees of obligation in his analysis, but treats the entire root/deontic meaning of *should* as weak.

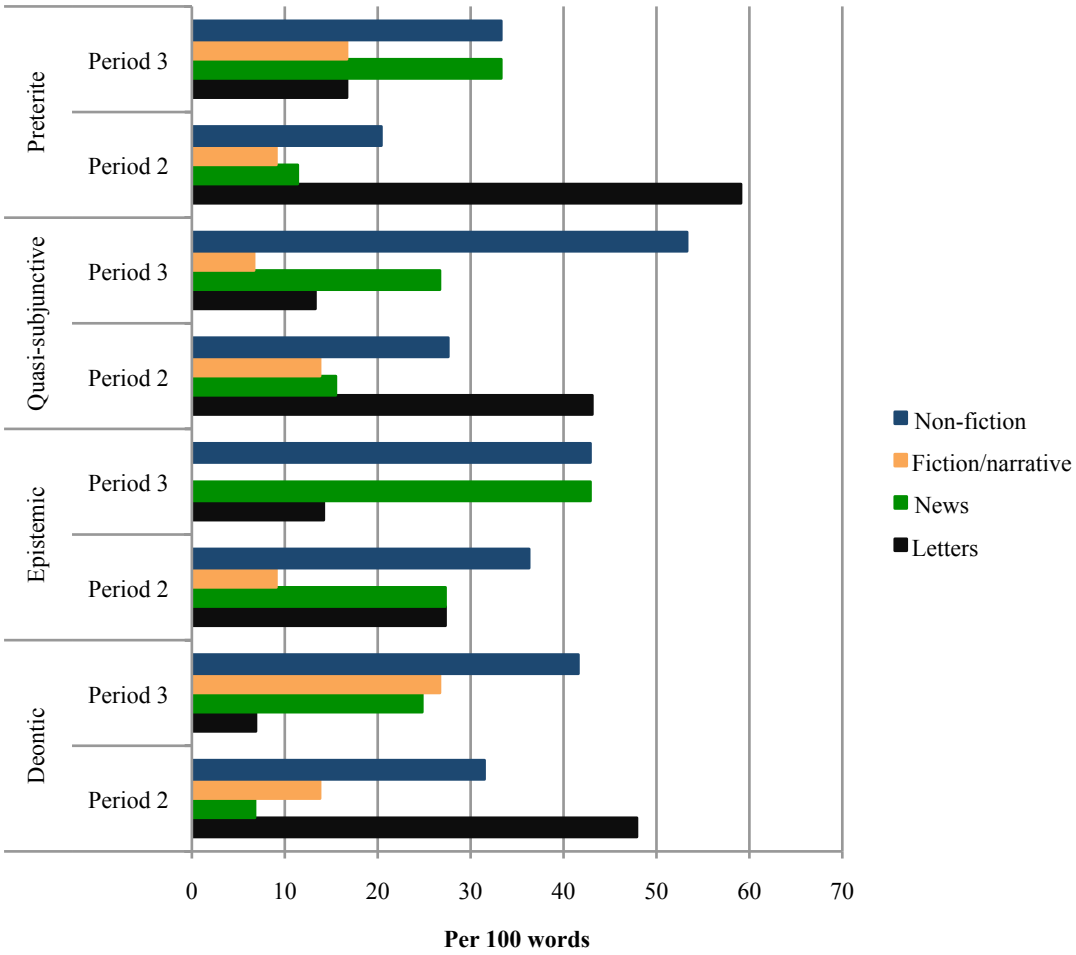

Figure 22

Macrosemantic frequency changes of SAfE *should* per register over the 20th century

Despite the increase of quasi-subjunctive *should* in news (+72,3%) and non-fiction (+93,1%), both letters (-69,1%) and fiction/narrative show a decrease (-51,4%), with letters being the register that contributes the most to the overall trend of decline for quasi-subjunctive *should* over the 20th century (see Figure 21). The sharp general decline of preterite *should* from Period 2 to 3, also seen in Figure 21, is mainly supported by its decrease in letters (-71,7%) and non-fiction (-63,2%), and is not impeded by its increase in news (+192,1%) and fiction/narrative (+83,5%). The letters register is therefore the only register in which all four the meanings of *should* decrease (most notably in the deontic meaning), and can hence be said to mainly support the general decrease of *should* in letters from Period 2 to 3, which in turn was found to be the main contributor to the statistically significant decline of *should* over

the 20th-century, as reported in § 4.2.1.4.1. This suggests that *should*, in all its meanings, is used less frequently in interpersonal contexts, like letters, and has become more frequently used in informative or argumentative contexts, like non-fiction. Compared to *must*, there is therefore a degree of stylistic, rather than semantic, specialisation present for *should*. The epistemic, quasi-subjunctive and preterite meanings of *should* will be discussed first, before focusing on deontic *should*.

4.3.3.1.1 Epistemic *should*

Figure 21 shows that, apart from indeterminate cases, the epistemic use of *should* is generally the lowest-frequency meaning for *should* in SAfE. Its overall trend across all periods is however one of increase: after it doubles in frequency from Period 0 to 1, despite its frequency remaining quite low, epistemic *should* rises by 28,3% into Period 2 and by a larger 54,2% into Period 3. In Period 2, the 11 raw instances of epistemic *should* are basically equally divided between the registers of non-fiction (4 raw instances, news (3 raw instances) and letters (3 raw instances), but only one such use occurs in fiction/narrative (see Table 31.1 in Appendix 5). As far as small numbers can tell, by Period 3 news catches up with non-fiction, when, out of the 14 raw instances, 6 such instances each are from news and non-fiction, and 2 are from letters. The increase in non-fiction and especially news (see Figure 22) supports the trend of increase for epistemic meanings over the 20th century, as discussed in § 4.3.3.1, despite the slight decrease in letters. Epistemic *should* in the non-fiction of Period 2 is exemplified in (218) and (219), where it is typically used to convey "a tentative assumption" or an "assessment of the likelihood of the predication" that is made (cf. Collins, 2009a:46). These uses were further typically objective, where "the proposition is deduced from known facts" (2009a:47).

(218) The United Kingdom has now also undertaken to buy at least 70 per cent of the Union's exportable surplus of canned fruit – an undertaking which **should** prove of considerable value to South African growers. (1950s; Non-fiction)

(219) With the memory of free access to land, with keen appreciation of their present exploitation, the natives **should** be easy to reach. (1910s; Non-fiction)

The same use of epistemic *should* as in the non-fiction of Period 2 becomes more frequent in the news of Period 3, as illustrated in (220) and (221), where *should* also expresses an objective deduction, assumption or "assessment of the likelihood of the predication", from known facts (cf. Collins, 2009a:46-7).

(220) The lack of any visible action on privatisation probably accounted for the market's initial disappointing reaction. In spite of these shortcomings, the overall direction remains encouraging and if properly implemented, **should** boost growth prospects. (1990s; News [W2C-010])

(221) Interest rates **should** drop sharply over the next five years, bringing relief to hard-pressed borrowers, if the Government's newly announced economic strategy works as planned. (1990s; News [W2C-005])

The fact that this particular objective use of epistemic *should* is most frequently found in both non-fiction and news by the 1990s is unsurprising, given the mainly informative and nature of these two genres, and indeed especially so for news, where the tentativeness of the assumption(s), as well as the deduction from fact(s), is shown by the use of *probably* and *In spite of* in (220), and by *if* in (221).

It was briefly mentioned in § 4.3.2.1.2 that, while epistemic *should* rises from Period 2 to 3 by 54,2%, as noted in the current section, epistemic *must* drops by 54,2% over the same time span. Evidence for a semantic trade-off between these two modals regarding the epistemic meaning is indeed very strong, seeing that their percentages of increase and decrease are precisely equal – making this a most interesting finding considering the opposite trend in other native Englishes, as also noted in § 4.3.2.1.2 (cf. Leech *et al.*, 2009:87-9). The quasi-subjunctive meaning is however still more frequent than the epistemic by Period 3.

4.3.3.1.2 Quasi-subjunctive *should*

What is termed the 'quasi-subjunctive' meaning by e.g. Coates (1983:67) and Leech (2003:233) is also called the 'putative' by e.g. Leech *et al*. (2009:86). This meaning includes "[s]pecial 'low-degree uses of *should*" that contribute "little discernable

modal meaning" and mainly occur in subordinate *that*-clauses (Collins, 2009a:48). Leech *et al*. (2009:86) add that the modal meaning is so weak that in these uses "*should* manifests mood rather than modality: it has no epistemic or deontic flavour, but instead expresses the non-factual nature of the predication pure and simple". The 'mood' they refer to here is of course the subjunctive mood, which is why this meaning is termed the 'quasi-subjunctive'. Five types of subordinate constructions that can convey quasi-subjunctive meanings are listed by Huddleston and Pullum (2002:187), viz. the mandative, adversative, purposive, emotive and conditional constructions. The mandative construction is e.g. a "more informal alternative to a subjunctive" (cf. Collins, 2009a:49), which underpins the term 'quasi-subjunctive'. Collins (2009a:49) further notes that *should* can "combine with a predicative item of strong modality in the matrix clause", e.g. items such as '(it is) insistent that' or '(it is) important that', which suggests that *should* "has undergone grammaticalisation in the mandative construction", i.e. it has extended or expanded its syntactic restrictions.

From Figure 21 it is visible that the quasi-subjunctive meaning is in general decline in SAfE. From Period 0 to 1 it drops by 23,8%, and remains fairly stable with a decrease of only 7,8% from Period 1 to 2, before declining again with some acceleration (by 37,1%) from Period 2 to 3. The latter trend of decline over the 20th-century is the same as the trends for both AmE and BrE from 1961 to 1991/2 (Leech *et al*., 2009:87).

Figure 23 again illustrates the declining diachronic trend of the quasi-subjunctive in SAfE over the 19th and 20th centuries, but this time as it is influenced by the five types of subordinate construction listed above. The data used to generate this figure is presented in Table 35.2 in Appendix 5.

It can be seen that the mandative construction is generally the most frequent, and retains the larger proportion of quasi-subjunctive uses over all periods. Its trend of decline over the course of the entire time span however carries the overall trend for the quasi-subjunctive meaning to a large extent: it declines by 27,1% from Period 0 to 1, by 19,5% from Period 1 to 2 (a deceleration), and finally by 8,6% from Period 2 to 3 (an even greater deceleration).

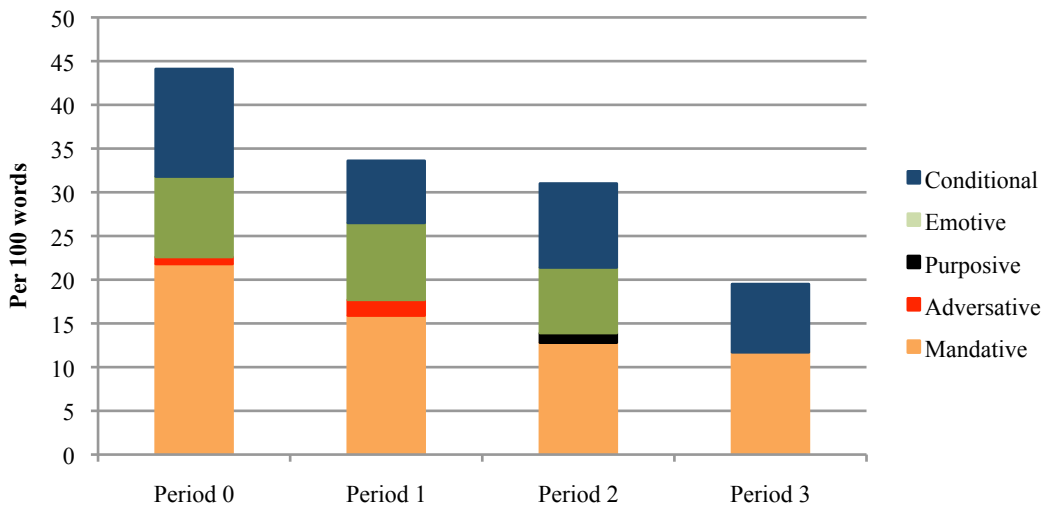

Figure 23

Frequencies of the five types of quasi-subjunctive uses of *should* in SAfE over time

In some cases the mandative construction "co-exists readily with the deontic meaning", in that it expresses what is desirable (Collins, 2009a:49). This can be seen in the following four examples for each of the four periods, where the mandative meaning is respectively conveyed by the constructions *it was deemed very undesirable that*, *we do not advocate that*, *it was agreed that* and *It is very important that*.

(222) ...a letter which I had written as to the desirableness of establishing fairs or markets within the Colony ...had been well-received, ...yet it was deemed very undesirable that any Missionary Society **should** be permitted to commence a Mission in Kaffraria. (1830s; Fiction/narrative)

(223) We do not advocate that we **should** run needlessly in debt, just to show that we enjoy good credit. (1880s; News)

(224) At that meeting at which the Bishop of Johannesburg was in the chair ...it was agreed that a great concerted effort **should** be made... (1920s; Business letters)

(225) It is very important that she **should** know what he has seen on the television before his afternoon snooze. (1990s; Fiction/narrative [W2F-005])

Furthermore, the disappearance of the emotive construction from Period 2 to 3 supports the decline of the quasi-subjunctive meaning across the 20th century. This construction often expresses surprise or some kind of evaluation, as seen in the following examples from fiction/narrative, which is the register in which these examples were most often found. Its highest frequency occurs in Period 0, which is exemplified by (226), but it gradually decreases over the course of Period 1 (-4,3%) (as illustrated in [227]), and accelerates in Period 2 (-14,8%) (as seen in [228]), before vanishing completely. The disappearance of the emotive construction in the fiction/narrative register therefore supports the decline of quasi-subjunctive *should* in this register over the 20th century, as seen in Figure 22. In the following examples the evaluative nature of the emotive use of the quasi-subjunctive is respectively seen in the constructions *it was better [that]*, *Strange that* and *how ridiculous it was that*.

(226) ...and the people who were thus to be deprived for a season of the residence of a Minister in the midst of them, believing that it was better they **should** submit to this temporary disadvantage, ...kindly acquiesced in the arrangement... (1830s; Fiction/narrative)

(227) She spoke Norwegian and I answered in English as a ride; but when words and gesticulations failed we had recourse to Zulu, in which we were both at home. Strange that a barbarous African dialect **should** form the only available means of communication between two European women. (1880s; Fiction/narrative)

(228) And when he gathered dimply that the defending advocate was explaining to the jury how ridiculous it was that a man **should** kill six people for the sake of an empty sack, he nodded his head slightly. (1920s; Fiction/narrative)

The conditional meaning also decreases from Period 0 to 3, but its trend is internally rather fluctuating. It first decreases by 42,3% from Period 0 to 1, but increases by 35,2% from Period 1 to 2, before decreasing again by 18,8% from Period 2 to 3. Its uses in each of these periods are illustrated in the next four examples. This construction can express a conditional meaning with the support of *if*-clauses (as in [231] and [232] or without (as in [229] and [230]).

(229) **Should** Eaton (to whom I have given him an introduction) not be in the way, be so good as give him a note to Thompson and Pillans, requesting them on my account to afford him any advice... (1830s; Social letters)

(230) I therefore collected my things together, on the principle of having them as light as possible, and in duplicate, the half of which I could leave en cache, when I had to quit my waggons, as a store to all back upon **should** I happen to meet with robbery or accident. (1880s; Fiction/narrative)

(231) Such was the situation at the beginning of September 1939: Smutsism and Hertzogism on the one hand, Krugerisme on the other; the one urging "in the first place a South African nationhood, and then, if the people **should** wish it, a Republic, in or out of the Commonwealth..." (1940s; Non-fiction)

(232) This significantly reduces contamination if a different product **should** be loaded afterwards. (1990s; Non-fiction [W2B-032])

The frequency of the adversative construction also contributes to the decline of quasi-subjunctive *should*, since it disappears entirely after Period 1, after increasing by 125%, but even in these first two periods it is seldom used, as seen in its low occurrence of raw instances in Table 35.1 in Appendix 5. Its use in Period 0 and 1 is respectively exemplified in the following extracts, where they are paired with a *lest*-clause, expressing a meaning close to 'in case'.

(233) They are afraid of going to the banks of some of their rivers by night, lest they **should** be dragged in by the spirits of the people living in the river... (1840s; Fiction/narrative)

(234) The delay is lest the Landdrost **should** in the meantime hear that the war had dried up when when anything we might do would only embarrass Government. (1870s; Business letters)

The only two raw instances of the purposive construction are found in Period 2, where, in both cases, it collocates with the phase 'in order that', as seen in (235).

(235) On the other hand, in order that our universities **should** be in a position to discharge their functions in this regard ...certain conditions will have to be fulfilled. (1950s; Non-fiction)

Like the quasi-subjunctive meaning, the preterite meaning of *should* is also one that is neither conveyed by *must* nor by *HAVE to*.

4.3.3.1.3 Preterite *should*

The preterite use of *should* refers to this modal "as the preterite counterpart of *shall*", involving backshift, where the meaning is grounded in the past time, or as part of "the apodosis[191] of an unreal conditional", even though *should* "cannot be used with independent time meaning" to the same extent as *could*, *would* and *might* (Collins, 2009a:51; cf. § 2.4.1.2). In this use, *should* is basically substitutable by *would*, and usually has a first-person subject (Collins, 2009a:51; cf. Leech *et al*. 2009:86). Formulaic uses can also occur, where the use of *should* makes the modality "politely tentative" (Collins, 2009a:52). Like quasi-subjunctive *should* in SAfE, preterite *should* in this variety has a higher frequency than deontic *should* in Period 0 and declines across all periods, to be much less frequent than deontic *should* by Period 3. In fact, preterite *should* all but disappears by the 1990s, with 3,9 normalised (6 raw) instances, when compared with its higher frequencies in especially Period 0 and Period 1. In Period 0, *should* is only slightly less frequent in its preterite use than in its quasi-subjunctive use, but the declines in both these uses into Period 1 renders them with near-equal frequencies in the latter period. Here preterite *should* declines by 13,2% from the period of input to the 1870s-1900s. From Period 1 onward, preterite *should* declines in an accelerated fashion: by 28,6% from Period 1 to 2 and by 83,4% from Period 2 to 3. The decline of preterite *should* in SAfE over the 20th century is consistent with its decline in BrE, as well as AmE, over the course of the 20th century (Leech *et al*., 2009:87).

[191] This refers to "the consequence or result expressed in the main clause of a conditional sentence; also called the **consequent**, and opposed to the **protasis**, which expresses the condition. In the sentence *We shall get in if we queue*, *we shall get in* is the apodosis, *if we queue* is the protasis" (Crystal, 2008:30; emphasis original).

In consideration of the reference to *would* replacing *should* in the piece on Cape English quoted in § 1.1.3, there is some support for this observation in the historical data from Period 0 to 2, and given that the author wrote this text in the 1870s, his observation matches these periods. The support for this claim is based on the simultaneous increase of *would* over Periods 0, 1 and 2 (see § 4.2.1.3) and decrease of the preterite use of *should* (in which *should* can be substituted with *would*) over the same periods (see § 4.3.3.1). From Period 2 to 3, however, the frequency of *would* decreases along with the frequency of preterite *should*, which suggests that this trend, where earlier *would* was perhaps the more popular option over preterite *should*, did not continue into the late 20th century. Yet, this, much like the case of *will* replacing *shall* (§ 4.2.1.3.2), cannot be regarded as an exclusive feature of SAfE, as the quoted piece suggests, considering the general increase of *would* in other native varieties over the early 20th century and a decrease towards the end of the 20th century (as noted in § 4.2.1.3.1), as well as the decline of preterite *should* in other native varieties over the 20th century noted above.

The use of *should* as the preterite counterpart of *shall* in Period 0 is illustrated in (236), where it could have read 'and requesting that you shall be confirmed in those lands' to the same effect. Indeed the use of *shall* in the preceding sentence reinforces this analysis. The use "in the apodosis of an unreal conditional", as described by Collins (2009:51), is seen in (237), and examples (238) and (239) illustrate the formulaic use of *should*. *Should* in (237) can be substituted with *would*, as in 'if they had attacked us then we would have cut a very poor figure'. The same applies to the latter two examples from Period 0, where the interchangeability with *would* is indeed shown by the alternating use of *would* in close proximity.

(236) I shall write to the governor, pointing out to his excellency the injustice that is attempted to be practiced on you, and requesting that you **should** be confirmed in those lands by some legal title, to preserve you from these disturbances. (1830s; Fiction/narrative)

(237) The Kat river was rather high and most of the soldiers had their guns and ammunition wet. We were now pretty well among the enemy and if they had attacked us then we **should** have cut a very poor figure. (1850s; Fiction/narrative)

(238) We were also shown many different sized milk pans (which we **should** imagine would be much valued by the housewives of Natal)... (1860s; News)

(239) I wonder if Rutherfoord received a letter I wrote him a fortnight ago with some balderdash in it, I wish you would ask him, as I **should** not like it to fall into the hands of the Rooks. (1820s; Social letters)

Most of the preterite uses in Period 0, 1 and 2 indeed have first-person subjects, as also illustrated above in (237), (238) and (239). In Period 1 the bulk of the uses with first-person subjects, which in turn are also the most frequent preterite uses overall, are formulaic and are mainly found in social letters, as in (240). Here the use of *would* in a collocating sentence as part of the exact same formula 'I would/should be so glad if you could' indeed shows that preterite *should* is substitutable by *would*.

(240) I would be so glad if you could send me a copy of Fred's poem. He has forgotten the words. I **should** be so glad if you could copy them for me. (1880s; Social letters)

By Period 2 the formulaic uses with first-person subjects are still the most frequent among the preterite uses, and are found with almost equal frequency in both social and business letters – letters being the register in which preterite *should* mostly occurs in Period 2, as seen in Figure 22. The occurrence in business letters is illustrated in (241).

(241) Under the circumstances I **should** prefer to withdraw my protest as the verdict could have no effect now on the positions obtained for the Corporation Cup. (1940s; Business letters)

Only one formulaic use however occurs by Period 3, where preterite *should* mostly conveys the purely backshifted use, substitutable by *would*, as seen in (242).

(242) I find it hard to understand why you **should** want to go to HK. It is such an overcrowded, polluted, materialistic place... (1990s; Social letters [W1B-004])

The frequent occurrence of formulaic preterite *should* in the social letters of Period 1, along with the apparent spread of these uses to business letters by Period 2 and the near disappearance of such uses by Period 3, suggests that the politeness formulae expressed by preterite *should* have become a marker of more formal contexts, and, as a consequence, have fallen out of use, which suggests a degree of colloquialisation.

4.3.3.2 Microsemantics of deontic *should*

Figure 24 illustrates the proportional differences in deontic frequency between high and median degrees of obligation for *should*, as well as formulaic uses. The number of the parameter that applied in each exemplified analysis will be provided in the same fashion as in § 4.3.2.2.

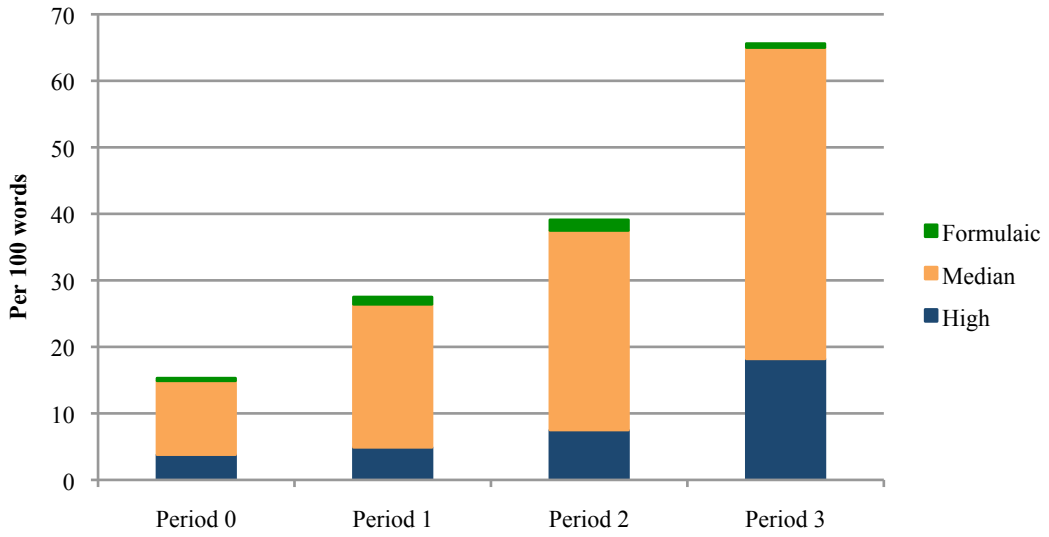

Figure 24

Microsemantic frequency changes in SAfE deontic *should* over time

From the above figure it is clear that the general trend of increase for deontic *should*, as noted in § 4.3.3.1, and seen in Figure 21, is mainly influenced by the gradual increases of both high- and median-degree obligation over the entire time span of the diachronic corpus, with median-degree obligation retaining the larger proportion of the meanings of *should* throughout. Overall, formulaic uses are very infrequent in

comparison with the other two uses. The proportional relationship between the uses that convey high- and median-degree obligation will firstly be considered.

4.3.3.2.1 *Should*: high and median degrees of obligation

The data in Table 33.2 in Appendix 5 show that in Period 0 median-degree uses (11,1 normalised instances) are almost three times more frequent than high-degree uses (3,8 normalised instances) with a ratio of 2,9:1 median relative to high obligation. Median-degree uses (21,5 normalised instances) become proportionally more frequent compared to high-degree uses (4,9 normalised instances) in Period 1, with a ratio of 4,4:1, and remains stable in Period 2, with a ratio of 4:1 (30 and 7,5 normalised instances for median- and high-degree uses respectively). By Period 3, however, median-degree uses (46,8 normalised instances) have lost some ground in comparison with high-degree uses (18,2 normalised instances), with a ratio of 2,6:1. The finding that high and median degrees of obligation have reached near-equal proportions for *must* over Period 2 and 3, after high-degree uses were more frequent in preceding periods, together with the fact that the proportional difference between high- and median-degree *should* also narrows in Period 3, is therefore similar to the tendency of *should*, which, like *must*, also appears to be moving toward a degree of polysemy on a micro-level in its deontic meaning. In other words, both *must* and *should* express similar proportions of high- and median-degree obligation by the end of the 20th century than previously, which in turn indeed suggests a move toward a degree of synonymy between these two modals in the deontic meaning (see also § 4.3.2.2.1), supporting the second part of Bowerman's claim regarding the interchangeability between *should* and *must* in terms of weaker obligation meanings in SAfE. As stated in § 2.4.2.2, the tendency for *should* is quite different in AmE and BrE, where *should* progressively expresses weaker deontic meanings over the course of the 20th century, and, specifically, weaker meanings than *must*. The trend towards micro-level monosemy for deontic *should* in other native Englishes does therefore not match the trend toward micro-level polysemy for deontic *should* in SAfE.

Table 13 presents the results for the further microsemantic analysis based on the source of obligation for high-degree *should*, to assess whether there is a tendency

for a particular degree of strength to co-occur with a particular source. As seen in § 2.4.1.1, the major distinction regarding a propensity for one modal or another to co-occur with a particular source is mainly made between *must* and *HAVE to*. However, no clear pattern for *should* appears to be present in this regard; Collins (2009a:45) notes that "[d]eontic *should* may be subjective, indicating what the speaker considers desirable, appropriate or right ...or objective, where the appropriateness or desirability of the course of action described stands independently of the speaker's endorsement". This part of the analysis will therefore also add an evaluation of this constituent in the micromeanings of *should*, especially regarding a comparison with *must*.

Table 13

Percentages of high-degree deontic uses with subjective, objective or indeterminate sources of obligation in *should*

Period	Subjective	Objective	Indeterminate
0	20	80	0
1	28,6	71,4	0
2	50	50	0
3	14,3	82,1	3,6
Average	28,2	70,9	0,9

On average, *should* with a high degree of obligation is 42,7% more frequently paired with objective sources than subjective sources. The frequency of objective sources remains higher than subjective sources throughout, with Period 0 and Period 3 having the highest objective frequencies, except in Period 2, where there is an exactly equal frequency between the two sources. Indeterminate sources account for very little in Table 14, and were only found in Period 3. The ratio between subjective and objective sources with a high-degree of obligation narrows from 1:4 (subjective relative to objective uses) in Period 0, to 1:2,5 in Period 1. As suggested above, the proportions between these uses are exactly equal with a ratio of 1:1 in Period 2, but in Period 3 the relative difference is even larger than it was in Period 0, with a ratio of 1:5,7. When considering these findings together with the late 20th-century trends and the average trends for high-degree *must* in this regard, a clear tendency arises. In Period 3, i.e. by the 1990s, high-degree *should* co-occurs more frequently with the objective source

than *must* (82,1% for *should* compared with 66,1% for *must*), and the same applies for the average scores (70,9% for *should* and 55,3% for *must*) (see Table 8). High-degree *must*, on the other hand, appears to prefer the subjective source more so than high-degree *should* by Period 3 (30,4% for *must* and 14,3% for *should*), as well as on average (41,9% for *must* and 28,2% for *should*). According to the average percentages and the percentages for the 1990s, it is evident that there is a link between the source and obligative strength as far as high-degree uses for *must* and *should* are concerned: for *should* objective sources tend to link with a high degree of obligation (see e.g. example [273]), whereas for *must* subjective sources tend to link with a high degree of obligation, when compared to high-degree *should* (see e.g. example [195]) (even though high-degree *must* in itself tends to link with the objective source [cf. § 4.3.2.2.1]). When the diachronic tendencies for median-degree obligation are compared in Table 14, a different picture arises.

Table 14

Percentages of median-degree deontic uses with subjective, objective or indeterminate sources of obligation in *should*

Period	Subjective	Objective	Indeterminate
0	24,1	75,9	0
1	36,1	63,9	0
2	57,1	42,9	0
3	54,3	45,7	0
Average	42,9	57,1	0

The pairing of median-degree obligation with subjective and objective sources of obligation is on average more evenly distributed between these sources than for high-degree uses, with only a 14,2% difference in the favour of objective uses. This reflects the tendency in Period 2 and 3, but before these periods median-degree *should* more often paired with objective sources. In Period 0 the ratio for subjective relative to objective uses is 1:3,1, and, like for high-degree *should*, the ratio narrows to 1:1,7 in Period 1, but the proportion basically reverses in Period 2 with a ratio of 1,3:1 in the favour of subjective uses. This tendency remains relatively stable in Period 3 with a ratio of 1,2:1, still in the favour of subjective *should*, but, as mentioned above, the

proportions are very close in the last two periods. This means that median-degree *should* does not have such a clear tendency to be paired with one source or another as is the case for high-degree *should*. With the subjective frequency being only very slightly higher than the objective frequency in Period 2 and 3, no real conclusion can be made as to median-degree uses preferring to co-occur with the subjective source, but rather that median-degree *should* has equal propensity to co-occur with both sources. When compared with the findings for median-degree *must* in Table 9, the average percentages for median-degree *should* and median-degree *must* are very close for both objective (57,1% for *should* and 53,5% for *must*) and subjective sources (42,9% for *should* and 44,2% for *must*), even if median-degree *should* is on average slightly more frequently paired with objective and median-degree *must* with subjective sources. A slightly clearer tendency is discernable in Period 3, although frequencies for median-degree *should* and *must* remain similar, where median-degree *should* is somewhat more frequently paired with a subjective source than median-degree *must* (54,3% for *should* and 35,9% for *must*), and median-degree *must* is somewhat more frequently paired with objective sources (54,7% for *must* and 45,7% for *should*). Here the link between the strength of the obligation and the source of the obligation for *should* and *must* is not so clear as for high-degree uses. A graphic representation of these comparisons can be seen in Figure 32 in § 4.3.5. By this account, these two modals share an even more specific kind of microsemantic synonymy: not only do *should* and *must* both convey more equal proportions of high- and median-degree uses by the 1990s than previously, but the median-degree uses of both these modals tend to co-occur with objective and subjective sources with near-equal propensity, both on average and specifically in the 1990s. Subsequent paragraphs will illustrate the occurrence of objective and subjective sources of *should* with its high- and median-degree uses over time.

The very frequent use of high-degree *should* with objective sources in **Period 0** was found to be linked to its use in legal or regulatory contexts (P6) in news and non-fiction, as seen in (243), (244), (245) and (246). In such objective uses the "appropriateness or desirability of the course of action described stands independently of the speaker's endorsement" (Collins, 2009a:45), since it is seated in an external rule, regulation or manifest of some sort. In the context of the reported speech of news and non-fiction, the obligation expressed in the latter three examples is technically grounded in the past, but despite this, the obligation is grounded in the future in the

context of the discourse that is reported (p4), and therefore has a high degree, as will be shown in the discussions below.

(243) Shipowners **should** be aware, that by Act 6 Geo.IV. cap.41, stamp duties are made no longer payable upon any deed of sale, or any conveyance, assignment, or other deed or instrument whatever... (1820s; News)

(244) On Friday afternoon last, the Bishop of Cape Town held a confirmation in St. George's Cathedral. After the usual service the Bishop addressed a solemn exhortation to the candidates... He also spoke of the good which might henceforward be done by them, referring especially to the efforts which **should** be put forth in aid of the Missionary work of the Church, the want of which was at the present time greatly felt. (1860s; News)

(245) The last advice which we have received, represent the state of affairs in Ireland to be still critical. It appears that the Catholic association having been put down by Parliament, has resolved to new model itself in such a manner, as to avoid direct breach of the act, and at the same time, to continue its efforts for obtaining emancipation. Instead then of meeting as heretofore for the purpose of pointing out to the Catholic body the means which they **should** pursue – instead of levying the Catholic rent, and directing its application, it is declared that a certain committee will in future meet for charitable purposes, but that of course digressions cannot be avoided nor legally objected to. (1820s; News)

(246) "But the interests of human happiness must clearly require that the relief which we afford them should be scanty. We may ...take upon ourselves ...to mitigate the punishments which they are suffering from the laws of nature, but on no account to remove them entirely. They are deservedly at the bottom in the scale of Society, and if we raise them from this situation we not only palpably defeat the ends of benevolence but commit a most glaring injustice on those who are above them. They **should**, on no account, be enabled to command so much of the necessities of life as can be obtained by the worst paid common labourer. The brownest bread, with the coarsest and scantiest apparel, are the utmost which they **should** have the means of purchasing". (1820s; Non-fiction)

In example (243) *should* is part of the matrix clause, which draws attention to the legal matters that is discussed in the subordinate clause. Hence, *should* has a high-degree of obligation due to its reference to legal obligation (P6) in the form of an *Act*, which renders the situation face-threatening and without possibility of non-compliance, and, indeed, with much gravity in the event of non-compliance (P5). The regulatory or legal nature of the text is also made clear by the use of the word *duties* and the phrase *deed of sale*. Examples (244) and (245) are both about church business. The source of the obligation in (244) tends toward the objective end of the continuum, signalled by the passive construction in the relative clause, which is in turn part of a non-finite subordinate construction, *referring ...to the efforts which should be put forth*, because, despite the fact that the utterance is made by a discourse-internal source in the context of the reported situation, *the Bishop of Cape Town*, his utterance is in fact now reported by a newspaper, rendering the source external to the current discourse. Furthermore, these constructions do not clearly indicate that the Bishop himself is actually imposing the obligation – he 'refers' to it, but it appears to be necessitated by circumstances, i.e. the obligation is independent of the Bishop's endorsement, signalled by *the want of which was ...greatly felt*, which further sways the example toward the objective end of the continuum. The high-degree of obligation is in turn indicated by the fact that the utterance is made in a context where a figure of ecclesiastical authority (P1) addresses *the candidates* (apparently young people at their confirmation ceremony), which indicates the impact of social norms in the situation (P5), despite the passive construction that would normally be a parameter for median obligation, but, as shown above, the passive is probably part of a technique in another discourse – that of the report itself. It is however not absolutely clear what the exact direct speech of the Bishop had been, and so this instance is another example of the scalar quality of these analyses according to the parameters, in that they usually cannot be treated in black and white terms. This is indeed not a prototypical example of the objective, high-degree use, as (243) more clearly is. The stronger impact of the suggestion is however supported by the apparent seriousness of that suggestion, indicated by *solemn exhortation*, which implies gravity in the event of non-compliance (P5). Example (245) is, on the other hand, more prototypical of the objective high-degree use than (244), despite also being part of a reported discourse context, because the utterance involves what is *resolved* upon by *the Catholic association* in the midst of the religious and political tension in 19th-century Ireland,

suggesting its regulatory and legal nature (P6). Here *should* expresses an obligation imposed by a previous regulatory model or manifest of the association in question, rendering it a high degree. *Should* is therefore not really used here as part of the current 'declaration', but rather refers to an older obligation – which the association in question now withdraws, before it is *declared* what will be done, as part of a new *model*, designed to uphold laws, i.e. *to avoid direct breach of the act* that has been made by an authority (P1 and P5) that trumps the Catholic association, namely the Irish *Parliament*. This suggests that the previous obligation *should* conveys was also part of a legal declaration. Example (246) is taken from a report of the address by The Rev. Dr. Philip at the Anniversary Meeting of the Subscribers to the Settlers' Fund Society, concerning charity to the poor and a former report made by a certain Mr. Malthus, apparently a great, rigid economist of high social standing (P1), as is mentioned earlier in the text the extract is from. In the context of his discourse, Mr. Malthus refers to obligations via *should* as part of a proposed manifest of the society in question, which indicates its regulatory nature (P6). Much like example (244), it can be argued that he is the subjective source of obligation. This example was however analysed to tend toward the objective end of the cline, since the obligation seems, once again, to be necessitated by circumstances, as the speaker also argues, i.e. it would apply whether the speaker endorses it or not, and the obligation is therefore just conveyed, but not necessarily imposed, by the speaker. This is of course an argumentative technique to convince his audience. The objective necessity for the obligations to be fulfilled is expressed in the phrase *the interests of human happiness must clearly require that* and in the reference to *the laws of nature*, which both distance the obligation from the speaker. The analysis is reinforced by the use of the passive in the first instance of *should* in this example – *they should ...be enabled to*, which does not make it clear who should 'enable them', i.e. who should enforce the obligation. The speaker however does make quite an emotional appeal (P5), indicated by adjuncts and expressions such as *clearly*, *on no account*, *entirely*, *deservedly* and *palpably*, which all contribute to intensify the force of the obligation (P2), and indeed is very face threatening (P5) to the beneficiaries of the obligation, namely the poor in question. Adjectives (P2) expressing disdain, as in *worst paid common labourer* together with superlative forms of adjectives in expressions such as *most glaring, brownest*, *coarsest*, *scantiest* and *the utmost* further reinforce the analysis for a high degree of obligation.

Objective median-degree *should*, which is also more frequent than median subjective uses in Period 0, is illustrated below.

(247) He would press upon the Directors to consider whether their present issue was not larger than necessary. If they looked abroad they would perceive the spirit of speculation which was afloat; a spirit which **should** be discouraged, but which a large issue tended to keep alive. He was himself of opinion, that the issue was larger than the wants of the country required (1820s; News)

(248) Thus remarks a paper before us:– ... Why, it may be asked, in the name of common sense, **should** the Hottentots be made a separate people? Why **should** a special law be made, giving them alone of all the other inhabitants of the colony, a perfect right to follow the bent of those vagrant habits, which are the last of strongest links of the chain that binds a people to heathenish and savage customes? Why **should** they be picked a culled from amongst civilized society, and thrown back upon their own resources? Why **should** all their old prejudices be kept alive? (1830s; News)

In example (247), objective *should* expresses median obligation, since the reported utterance is made by a Mr. Young, apparently the secretary of the bank in the context of a large Bank meeting, as preceding information in the text discloses, to *the Directors*, which would be above him in rank (P1), indicating that a person of lower professional rank is speaking to persons of higher rank than himself. The utterance is therefore not threatening to the face of the 'directors', and leaves them with the possibility of non-compliance (P5). Furthermore, he is giving them his *opinion,* which further renders it median obligation (P6), despite the reporter using the phrasing *would press upon*, where *press* may strengthen the force, but directly after this, the phrase *to consider* weakens the force again. Once again, much like some of the examples of high-degree objective *should* discussed above, this is not a prototypical example of an objective source, since Mr. Young is internal to the reported discourse, but the use of the passive distances the obligation from him, seeing that it does not make clear who or what would enforce the obligation. Example (248) on the other hand illustrates the quite dense collection of objective median-degree instances of *should* in idiomatic or rhetorical *why*-questions (cf. Collins, 2009a:46) that are found

in news during this period. The uses in (248) tend toward the philosophical and hypothetical (P6), according to their rhetorical nature, since no immediate compliance with the obligation, which these questions imply, is required. There is therefore a possibility of non-compliance (P5), rendering it a median degree of obligation, despite some parameters for high-degree obligation being present, such as the emotional expression *in the name of common sense*. Furthermore, these questions "can be paraphrased as 'Is x really necessary?', with the implicit answer being 'no'" (Collins, 2009a:46). Yet the implied obligation of e.g. 'the Hottentots should not be made a separate people' in the first instance of *should* in (248) remains based on opinion, also rendering it a median degree of obligation (P6). The same applies to all the uses of *should* in this example. This contrasts with the high-degree objective uses above, where the utterances were made as part of official, legal or regulatory addresses, manifests or legislation, whereas in the median-degree objective uses no such context is present.

Where deontic should is subjective, the speaker indicates what he/she deems "desirable, appropriate or right" (Collins, 2009a:45). The only two instances of subjective, high-degree *should* in Period 0 are illustrated below.

(249) And it is very important how any not helping it will be for you and Mr. XXX[192] to consider whether it might not be well for you to XXX a native female with you who can XXX Kaffir ...you **should** at once set her to work under Mr. XXX. (1830s; Social letters)

(250) We would resist the tyranny – we would be trampled on no farther – we would assert our rights... The money, whether it wore the beggarly six thousand or six millions a-year, **should** be granted without robbing us of our rights, peace, and honour, or we would never suffer it to leave the Exchequer. (1820s; News)

Example (249) is an extract from a letter containing advice to a new missionary, apparently by someone on the mission council, although it is unclear exactly who. The addressee is simply addressed as 'Brother', which is probably an ecclesiastical

[192] The symbol 'XXX' indicates a word that is illegible from a handwritten text.

address, and not a familial one. This indicates that there is a degree of social distance between the interlocutors (P5) and hence that the writer may have authority over the subject (P1). The obligation is further strengthened with the use of the second-person pronoun *you* as the animate subject (P1) and the use of an agentive verb phrase *set her to work* (P1). The indication that the writer is interested in the action being performed, as indicated by the adjuncts *important* and *at once* (P2) further contributes to the analysis for a high degree of obligation. These above-mentioned parameters operate on the point where the strength and source continua interact to produce a prototypical or core example of a modal with strong illocutionary force, like *must* (see § 2.4.1.1; cf. Coates, 1983:33-6). This shows that even traditionally weaker *should* can be semantically strong in some contexts. There is however one difference between this example and prototypically strong *must*: this example has a slightly polite tone, indicated by *to consider whether it might not be*, which renders it somewhat more tentative feel (i.e. one parameter for median-degree obligation, but not enough to sway the analysis) than a prototypical example of high-degree, subjective *must*, as in (168), for example. Example (250) is from the same text as (245), which was also analysed as a high-degree example, and is part of a review of a parliamentary session in Ireland. Here the interpretation of the subjective source is not prototypical, given the passive construction in *should be granted*, but it is clear from preceding clauses that the elided subject of the clause containing *should* is the plural first person, or in more precise terms the clause can be written as *the money ...should be granted [us]*, or *[we] should be granted [the money]* to the same effect. Here the appeal is made to the parliament by one of the members, as the context shows, so it is evident whom the obligation is imposed upon, seeing that the passive construction in this case is more a rhetorical device in the context of a parliamentary speech than an indication of an objective source. The high-degree of obligation is in turn revealed by the emotional motivation (P5) as shown by the use of noun phrases such as *tyranny*, *our rights*, *peace and honour*, verb phrases such as *resist, trampled on*, *assert*, *robbing* and *suffer to leave*. Some adjuncts (P2) also contribute to the emotional tone, like *beggarly*, *no farther* (with negative polarity [P4]) and *never* (conveying a high degree of usuality). Also, the clause *or we would never suffer it to leave the Exchequer* indicates that there would be a consequence, i.e. gravity, in the event of non-compliance (P5).

The subjective uses of median-degree *should* in Period 0 are mostly found to have first-person subjects, as illustrated in (251) and (252).

(251) I **should** have mentioned that wen [sic] Mr. Wait went to Cape Town wile [sic] we laid in Simon's Bay he ingaged [sic] a Mr. Roades [Rhodes] a schochman [sic] a verey [sic] respectable man... (1820s; Fiction/narrative)

(252) THE patience, or – perhaps we **should** speak more correctly to say – the apathy with which the Inhabitants of this colony have submitted for a series of years to the goulest detraction and calumny must be fully known to be credited. (1830s; News)

The non-actualisation of, or non-compliance with (P5), the obligation expressed by *should* in (251) lowers its deontic force, rendering it an example of median obligation, despite its first-person subject. This particular past construction (P4) with a deontic meaning is not possible for *must*, seeing that *must have* expresses epistemic modality, making *should* more likely to occur in such contexts. In this context *I should have mentioned that*... is almost a writing formula, weakening the force even more. The same applies to the use in (252), in that it has a formulaic character with very little deontic force. The obligative force is further weakened by the low-degree probability adjunct *perhaps* (P2), which is used in search of the 'correct' word for the situation.

High-degree uses of *should* with objective sources, which are still prevalent in **Period 1**, are illustrated below.

(253) The Governor is pressing hard to get you and the members of the Executives of the two Chambers up. At the beginning of the month I submitted the enclosed proposal as to the order in which people **should** be allowed up. ... I am convinced that they will be obliged to allow people up very soon, as we cannot go on like this indefinitely... (1900s; Social letters)

(254) The only condition that **should** be insisted on, on behalf of the public, is that it **should** not establish for itself an unjust monopoly of mineral and other rights... ... If such a safeguard is secured, we **should** hall the establishment of such a Company; but the utmost vigilance **should** be used by both the Colonial Government and the public, to prevent the RUDD-RHODES Syndicate from filching away on any pretence the vast treasure-deposits of Matabeleland, which belong to British subjects at large. (1880s; News)

(255) By the terms of this deed the assets were to be realized, and to be applied: ... Thirdly,– That Adler and Co. **should** accept the balance in satisfaction of all claims due to them, and **should** release the plaintifs from all further liability. (1870s; News)

In (253) the high-degree of obligation is signalled by the fact that is written in the context of official war business, concerning a *proposal* of what is *allowed*, i.e. it is regulatory (P6), and apparently endorsed by *The Governor*, a figure of authority (P1), who *is pressing hard* for the fulfilment of the obligation. The subsequent sentence indicates the writer's conviction that the obligation will in fact be *obliged* or complied with (P5). The objectivity of the source is indicated by the passive voice, as it is not clear from the text who *they* – the ones that should comply with the obligation – are. The motivation of the obligation also comes from circumstances, as the writer states a reason for referring to it, namely that *we cannot go on like this indefinitely*. In example (254) the first and fourth uses are deontic. The second use is a mandative construction of the quasi-subjunctive and the third use is carries a preterite meaning. The two deontic instances are objective, which for both is signalled by inanimate subjects and the passive voice. The analysis for a high degree of obligation in the first instance is supported by the verb *insisted*, indicating the intensity of the obligation, and by the fact that *should* is part of the relative clause that postmodifies the noun *condition*, which further indicates that there is a degree of gravity in the event of non-compliance (P5). The negative polarity (P4) in the subsequent sentence containing the mandative construction, as well as the adjunct *unjust* (P2), further contributes to the analysis in the surrounding context. The second instance of *should* has a high-degree of obligation, since an emotional tone is apparent (P5) – reinforced by the adjunct *utmost* (P2). The verb phrases *prevent* and *filching*, and the noun phrases *vigilance* and *pretence* also add intensity to the obligation, as they point to gravity of non-compliance (P5). Example (255) illustrates more prototypical objective high-degree obligation, in that it is purely legal (P6), with very little possibility and much gravity of non-compliance (P5), and naturally a face-threatening implication (P5). The legal nature of the obligation is shown in the noun phrases *terms*, *deed*, *assets*, *claims*, *plaintiffs* and *liability*. The lack of a possibility of non-compliance is also supported by the verb phrases *accept* and *release*.

As noted above, median-degree uses of *should* with objective sources are also more frequent than with subjective sources in Period 1, as they are in Period 0. Such uses are often found in passive constructions in news, as exemplified below.

(256) Great vigilance **should** have been observed in case the shop unduly neared the land. No such vigilance was, however, exercised; there was no look out, neither was there an officer on the bridge. (1870s; News)

(257) In Natal one sees raw brick houses and sod houses (build of sods of very dense shining pot clay), coated outside very cheaply with coal tar, then whitewashed with quick lime. What proportion of salt and water **should** be added to quick lime to whitewash a house outside or inside? (1880s; News)

(258) What **should** be the basis of representation? It **should** be in no respect narrower than that now existing in any of our States. ...it **should** be adult franchise, with a high educational test. (1990s; News)

Example (256), much like (251), expresses the non-compliance with a past obligation (P5), and criticism at its non-fulfilment, indicated by the adjunct *however* (P2), which lowers its degree of obligative strength (cf. Collins, 2009a:45-6). In the verb phrase *should have been observed* the perfect aspect grounds it in a past time, and the passive construction renders its source objective. Example (257) is an informative piece in a newspaper titled 'Building with Raw Brick'. Here *should* is used objectively as part of a rhetorical question, as the answer is never given later in the piece. In the context of the question, the obligation has quite a neutral degree of force, since the answer, e.g. 'the proportion should be...', would imply a suggestion (P6) or instruction with a neutral or unmarked value, which *should* traditionally conveys, considering its lower degree of force. It does therefore not involve social norms and is not face threatening in any way (P5). In this example the force of obligation is therefore much like that of *must* in the examples from spoken contemporary SAfE in (211) and (214), which make them all the more interesting. Another question is found in (258), but this time with an answer – with both containing *should*. This use is however not neutrally instructive, but is grounded in hypothetical circumstances (P6), and is based on an opinion (P6), even if the source is objective in the context of a newspaper piece,

which is signalled by the impersonal pronoun *it* (P1). Here indeed the clines for weaker force and an objective source intersect, rendering it a prototypical example of obligation at the lower end of the scale. This remains based on the writer's opinion, despite the phrase *in no respect*, which is used by the writer as part of the argumentative style of the piece.

The less frequent subjective high-degree use of *should* in Period 1 is illustrated in (259).

(259) In my petition Dutch would sign on Saturday since forgathering at kirk with ministers on Sunday they all say Jim **should** be hanged! – and many add, Rhodes too; and some add, Hercules Robinson too; and a few add, Imperial Government altogether. (1890s; Social letters)

This example was interpreted to have a higher degree of strength on the continuum, although it is not at the strongest end, owing to its emotional tone (P5) in the context that is being reported, which is reinforced by the exclamation mark at the end of the clause containing *should*. This could be either an indication of the writer's surprise or alarm at the reported speech, or of the emotional tone with which the obligation was conveyed in the first place. The writer's alarm would also imply the emotional value of the utterance, so this parameter applies in both cases. Furthermore, the obligation comes from *ministers*, who are the subjective sources of the obligation with high social status (P1) in the context of promoting the death sentence, which implies the role of social norms (P5), and a very face-threatening situation for *Jim* (P5).

Examples (260) and (261) illustrate the equally infrequent subjective median-degree use, compared to the objective median-degree use, during this period.

(260) You **should** hear Mrs Brown, who was at Lovedale when you were at Katberg, on the subject of Mina. She is in love with her all over. It was a great pleasure to me to listen to her high praise. (1880s; Social letters)

(261) I believe policy **should** be to keep shooting for extreme cases, as heavy fines and prison serve the same end and makes less "martyrs" and to to keep hanging for those cases where there is a taint of trechery or crime... (1900s; Social letters)

Should in (260) has very little illocutionary force, as it purely implies an act of socialising (P6) or gossiping in a friendly situation (P5), namely a private letter to a friend. There is nothing face threatening about the obligation expressed in this example (P5), despite the presence of a second-person subject and the interpersonal content of the letter, which would normally be possible indicators of a higher degree of obligation. The level of familiarity between the two correspondents, apart form the opening and closing of the letter, can also be seen in this extract in the mutual knowledge of several persons, places and situations. The analysis for median obligation in example (261) is firstly supported by the indication that the opinion (P6) of the subjective source is being given in the matrix clause, viz. *I believe*, which *should* in the subordinate clause is complement to. Secondly, the fact that the correspondence is friendly (P5) and does not threaten the face of the addressee (P5), but rather conveys some philosophical considerations (P6) about a moral topic, the death penalty, further indicates that the obligation does not have a high degree of social impact at all. Thirdly, the circumstances discussed are hypothetical (P6), which is supported by *for extreme cases* and *for those cases*.

As mentioned above, equal amounts of objective and subjective sources were found to co-occur with high-degree *should* in **Period 2**. The following examples illustrate the objective high-degree use, e.g. in the passive voice, which was mostly found to be regulatory in nature. Most of the examples provided for Period 2 are from letters, either social or business, since this is the register with the highest occurrence of deontic *should* during this period.

(262) As the enclosed memorandum shows... If the expansion of Native Education is to be based upon the poll tax, then so much of the poll tax **should** be available as the budget requires. (1920s; Business letters)

(263) The clerk read the letter, and explained to Baden how he **should** go about sending a remittance to his languishing wife and family. Paper money **should** be put in the letter, and upon the envelope a blue cross **should** be made, to signify that its contents were valuable. (1930s; Fiction/narrative)

In (262) the regulatory content of the text (P6) is clear from the fact that it is part of a report on an *enclosed memorandum* to the letter, and *should* is here indeed used as part of the expression of a regulation, which is implied by the conditional construction, *If ...then*, by the verb *requires* and by the phrase *so much*, where 'much' is a pronoun conveying the quantity 'required' by the obligation. The official nature of the obligation is further reinforced by the mention to the issue of *Native Education*, *poll tax* and *the budget.* The three instances of *should* in (263) convey obligation based on postal regulations (P6). *The clerk* is therefore not the source of the obligation, although he conveys it to *Badeni*, but the regulations are, rendering it objective. The obligation itself has a high degree of strength, seeing that non-compliance would have serious implications (P5), namely that the letter containing the money would never reach the *languishing wife and family*, and they would ergo not receive their relief.

The next three examples show the use of high-degree *should* with subjective sources in Period 2.

(264) I don't care if you never come top of the class, but you **should** be in the first five. (1950s; Social letters)

(265) It is important to know what is worthwhile in life so far as you are concerned. Health, friendship, and self-respect are more important than money, glory, or titles. The former **should** never be sacrificed for the latter, nor **should** greed or acquisitiveness make you lose your sense of proportion. (1950s; Social letters)

(266) "You made a mistake. No one shouted "bloody swine". ... "All right, then I **should**n't have shouted at you from my car." (1950s; Fiction/narrative)

In example (264) the subjective source a father writing to his son, so a degree of social distance (P5) and authority (P1) can be detected. Here the father states his expectations for his son's performance at school, which is a face-threatening topic coming from a parent (P5). A slightly emotional tone (P5) is also detectable in the expression *I don't care if....* Furthermore, the coordinator *but* activates the contrast between *never* (a modal adjunct of polarity with a high-degree of usuality) and the implied 'always' (the strong counterpart of 'never'), i.e. the son is expected to 'always' be *in the first five* (P2). This in turn implies gravity of non-compliance (P5).

Many other parameters for high-degree obligation are also present here, like the second-person subject and a projection into future time (P1). The high degree of obligation in (265) is also signalled by the use of *never* (P2), but also by its complement *nor*, as well as by the adjective *important*. This implies gravity in the event of non-compliance (P5). A somewhat emotional tone (P5) is also taken by the subjective source, who addresses the addressee in the second-person (P1), as is discernible from the verb phrases *be sacrificed* and *lose*, as well as by e.g. the noun phrases *self-respect, greed* and *sense of proportion*. Despite the slightly philosophical nature of the letter, the parameters in favour for a high-degree analysis are dominant in this case. Example (266) expresses self-criticism on an implied obligation (e.g. 'I should not shout at you from my car') that clearly had not been fulfiled, but despite this, the obligation remains of a high degree, seeing that the gravity of non-compliance is realised (P5) in the concession *All right*. The sketched situation, being an argument between two strangers, also appears to be quite tense, i.e. emotionally charged (P5), and indeed face threatening (P5), as shown from the insult *bloody swine* that was in fact *shouted* at a passerby just before the utterances in this extract take place. The speaker in this example is therefore lying when he says that *No one shouted [it]*, which implies that even the concession made by the speaker is false, rendering the situation even more face threatening.

Median-degree *should* co-occurring with subjective sources, being slightly more frequent than median-degree *should* co-occurring with objective sources, and mostly occurring in the letters register, is exemplified in the next three extracts from letters, as well as one from non-fiction.

(267) I personally think that you **should** give it your most careful consideration and see whether it could not be published cheaply, as he suggests... (1930s; Business letters)

(268) I **should** like one day to discuss with you why it is that the word 'absconder' looms so large. I feel sometimes that it **should**n't. I feel almost displeased that it **should** be so... (1930s; Social letters)

(269) Art bows to the dictates of fashion. What was considered good taste a century ago may be considered vulgar today. The reverse may also apply. My opinion

is therefore just an opinion and **should** not be accepted as dogma or authoritative. (1950s; Social letters)

(270) In their view, the Afrikaner **should** deliberately strengthen himself in social groups where hitherto he had been weak – as for instance the world of business. (1940s; Non-fiction)

Median obligation in (267) is signalled by the fact that the obligation that is imposed by the subjective source in the first person is based on his opinion (P6), as seen in the hedge (P5) *I personally think*. In fact, the obligation further appears to have come from another source, *he*, in the first place, which too carries less force by only being 'suggested' (P6). Despite some indicators of a higher degree of deontic force, such as the superlative from of *most careful*, the hedging (P5) found in *and see whether it could not be...* renders it a median degree. In (268) the first instance of *should* has a preterite (formulaic) meaning, whereas the third instance of *should* is an emotive construction of the quasi-subjunctive meaning. The second instance carries the deontic meaning, which is interpreted as having a median-degree force, because it too is based on personal opinion (P6), signalled by *I feel*. The tentativeness of the obligation is further shown by the use of the adjuncts *sometimes* (conveying a median degree of usuality) and *almost* (P2), despite the emotional tinge expressed by the adjective *displeased*, which is essentially part of the emotive, quasi-subjunctive construction. Much the same applies to (269), where the obligation too is based on the *opinion* of the subjective source (P6). Essentially, the obligation concerns this opinion. The writer implies the lower degree of the deontic force when he denies his 'opinion' on *Art* and *fashion* to be *dogma* (i.e. objective and high-degree) or *authoritative* (i.e. subjective and high-degree), seeing that *The reverse may also apply*. Furthermore, the content is philosophical in nature (P6) – also supporting the analysis for median obligation. The obligation in (270) is also based on opinion (P6) from subjective sources, which is being reported in the text. This is reinforced by the writer's statement that the obligation stems from *their view*. Median obligation is further indicated by the hypothetical nature of the reported utterance (P6), in that generic terms are used, namely *the Afrikaner, himself* and *he*.

Examples (271) and (272) show the median-degree use of *should* with objective sources in Period 2.

(271) In the large class of cases of which this is a type, the functions of good administration and good citizenship **should** combine to prevent the prosecution of innocent Natives. (1910s; Non-Fiction)

(272) The matron **should** be a motherly woman of fair age who would see to the cooking and domestic affairs generally. (1910s; Business letters)

The hypothetical nature of the text in (271), as indicated by the expression *In the large class of cases...*(P6), renders it an example of median obligation. There is no rule or regulation that underpins the obligation here, but it is rather a suggestion (P6) about what needs to be done to meet a certain end, shown by the infinitive construction *to prevent*. Circumstances therefore necessitate the obligation, contributing to its objectivity. These parameters overrule the presence of an emotional tone, conveyed by e.g. the adjuncts *good* and *innocent*, and the noun *prosecution*. *Should* in example (272) conveys an even more directly hypothetical (P6) meaning, where the desirable attributes of a potential *matron* are listed, which also renders it an example of median-degree obligation.

Objective, high-degree uses of *should*, being so much more frequent than subjective high-degree uses in **Period 3**, as seen in Table 13, still mostly convey regulatory meanings, much like in previous periods. This suggests that the propensity for *should* to be used in such contexts has increased over the 20th century, and has returned to, and even exceeded, the proportions of subjective and objective uses in Period 0. The regulatory nature of these uses can be seen in the below examples. The bulk of the examples for Period 3 are from news, fiction/narrative and non-fiction, these being the registers that all show an increase in the frequency of deontic *should* over the 20th century.

(273) ...citrus oils such as in eau de cologne, are more susceptible to heat and oxidation, and **should** therefore be processed within a shorter period. (1990s; Non-fiction [W2A-023])

(274) Students are also required to keep a log book in which they must detail their weekly activities. The log book **should** be signed and dated by the employer at the end of each month. (1990s; Business letters [W1A-018])

(275) If you do not already have a will, you **should** make one immediately – dying without one always leaves complications… (1990s; Non-fiction [W2D-004])

As part of an academic piece on perfumery, the obligation in (273) essentially houses a regulation (P6) involving a chemical reaction, which, if not complied with, will in all probability have serious (chemical) consequences (P5), which is indicated by the adverb *therefore* (P2), rendering it a high degree. Ergo, in this instance *should* is basically part of a warning. Example (274) involves a regulation for students (P6), as is evident from the verb phrase *are ...required*. It is interesting to note the use of *must* in this context, which is also objective with a high degree of obligation, and essentially conveys the same degree of force as *should* in this example. In this particular context these two modals are therefore used interchangeably, which indicates a level of synonymy on the high-degree obligation end of the strength continuum as well, since this was already reported for objective an subjective uses on the median-degree end of the continuum in § 4.3.2.2.1 (see examples [183], [184], [186], [194] and [197]). Example (275) was interpreted as having an objective source of obligation, seeing that, in the context of the non-fiction text, the generic *you* is addressed as a rhetorical strategy in administrate/regulatory writing, specifically in an informative piece on wills. A specific person is therefore not addressed, and is not addressed by a specific person. This is reinforced by the conditional construction *If you do not ... you should*, which grounds the motivation for the obligation in circumstances, as obviously the it would not apply if the generic addressee *already [had] a will*. Naturally the legal or regulatory topic (P6) flags the instance for high-degree obligation, but the adverbs *immediately*, which conveys a sense of urgency (P2), and *always*, which conveys a high degree of usuality also contribute to the analysis. Furthermore, the fact that it is mentioned that non-compliance with the obligation would have consequences, here described as *complications* further renders the obligation a high-degree.

Example (276) illustrates the subjective high-degree use of *should* in Period 3, which is the period with the least frequent occurrence of such instances, as mentioned previously. Out of the four raw instances of subjective high-degree uses, two are from fiction/narrative, one from news and one from non-fiction. An example of this use in fiction/narrative is given below.

(276) I was taught that a panty is the most intimate thing ... my mother told me no one else **should** even see my panty. (1990s; Fiction/narrative [W2F-002])

Here the source of obligation is the speaker's *mother*, which indicates status and authority over the subject (P1), and the situation involving a moral regulation is clearly governed by social norms (P5), which is reinforced by the mention of a piece of underwear, the speaker's *panty*, as well as the superlative form of the adjective in *most intimate*. The high degree of the obligation imposed here is further flagged by the use of the verb phrase *was taught*, as well as the phrase *no one else*, indicating restrictiveness, and the adjunct *even*, conveying an exceeding level of counterexpectancy (P2).

Objective, median-degree *should* remains only slightly less frequent than subjective, median-degree uses of this modal, the former of which is exemplified below in the registers these uses frequently occur, namely news and non-fiction.

(277) Such a strategy, which is presently being developed, **should** not only clearly redefine the role of the police in combating crime but **should** also make others in the private and public sectors co-responsible for combating crime. (1990s; News [W2C-002])

(278) For optimum odour stability, emulsifiers with aleate or ricinoleate groupings **should** preferbly be avoided, as exposure to these emulsions to air often results in rancidification. (1990s; Non-fiction [W2A-023])

(279) American ambassador James Joseph spoke enthusiastically at the occasion of his country's Independence Day yesterday of developments in South Africa. He said people of both countries **should** be celebrating democracy and its spread around the globe. He did not mention Russia in that context. Perhaps he **should** have. (1990s; News [W2E-001])

In example (277), the content is hypothetical (P6), as shown in the predeterminer *Such*, as well as in the relative clause *which is presently being developed*, rendering both instances in this extract examples of median obligation. Example (278) is from the same academic text as (273), which was classified as having a high-degree of

obligation with an objective source, due to its indication of gravity of non-compliance and its regulatory nature. In this extract, however, *should* is interpreted as having a median degree of obligation with an objective source. Its nature is still instructive, but its deontic force is definitely lower, in that this is a suggestion on how to achieve the best results (*For optimum odour stability*), i.e. it is not such a strong warning, as in (273), but rather advice that is being offered (P6). This analysis is strengthened by the use of the comment adjunct *preferably*, which indicates a level of tentativeness, and by the adjunct *often*, which conveys a median degree of usuality (P2). Finally, the verb phrase *be avoided* indicates that there is a possibility of non-compliance (P5), also signalling a median degree of obligation. In (279) the first use of *should* is median subjective, but the second instance is median objective, where criticism is given as to the non-fulfilment with an implied obligation (P5) (cf. Collins, 2009a:45-6). A degree of tentativeness is also suggested by the median-degree probability adjunct *Perhaps* (P2).

The use of median obligation *should* with a subjective source, as illustrated above in the first instance of *should* in (279), is often found in direct speech representation in fiction/narrative texts, and can be also seen in the next few examples.

(280) What do you think, Nomsa? **Should** Beauty accept to go out with Pule or Bhoikei?" (1990s; Fiction/narrative [W2F-014])

(281) Can you believe that? This woman would leave her bath water for me to let it out? I'm not saying she **should** wash the tub. Hey, she's paying me to do that – O.K. (1990s; Fiction/narrative [W2F-002])

(282) One did not attend a funeral in order to pry into the circumstances of a death. "You think we **should** go?" Emma asked in tones suggesting she did not herself consider it at all necessary. (1990s; Fiction/narrative [W2F-012])

(283) "Did I wake you up? You weren't already asleep, surely? A young woman like you, you **should** go out a bit you know. Then again, what nonsense am I talking ... when would you go out since this woman you work for keeps you in her kitchen so late." ... "Maybe you **should** go home and send one of your daughters to help." (1990s; Fiction/narrative [W2F-002])

In (280) *Nomsa* is asked for her advice (P6), as clearly seen in the question *What do you think [?]*, where the hedge *think* in itself weakens the force. The topic of the conversation, namely choices related to dating, as well as the mention of mutual acquaintances on their first names, indicate the familiar and hence friendly nature of the relationship between the speaker and the addressee (P5). In example (281), however, the median degree of obligation is indicated by the negation in the matrix clause *I'm not saying* (in this case a counterproductive parameter), which shows that, essentially, the speaker mentions an obligation she rather does not impose, which weakens the illocutionary force. Furthermore, the implied situation is hypothetical (P6), supporting the analysis for median obligation despite e.g. the emotional tone. The obligation implied in the question posed with *should* in (282) is very plainly median degree, as seen in the preceding and subsequent narrated sentences of the quote. The preceding sentence suggests that the motives of *Emma* and the addressee for going to *a funeral* are not pure, which are here suggested to be *to pry into the circumstances of a death.* The subsequent sentence indicates the intention of the speaker in asking the question, seeing that the narrator clearly mentions her *tone* to be 'suggestive' of her belief that they 'should' not go, as seen in *she did not herself consider it at all necessary*. Hence, the obligation implied in the question has much possibility and all probability of non-compliance (P5). In the question itself, the hedge *think* (P5) shows that the implied answer will be based on opinion (P6), which further supports the analysis for median-degree strength. In example (283) the median obligation of the first instance of subjective *should* is flagged by the apparent friendly context of the utterance (P5), seeing that the speaker shows concern for the addressee in the question *Did I wake you up?*. Familiarity and an informal conversation are also implied in expressions like *surely*, *a bit* and *you know* (P2), rendering the interaction non-face-threatening (P5). *You know* also shows that the speaker is making a suggestion as an act of socialising (P6). When the speaker discredits her own suggestion in the subsequent utterance, shown in *what nonsense am I talking*, by stating the reason the obligation would not be possible to be complied with (P5), conveyed by *since...*, it further lowers the force of the obligation. For the second instance of *should*, the tentativeness, i.e. the hypothetical nature of the suggestion (P6), is indicated by *Maybe*, a low-degree probability adjunct (P2).

This being an account of high- and median-obligation in the deontic domain of *should*, the formulaic use wants discussion, to conclude the section on the microsemantics of deontic *should*.

4.3.3.2.2 Formulaic *should*

As seen in Figure 24, the formulaic deontic use of *should* remains infrequent throughout the chronology of the study. It rises from 0,4 normalised in Period 0 instances to 1,1 normalised instances in Period 1, and rises again into Period 2 with 1,6 normalised instances. The formulaic frequency however drops to 0,6 normalised instances in Period 3, which is similar to its frequency in Period 0. As far as these small numbers permit such a conclusion, there is therefore some slight indication of colloquialisation across the 20th century, which involves the decline of formal writing formulae in this case. Table 15 shows the co-occurrence of formulaic deontic *should* with subjective, objective and indeterminate sources of obligation.

Table 15

Percentages of formulaic deontic uses with subjective, objective or indeterminate sources of obligation in *should*, with raw numbers in brackets

Period	Subjective	Objective	Indeterminate
0	0	100 (1)	0
1	0	100 (3)	0
2	33,3 (1)	66,7 (2)	0
3	0	100 (1)	0
Average	8,3	91,7	0

The results for Period 2 stand out in this table, being the only period when subjective sources are detected (where they indeed comprise a third of all sources), which supports the higher frequency of formulaic *should* in Period 2 mentioned above. For all other periods formulaic *should* is exclusively paired with objective sources. Because the raw instances of formulaic *should* as part of the deontic meaning, which contrasts with the formulaic use of *should* as part of the preterite meaning, are so few (respectively a total of 1, 3, 3, and 1 raw instances across Periods 0, 1, 2 and 3), no

real conclusions can be made as to the changing proportions of subjective/objective pairing with such uses. All such instances in Period 0 and 1 are found in news, whereas in Period 2, 2 of the 3 raw instances occur in non-fiction and 1 in fiction/narrative. In Period 3 the only raw instance occurs in non-fiction. Hence, the only deduction that can perhaps be made, considering the small raw numbers, and as far as these numbers permit, is that formulaic *should* is preferred in news in the 19th century and in non-fiction in the 20th century. The use of formulaic *should* with objective sources across all four periods is illustrated in chronological order below. All of these uses occur in the passive voice, as part of informative writing formulae.

(285) And here it **should** not be forgotten, that Dr. Morrison's own own studies were prosecuted without any of these helps which he has thus furnished to others... (1820s; News)

(286) And it **should** be kept in mind the evils which spring out of it are of vigorous and rapid growth. (1870s; News)

(287) It **should** be remarked that the fish was landed by the Indian; it was a 'springer', rather over five feet long, the shape and colour of a torpedo. (1930s; Fiction/narrative)

(288) It **should** be noted that spillage from Vaal Dam during the 1995/96 season was sufficient to refill the dam at least three times! (1990s; Non-fiction [W2A-036])

The pairing of formulaic *should* with subjective sources, which occurs once in Period 2, is exemplified in (289).

(289) Suffice it for me to say, therefore, that it must come – sooner or later; and I **should** mention that our own Minister of Finance, Mr Havenga, has been in the forefront of this movement. (1950s; Non-fiction)

In this use the writer is the first-person subject and the formula is presented in the active voice, which is a slightly different writing formula than that of the objective uses above, but, essentially, does the same job – fixing the attention of the reader on the new idea the writer introduces.

To conclude this section, the semantic ecology of *should* in SAfE is therefore increasingly dominated by the deontic meaning over the 20th century (macro-level monosemy), and this meaning is moreover characterised by a tendency toward a smaller proportional difference between high- and median-degree uses (a tendency toward micro-level polysemy), i.e. not a distinct move toward weaker obligative meanings over the 20th century (no clear monosemy), as is the case in other native Englishes. This in turn points toward a degree of synonymy with *must* on both a macrosemantic level, as regards tendencies toward deontic meanings (monosemy), and on a microsemantic level, involving tendencies toward a smaller proportional difference between high- and median-degrees of obligation (polysemy), as well as the tendency for the median-degree uses of both these modals to co-occur with objective and subjective sources with near-equal propensity, proving Bowerman's (2004b:477) claim regarding the interchangeability between *must* and *should* in SAfE. Furthermore, high-degree uses of SAfE *should* tend toward having more objective sources toward the end of the 20th century, whereas median-degree uses tend toward co-occurring with both sources on near-equal levels, as mentioned above, albeit with a slight tendency toward subjective sources. As far as *should* is concerned, there is some evidence of a link between its force and source of obligation by the 1990s, but, at least more so than for *must*, yet this link is still not very clear, especially in the case of median-degree uses. However, by the 1990s, SAfE *should* still has a slightly stronger tendency for conveying median-degree meanings than *must*, and the slight tendency toward subjective sources in this meaning, which suggests that the degree of synonymy is not complete, as any kind of synonymy rarely is, with *should* possibly retaining a tinge of politeness in these contexts. But, either way, the degree of synonymy is detectable to such an extent as to contrast strongly with the tendencies of other native Englishes, where *must* and *should* are options for very different contexts of use, and have both therefore found their own niche in the ecology of the obligation and necessity group. The picture in SAfE, on the other hand, is one of a shared, or at least overlapping, ecology. The next section explores the semantic ecology of the third major contender in the obligation and necessity cluster, quasi-modal *HAVE to*.

4.3.4 *HAVE to*

For comparison with the macro- and microsemantic trends of rival modals in the obligation and necessity cluster, *must* and *should*, the highest-frequency quasi-modal, *HAVE to*, is the last auxiliary to be discussed in these terms. Following the same order of discussion as for its rival modals, the section on *HAVE to* will firstly consider macrosemantics and then move to microsemantics.

4.3.4.1 Macrosemantics of *HAVE to*

Figure 25 illustrates the normalised diachronic macrosemantic trends for *HAVE to* in SAfE, as generated from the data in Table 29.2 in Appendix 5.

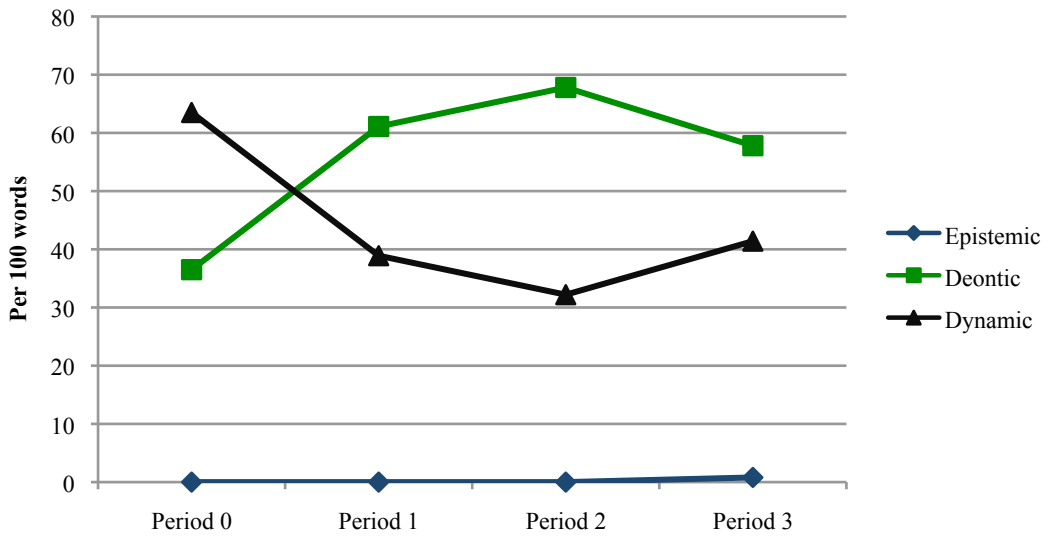

Figure 25

Macrosemantic frequency changes of *HAVE to* in SAfE over time

The trends of dynamic and deontic *HAVE to* are immediately striking in Figure 25: the frequencies of these meanings appear to move toward a point of intersection in some future time after Period 3, as they indeed have between Period 0 and 1. The epistemic meaning is absent for *HAVE to* from Period 0 to 2, and appears with a very low frequency in Period 3 (only a single instance). The overall trend of increase for this

quasi-modal in the diachronic corpus from Period 0 to 2 (see § 4.2.1.4.1) seems to be supported by the rise of the deontic meaning over the same time span (the declining trend of dynamic meanings apparently slows down the rate of overall increase, but does not prevent it), and the stabilisation it achieves between Period 2 and 3 (§ 4.2.1.4.1) appears to be supported by the change of direction in both the trends of dynamic and deontic meanings during this era, where they move in opposite directions. After the dynamic meaning is the most frequent among the other meanings in Period 0, the rise of deontic meanings from Period 0 to 1 by 67,4% (from 36,5 to 61,1 normalised instances) renders it the highest-frequency use from Period 1 onward. Its increasing trend however decelerates between Period 1 and 2, with an increase of only 11% (to 67,8 normalised instances in Period 2), and, as mentioned above, changes direction between Period 2 and 3 to decline by 14,7% (ending at 57,8 normalised instances in Period 3).

By Period 3, the semantic ecology of *HAVE to* in SAfE consists of 57,8% deontic, 41,4% dynamic and 0,8% epistemic meanings. The deontic meaning is far more frequent in contemporary AmE, BrE and AusE combined, occupying 76,1% of the semantic ecology of *HAVE to*, with dynamic meanings only occupying 22,4%, epistemic meanings 0,7% (similar to SAfE) and indeterminate meanings 0,8% (Collins, 2009a:60). Considering Leech *et al.*'s (2009:115) report that *HAVE to* tends to expresses general root modality in Contemporary English (cf. § 2.4.2.2), the SAfE data are in agreement, with both deontic and dynamic uses remaining dominant for this quasi-modal in the 20th century. The second part of Leech *et al.*'s (2009:115) finding in this regard involves that *HAVE to* tends toward meanings that are not face-threatening and have an objective source of modality (cf. § 2.4.2.2; cf. Mair & Leech, 2006; Collins, 2009a) – these tendencies will be investigated for SAfE *HAVE to* in § 4.3.4.2, after dynamic and epistemic *HAVE to* are discussed.

First, however, the changing tendencies in the occurrence of the macromeanings of *HAVE to* in the different registers over the 20th century needs to be considered, in order to account for this quasi-modal's general frequency patterns detected in § 4.2.1.4.1. Figure 26, generated from the data in Table 32.2 in Appendix 5, illustrates these tendencies. Since there was only one raw instance of epistemic *HAVE to* in the diachronic corpus, this meaning is not included in the graph.

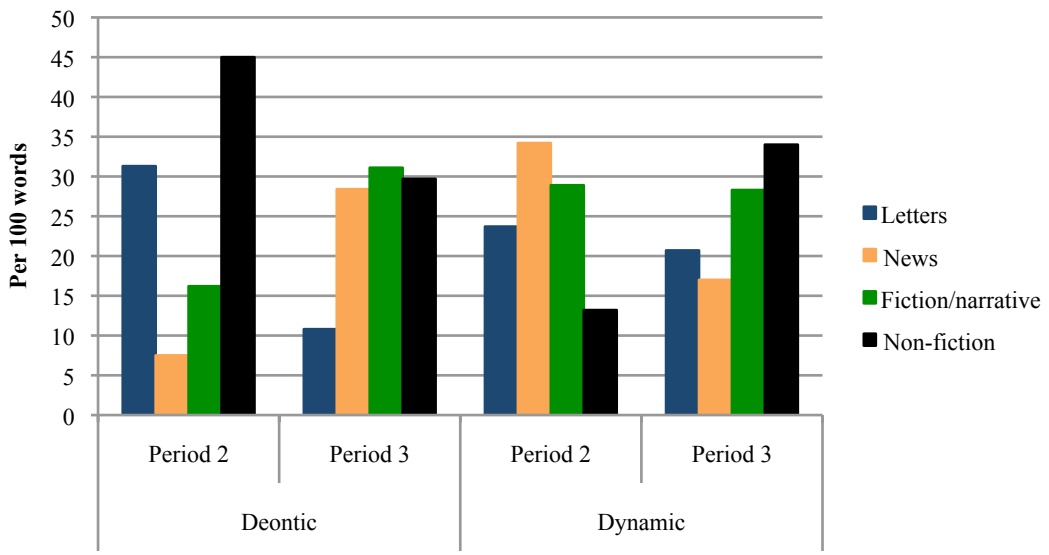

Figure 26

Macrosemantic frequency changes of SAfE *HAVE to* per register over the 20th century

Deontic *HAVE to* declines in two registers over the 20th century, namely letters (-65,5%) and non-fiction (-34%), supporting the downward trend of this meaning between Period 2 and 3 seen in Figure 25, but the simultaneous increase of deontic *HAVE to* in news (+278,7%) and fiction/narrative (+92%) no doubt slows down this trend. Dynamic *HAVE to* decreases in news (-50,3%), and more slightly in letters (-12,7%), whereas it remains rather stable in fiction/narrative (-2,1%), and increases (+157,6%) in non-fiction from Period 2 to 3. Non-fiction is therefore the main register supporting the general upward trend for dynamic *HAVE to* over the 20th century illustrated in Figure 25. Striking trade-off relationships between deontic and dynamic meanings are clearly present in firstly the non-fiction register, with the shift of popularity in favour of dynamic meanings by the 1990s, and secondly in the news register, this time with the shift in favour of deontic meanings by the end of the 20th century. Slightly less prominent trade-offs occur between these meanings into Period 3 in letters, in favour of the dynamic meaning, as well as in fiction/narrative, in favour of the deontic meaning. Thus, by the 1990s, the dynamic meaning is more often conveyed in non-fiction and letters, and the deontic meaning in news and fiction/narrative. The 20th-century increase of deontic meanings in news mainly supports the general increase of *HAVE to* in news from Period 2 to 3, which was noted

in § 4.2.1.4.1, whereas the decrease of deontic meanings in non-fiction mainly supports the overall decline of *HAVE to* in this register over the same period. Furthermore, the 20th-century trade-off in favour of dynamic meanings in letters seems to support the overall increase of *HAVE to* in letters, and the relative stability of *HAVE to* in fiction/narrative from Period 2 to 3 is mirrored in the stability of dynamic meanings over the same periods (see § 4.2.1.4.1). The next subsection will focus on the dynamic meaning of this quasi-modal.

4.3.4.1.1 Dynamic *HAVE to*

After dynamic *HAVE to* (in Figure 25) decreases by 38,7% from Period 0 to 1 (from 63,5 to 38,9 normalised instances), and decelerates as it declines by 17,2% from Period 1 to 2 (to 32,2 normalised instances), it changes direction into Period 3 with an increase by 28,6% (ending with 41,4 normalised instances). By comparison with deontic meanings, the ratio of deontic relative to dynamic *HAVE to* basically reverses from 1:1,7 in Period 0 to 1,6:1 in Period 1, and the differential widens to 2,1:1 in Period 2. In Period 3, however, the ratio narrows again to 1,4:1, indeed suggesting that the frequencies of these two meanings are moving closer together, as also noted above. Dynamic necessity meanings, where the "factors facilitating the activity reside in the situation" (Collins, 2009a:62), are often found in the fiction/narrative register of Period 0, which contributes greatly to the high frequency of dynamic *HAVE to* in this period. Here dynamic *HAVE to* is often used as part of a description of a journey, where external circumstances dictate a need for an activity. Since *HAVE to* is here used as part of narration, the past form *had to* is prevalent (cf. Banfield, 1982). These abundant uses are exemplified below for each of the five decades that comprise the data for Period 0.

(290) About Six oclock [sic] the wagones [sic] came up and we **had to** travel on until we came to Branfort [Brentford] five Miles from London: at this Town we haulted [sic] and Got brakefast [sic]... (1820s; Fiction/narrative)

(291) It was requisite that we should enter Kaffraria by way of the Chuinie, and pass through the country of Gaika; and, consequently, our journey being very

circuitous, we **had to** travel upwards of one hundred and sixty miles... (1830s; Fiction/narrative)

(292) The first part of the journey was somewhat discouraging, in consequence of the wide extent of country through which we **had to** travel, covered with thicket, which is inhabited by elephants, wolves, tigers, hyenas, buffalos, jackals, wild dogs, wild hogs, baboons, and other wild animals. (1840s; Fiction/narrative)

(293) It was so steep down into the Kloof they **had to** slide on their breeches which of course made some noise. (1850s; Fiction/narrative)

(294) I spanned in and proceeded to Sunday's River, where I stuck fast, the oxen not being able to pull the wagon out. I **had to** unload and carry the principal part of the freight out of the drift and, meanwhile, the oxen had a good feed. (1860s; Fiction/narrative)

In Period 1 the decline in dynamic *HAVE to* is greatly influenced by the reduced density of such uses in the fiction/narrative register, where dynamic necessity meanings were found in more equal proportions among the four registers. In the next three examples, dynamic *HAVE to* in Period 1 is exemplified in the news, non-fiction and letters register. In all three these examples the necessity is derived from some external circumstance, sometimes expressed in the matrix clause, here in the form of a rhetorical question (as in [295]), an opinion (as in [296]) and an assurance (as in [297]), where the necessity itself is usually conveyed in the subordinate clause.

(295) ...and, what exactly, are the nature of the chief difficulties which **have to** be overcome... (1870s; Non-fiction)

(296) For the rest, I am of opinion that direct taxation of some sort will **have to** be resorted to... (1880s; Social letters)

(297) ...readers may be assured that ...we **have to** face a necessity which will entail considerable outlay... (1890s; News)

The same tendency for dynamic *HAVE to* to occur in all registers with similar propensity broadly applies for Period 2, for which these uses are illustrated in the

letters and fiction/narrative register below. Here the necessity is derived from an external circumstance in (298) and from the subject referent, i.e. an internal necessity, in (299).

(298) I'll **have to** have the car decarbonised, & when I arrange it I should like to know if you are game to leave Ray field on Saturday for P.M.B.? (1930s; Social letters)

(299) Her legs are bad, she can't work, said Caroline, shrugging. I know, I said, but they **have to** eat. (1950s; Fiction/narrative)

In Period 3 the rise in dynamic *HAVE to* is largely supported by its continued use in all registers (as shown in [300] for e.g. social letters), but especially by its increased use in the non-fiction register (see Figure 26), where an external need is often expressed as part of the informative nature of these registers, as illustrated in (301). This use also occurs in news (see [302]).

(300) I hope you, Paul and Jacqui are still enjoying your jobs. Sorry everything **has to** be so hectic. That seems to be the way life is these days. (1990s; Social letters [W1B-004])

(301) The biggest problem is that more than 70 per cent of the people have **had to** live in only 13 per cent of the land. This causes overcrowding, overgrazing, and soil erosion... (1990s; Non-fiction [W2D-004])

(302) As with all pre-schools, Hillbilly has **had to** face dramatic Government cutbacks in funding... (1990s; News [W2E-002])

In (301) and (302) the necessity is grounded in the past, which is expressed by *had to* in both cases. One reason why *HAVE to* is the more frequent option for expressing dynamic necessity meanings amongst its rival for this meaning, *must* (see Figure 30 in § 4.3.5), can be located in the fact that *must* does not have the grammatical capacity to be used in this way, e.g. in these two examples *HAVE to* is not replaceable by *must*. *Must* is however much more frequent than *HAVE to* in conveying the epistemic

meaning (see Figure 28 in § 4.3.5), which is much like the situation in other native varieties (Collins, 2009a:63).

4.3.4.1.2 Epistemic *HAVE to*

As mentioned earlier, epistemic *HAVE to* only appears in Period 3 with a single instance in non-fiction (calculating to a normalised frequency of 0,8), which is given below.

(303) ... a small contribution to project costs with the rights to exploit 100% of the results **has to** be an attractive proposition... (1990s; Non-fiction [W2A-036])

Collins (2009a:63) notes that *HAVE to* conveying epistemic necessity is a 20th-century innovation inspired by AmE (cf. Coates, 1983:57), and is typically associated with younger speakers (cf. Hughes and Trudgill, 1979:23), but that it has now become established in BrE and AusE. As far as epistemic *HAVE to* being 'established' in SAfE, one raw instance in the entire diachronic corpus is hardly an indication thereof. The fact that this one instance occurs in the 1990s does however suggest that this use has found its way into the contemporary variety, even if its establishment as a semantic option might still be in an early stage. The next section will discuss deontic *HAVE to* and its microsemantics.

4.3.4.2 Microsemantics of deontic *HAVE to*

Figure 27 illustrates the proportional differences in the deontic frequency of *HAVE to* between high and median degrees of obligation, as well as formulaic uses, as drawn from Table 33.2 in Appendix 5. Once again, parameters that applied to the exemplified uses will be indicated.

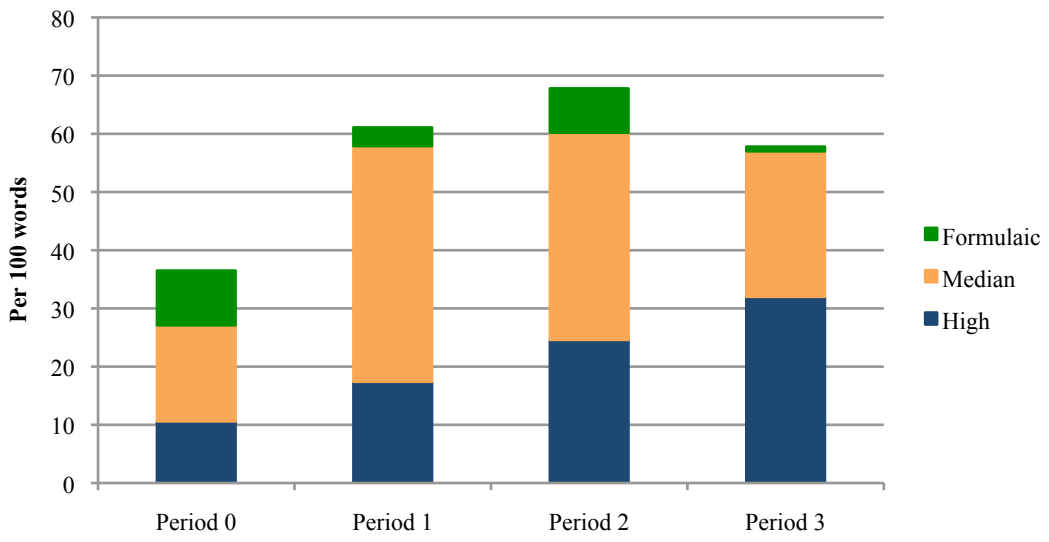

Figure 27

Microsemantic frequency changes in SAfE deontic *HAVE to* over time

The overall rising trend of deontic *HAVE to*, with its drop in Period 3, can once again be seen in Figure 27, as was also shown in Figure 25. Here the changing proportions between high-degree and median-degree obligation indeed support the rising trend of deontic *HAVE to* from Period 0 to 2, as well as the trend of decrease into Period 3. These changes will be considered in the following subsection. Formulaic uses are the most frequent in Period 0 (where they are almost as frequent as high-degree uses) and Period 2, and all but disappear in Period 3, also contributing somewhat to the deontic decline into the 1990s.

4.3.4.2.1 *HAVE to*: high and median degrees of obligation

With the increase of median-degree *HAVE to* from Period 0 to 1, the ratio between high- and median-degree uses widens from 1:1,6 in Period 0 to 1:2,3 in Period 1. In Period 2, however, the ratio again narrows to 1:1,4, with the rise of high-degree uses, returning to a more similar ratio to that of Period 0. By Period 3, the ratio has basically turned around (and narrowed somewhat) to 1,3:1, this time in favour of high-degree uses. A proportional difference of 0,3 is however not much, and considering that, despite its turn in Period 3, this is the period with the smallest proportional difference between these uses across the entire chronology. This

equalising trend is also visible in Figure 27. When considering the finding that both *must* and *should* have moved toward more equal proportional amounts of high and median uses by the end of the 20th century (a micro-level of polysemy), the fact that more or less the same tendency can be observed for *HAVE to* suggests that this quasi-modal might be moving toward a micro-level of polysemy itself – i.e. an equal propensity for expressing high- and median degree uses – which in turn suggests a level of synonymy between the three major options in the obligation and necessity cluster in SAfE regarding the deontic meaning, since the same kind of micro-level polysemy is present for *must* and *should* as well. The finding that SAfE *HAVE to* expresses both high- and median-degree uses in near-equal proportions by the 1990s is different from Leech *et al.*'s (2009:115) above-mentioned report that this quasi-modal expresses obligation of a non-face-threatening nature in Contemporary English, i.e. a lower degree of obligation, in the large native varieties. In itself the median-degree use of this quasi-modal has declined by 29,8% from Period 2 to 3 in SAfE, and the high-degree use has increased by 30,1% over the same era, with the high-degree use being slightly dominant by Period 3 (suggesting a trade-off between these two meanings over the 20th century), which does not correlate with the tendency in other native varieties.

Table 16 presents the results of the further microsemantic analysis of deontic *HAVE to*, regarding the co-occurrence of different sources of obligation with its high-degree uses over time.

Table 16

Percentages of high-degree deontic uses with subjective, objective or indeterminate sources in *HAVE to*

Period	Subjective	Objective	Indeterminate
0	0	100	0
1	9,1	86,4	4,5
2	6,9	93,1	0
3	26,8	73,2	0
Average	10,7	88,2	1,1

The above table shows that the high-degree use of *HAVE to* on average more frequently combines with the objective source of obligation, and this is influenced by the much higher objective percentages throughout. After all high-degree uses are paired with objective sources during the period of input, objective instances in Period 1 occur almost ten times more frequently than subjective sources (with a ratio of 1:9,5 subjective:objective uses), which is influenced by some indeterminate sources being present. Over Period 2 and 3, however, the differential between the subjective and objective sources widens, with a ratio of subjective relative to objective sources of 1:13,5 in Period 2. Overall, the objective source remains dominant over the subjective across the early 19th to mid-20th centuries. There is however a noteworthy shift which occurs in Period 3, where the ratio narrows greatly to 1:2,7, still in favour of objective uses, but with a much more balanced proportion than in preceding periods. There is therefore an indication of a trend for *HAVE to* toward a smaller degree of preference for one source over another regarding high-degree uses over the 20th century. Notwithstanding, the above-mentioned report of Leech *et al*. (2009:115) suggests that the uses of AmE and BrE *HAVE to* that are not face-threatening, in the terms of this thesis the median-degree uses, are more frequently objective, and as far as this comparison goes, SAfE is not so different.

When considering the late 20th-century and average findings for high-degree *HAVE to* together with those for high-degree *must* and *should* in this regard, some tendencies can be discerned (see Figure 26). In Period 3 high-degree *must* occurs more frequently with the subjective source than *HAVE to* (30,4% for *must* compared with 26,8% for *HAVE to*), and when considering that the percentage for high-degree subjective *must* is higher than for high-degree subjective *should* (14,3%) in Period 3, this renders *must* the option among its rivals with the greatest propensity to occur in subjective contexts with a high-degree of obligation. With a difference of only 3,6% between the frequencies of subjective *must* and subjective *HAVE to* in Period 3, the precedence of *must* over *HAVE to* in these contexts is not very substantial. On average, high-degree *must* also occurs more frequently with subjective sources than *HAVE to* (41,9% for *must* compared with 10,7% for *HAVE to*). The average high-degree subjective percentage of *should* is higher than that of *HAVE to* (28,2% for *should*), which shows that *must* has the greatest propensity amongst its rivals to convey this specific meaning across the entire time span, and not just in Period 3.

As for objective sources pairing with high-degree uses, *HAVE to* has a higher frequency than *must* in Period 3 (73,2% for *HAVE to* and 66,1% for *must*), which matches the finding that high-degree *must* prefers subjective sources. With a 7,1% difference the precedence of *HAVE to* over *must* in conveying these meanings is slightly more prominent than for subjective high-degree *must* over *HAVE to*, as noted above, but the difference is, once again, not massive. When compared with high-degree *should*, this modal however has a greater propensity to occur with objective sources than both its rivals (82,1% for *should*). On average, high-degree *HAVE to* also more often pairs with objective sources (88,2% for *HAVE to* compared to 55,3% for *must*), and when compared with high-degree *should*, *HAVE to* maintains precedence over its rivals (70,9% for *should*), making *HAVE to* the modal amongst its rivals with the greatest propensity on average to occur with objective sources in its high-degree use.

As far as the average percentages indicate traditional tendencies, the fact that the average tendencies and those for Period 3 correspond for high-degree *HAVE to* and high-degree *must* regarding both subjective and objective sources, suggests that the traditional uses are remaining popular into the late 20th-century, i.e. the tendency for high-degree *HAVE to* to be traditionally associated with objective sources continues into the 1990s, and the tendency for high-degree *must* to be traditionally associated with subjective sources also continues. Accordingly, a link between the strength and the source of obligation is suggested for *HAVE to* in Period 3 and on average, where objective rather than subjective sources tend to link with high-degree obligation. A comparative summary of co-occurrence between force and source in Period 3 for all three auxiliaries discussed in this main section will be given in § 4.3.5. However, the tendency for the frequencies of high-degree *must* and *HAVE to* to be much more similar during Period 3 than on average, suggests that the prototypically traditional uses are losing ground, and that high-degree *must* and high-degree *HAVE to* may be moving toward a degree of synonymy on this micro-level.

Table 17 shows the frequencies for the median-degree uses in SAfE according to source.

Table 17

Percentages of median-degree deontic uses with subjective, objective or indeterminate sources in *HAVE to*

Period	Subjective	Objective	Indeterminate
0	0	92,9	7,1
1	9,8	88,2	2
2	2,4	90,5	7,1
3	12,5	84,4	3,1
Average	6,2	89	4,8

A slightly similar proportional pattern as for high-degree uses also arises for median-degree *HAVE to* concerning the two sources. Just like the high-degree uses, the objective source enjoys exclusivity with the median-degree uses of Period 0, with no subjective uses. In Period 1, subjective relevant to objective uses have a ratio of 1:9, which is very similar to the ratio for high-degree uses during the same period. The ratio also widens for median-degree uses in Period 2, just like for high-degree uses, but with a much greater difference, to 1:37,7. The same trend that occurs for high-degree uses over Period 2 and 3, namely the narrowing differential between the two sources, also occurs for median-degree uses, as median-degree *HAVE to* has a ratio of 1:6,7 in Period. One difference, however, is the fact that the objective use remains clearly dominant in the 1990s for median-degree *HAVE to*, whereas the subjective source has become slightly more dominant for high-degree *HAVE to* by this time, and with a much narrower ratio than for median-degree uses. Leech *et al.*'s (2009:115) above-mentioned report for *HAVE to* in other native varieties to be dominantly objective therefore correlates more with the tendency for median-degree than with high-degree *HAVE to* in SAfE by the 1990s, although objective uses appear to be losing ground in SAfE by the end of the 20th century for both high- and median-degree uses.

When comparing these tendencies for median *HAVE to* with those of median *must* in Period 3, *must* has the greatest propensity to occur with subjective sources (35,9 for *must* and 12,5% for *HAVE to*), and on average the same pattern applies (44,2% for *must* and 6,2% for *HAVE to*). This is similar to the trend mentioned above

for high-degree subjective *must* to be more frequent in Period 3, but even more frequent on average, than high-degree subjective *HAVE to*. In Period 3 median-degree subjective *should* however occurs more than *HAVE to* and *must* (54,3% for *should*), which makes this the modal amongst its rivals with the greatest propensity to be used in median-degree contexts with subjective sources. On average, the similar frequencies of median-degree subjective *must* and *should* (42,9% for *should*) are both higher than for *HAVE to*, which suggests that these modals have the greater propensity to occur with subjective sources in their median-degree use than *HAVE to*.

As far as objective sources and median-degree obligation are concerned, the average percentages and the percentages for Period 3 between *HAVE to* and *must* form a very similar pattern: 84,4% for *HAVE to* and 54,7% for *must* in Period 3, and 89% for *HAVE to* and 53,5% for *must* on average. The preference of median *HAVE to* over median *must* to be used with objective sources is therefore a constant tendency over the entire time span of the diachronic corpus, as well as by the late 20th century. Seeing that median *should* with objective sources has an even lower frequency than *must* in Period 3 (45,7% for *should*), *HAVE to* indeed emerges as the option with the greatest propensity to be used with this microsemantic combination of force and source. These proportional tendencies are illustrated in Figure 32 in § 4.3.5. According to the average percentages, median-degree *should* has a slightly higher rate of co-occurrence with objective sources than *must* (57,1% for *should*), which matches the tendency for median *must* to co-occur infrequently with objective sources, when compared with its rivals. Subsequent paragraphs will exemplify the high- and median-degree uses of *HAVE to* as they co-occur with subjective and objective sources.

As noted for the data in Table 17, high-degree *HAVE to* only has objective sources during **Period 0**. This use is mostly found in the fiction/narrative register, but there are also occurrences in news. Examples from these two registers can be seen below. To a lesser or greater degree, most of the objective high-degree uses of *HAVE to* found in the diachronic SAfE corpus are close to the dynamic meaning on the root-modality continuum, since circumstances often influence the modality in both cases, but in the objective deontic cases, circumstances or other objective sources always impose an obligation, whereas in dynamic uses circumstances give rise to a need. In (304) and (305) the form *had to* grounds the obligation in the past time, as would be expected form the fiction/narrative register (cf. § 2.2.2.1; Banfield, 1982). Collins (2009a:61) notes that his data show a correlation between objectivity and third-person

subjects for *HAVE to*. This applies to examples (304) and (306), but this was found to not always be the case in SAfE, as illustrated in (305), where *HAVE to* has a first-person subject.

(304) I dined with the officers and after dinner the conversation turned upon the war and the many disagreeable duties the troops **had to** perform, especially escourting wagons. (1850s; Fiction/narrative)

(305) But as this was a civil service **I had to** get leave of absense as I was now considered by our Commandant of Grahams Town (Burnaby) to be a military man. (1850s; Fiction/narrative)

(306) Pretoria is at present the seat of government and the members of the Executive Council **have to** reside here. (1860s; News)

In (304) the obligation imposed by the external source that comes from certain military *duties*, which, despite being *disagreeable*, *the troops* are obligated *to perform*, which renders the obligation inescapable with no possibility of non-compliance (P5), as the 'duties' are imposed by military regulation (P6). Social norms therefore play a role in this apparently habitual obligation, making it face threatening for 'the troops' (all P5). The first-person subject of *HAVE to* in (305) does not point toward a subjective source, since the subject is the receiver of the obligation. The objective source is apparently some or other civil law or military regulation (P6), perhaps springing from the *Commandant*, a figure of authority (P5), which, like in the above example, renders it impossible not to comply with the obligation (P5). The obligation in (306) is clearly derived from circumstances, and, once again, from inescapable (P5) regulation (P6).

Just like for high-degree *HAVE to*, median-degree *HAVE to* does not occur with subjective sources in Period 0, but is dominated by objective sources, as exemplified below, also by means of two uses with third-person subjects (examples [307] and [308]) and one with a first-person subject (example [309]). Once again, some degree of overlap with the dynamic meaning is present in some cases.

(307) Providently, no lives were lost; but Great Britain **has to** lament the loss of Sir Stamford Raffles's invaluable manuscripts... (1840s; Fiction/narrative)

(308) It would cause a system of peculation on the part of the officers; men would be taken on by them who would be unable to equip themselves and the officers would **have to** do it and charge their own price for the articles. (1850s; Fiction/narrative)

(309) I am very glad to hear that there is a prospect of my going out in November, as this summer has been a very trying one in point of suspense and though, if I saw it was God's will, I would try to wait longer, still I must confess it would be a very heavy blow for me to **have to** do so. (1860s; Social letters)

In example (307) the source of obligation is clearly external, derived from circumstances – perhaps even 'providence', as could in a manner of speaking be paraphrased as 'what happened providence now requires lamenting'. The obligation is however median-degree, since it is almost a formula used by the writer – one with which to convey bad news, similar to a hedge (P5) – albeit with very little modal force, insofar that it is close to dynamic necessity on the deontic-dynamic cline. Since many dynamic uses would have some implications if the necessity were not fulfilled, due to it being grounded in circumstances or some internal need, this particular instance was not interpreted as overtly dynamic. What renders the instance both deontic and median degree is indeed the fact that there would be no gravity in the event of non-compliance, i.e. if *Great Britain* did not *lament the loss of [the] ...manuscripts*, there would be no consequences (P5) from the source that made the suggestion (P6) or from any other source for that matter. On the other hand, the hypothetical nature (P6) of the utterance in (308) flags the instance of *HAVE to* as conveying a median degree of obligation, as shown by the use of the modal *would* (see § 2.4.1.2; cf. Bybee, 1995:503; 511), as part of the bigger verb phrase containing *HAVE to*, as well as elsewhere in the extract. Here the hypothetical obligation *the officers* 'would' have to comply with is based on a consequence derived from circumstances, which lends it some overlap with the dynamic meaning. The different textual natures of the fiction/narrative and letters registers affect the use of the first-person subject in (309), where the writer of the letter refers to a hypothetical situation that may or may not actualise (P6), flagged by the use of the *to*-infinitive as part of the bigger verb phrase containing *HAVE to*, which lowers the force of the obligation,

albeit not considerably, due to some factors discussed below. The median-degree analysis applies despite the reference to *God*, which would be an indication that the obligation is imposed by an arguably subjective source with an indication of authority and inevitability, which would be indicative of a high-degree use. However, the writer mentions that he might have to act on what he *saw*, i.e. perceives, as, *God's will*, which is similar to the idea of 'providence' in (307), orientates the obligative source toward circumstances, i.e. an objective source. Of course it can also be argued that the writer himself is the subjective source of the obligation, since he himself 'sees' it, but this argument is not so strong as the one for external circumstances being the source, where it borders on a dynamic use. Furthermore, there is a possibility of non-compliance with the obligation implied in the reference to what he 'sees' as *God's will*: the writer states that he will only comply *if* he regards it as such, which leaves a possibility of con-compliance (P5) and counters the idea of inevitability noted above. On both the subjective/objective cline and on the high-degree/median-degree obligation cline, this is not a prototypical example of objective, median-degree *HAVE to*, but it has a tendency toward that analysis, and the tendency is not so weak as to render it indeterminate. This is however, once again, a testimony to the gradience of modal semantics, on both a macro- and micro-level. This is a special case, to illustrate the aforementioned point, as most of the objective, median-degree uses were not as ambiguous as this instance.

It was noted above that both high- and median-degree uses gain some subjective sources in Period 1, despite still being outnumbered by objective sources in both cases. The ratios of subjective relevant to objective sources are also very similar for median-degree and high-degree uses. The next example illustrates the subjective, high-degree use of *HAVE to* in **Period 1**, where these uses were only found in the fiction/narrative register in reported speech, suggesting that this use was very much only a spoken feature during this period.

(310) I was told that when once we had hold of the bird we were to hang on like death... ...you always **have to** be careful to keep it right at the bird's head, to put it against his breast would be worse than useless... The cock bird ...came flying down the camp towards us at full speed. "Stand firm there!" was the order... (1880s; Fiction/narrative)

In the above example, the high-degree of obligation conveyed by *HAVE to* is motivated by the instructive nature of the utterance (P6), since the speaker (which is the second-person subject *you* in this context [P1]) *was told* by a subjective source that is unnamed how to go about catching breeding birds, as is apparent from some context provided earlier in the text. The clause containing *HAVE to* serves as a warning (P6), as seen in the phrase *be careful*, which in turn is indicative that there will be consequences if the instruction is not complied with (P5), which is here implied to be both failure to catch a bird (the natural consequence) and moreover failure to complete the mission (implying social consequences). The implication of social norms (P5) is suggested by the quotation of a direct *order* in the imperative mood (P3), showing that the speaker is under obligation from someone with a higher social rank (P1). The emotional tone further contributes to the interpretation for a high-degree of obligation, as is shown in the expressions *hang on like death* and *worse than useless*. The adjunct *always*, which expresses a high degree of usuality (P2), further strengthens the force, by adding the sense that the obligation has a habitual nature (P5), that is to say, perhaps the speaker will be required to perform the same task in the future. Despite this layer of habituality, in the context of the utterance the obligation is grounded in the present (P4), as it applies to the speaker's current situation. This instance is a rare case, in that all six parameters apply to its interpretation as a high-degree obligative use, i.e. it is a prototypical use, and, reciprocally, this kind of use remains a rare find in Period 1.

The dominant objective high-degree use of *HAVE to* can be seen in the following examples.

(311) Well I am very much afraid this war will be more serious than any of us thought. It seems the only way now open for the military will be to burn all farms + railways that do not come XXX their lines + take all stock + so starve them out... ...it will be famine all over the country when this war is finished if it lasts much longer, as the Boers are planting nothing, then again, when we burn a farm house we **have to** take + feed the women + children so that there is so many less for the Boers to feed + you know a Boer can live on a few mealies[193] + a piece of meat for quite a long time. (1900s; Social letters)

[193] Corn or maize.

(312) The Free State shopkeeper on the north bank of the Orange River was able to undersell the colonial shopkeeper on the South bank of that River, as at Aliwal North, because the latter **had to** pay the high customs tariff of the Colony, and the former paid but [sic] fraction of that amount. (1880s; News)

(313) ...I have never failed to obtain as many men as there were vacancies for. The B troop P. A. M. P. which Gillibrand was sent to raise has been filled up long since and I have **had to** refuse the services of at least 20 men besides dismissing three. (1870s; Business letters)

In the historically interesting example (311) from a social letter, the topic of discussion is the Anglo-Boer War (as discussed in § 2.3.3.4.1), and specifically the so-called 'scorched earth policy' the British adopted in the later stages of the war (Giliomee & Mbenga, 2007:214-5). The clause containing *HAVE to* conveys a high degree of obligation, since it is grounded in an impending military situation, governed by orders, and, hence, regulation (P6). Here the obligation is part of an inevitable consequence of the policy, namely the need to house and feed the boer (Afrikaans) women and children after their houses have been destroyed (which led to the erection of concentration camps), rendering the source objective, and the obligation itself with little possibility of non-compliance (P5). Despite the fact that the utterance may appear to have a hypothetical layer, because the events discussed had not occurred yet, the inevitability and the gravity of the situation are revealed by the phrases *very much afraid*, *more serious* and *the only way*, which still indicates high-degree obligation. In (312) the objective source is clearly a law or regulation (P6), namely the *customs tariff of the Colony*, which strengthens the force of the obligation. For the *colonial shopkeeper* there is therefore no possibility of non-compliance (P5). However, in example (313) the source of the obligation is grounded in external circumstances, which shows some overlap with the dynamic meaning. In a way, it is a regulatory measure that prompts the obligation in this case (P6), and, as for the former two examples, there is little or no possibility of non-compliance (P5).

The infrequent use of median-degree subjective *HAVE to* is exemplified by the following instance.

(314) ...in the meantime we have not taught the black man at least two things. We **have to** teach him that there is justice – even-handed justice – for him... (1900s; Non-fiction)

The extract in (314) is from a philosophical opinion piece (P6) on Christianity and Africa. The source of the obligation is here interpreted to be the writer himself, since it is his opinion that is given by stating *We have to*. The obligation to *teach* the *black man* about *justice* is grounded in a hypothetical situation, which may never be actualised (P5), and with no social norms governing the obligation (P5).

Much more frequent objective median-degree *HAVE to* is illustrated below.

(315) ...it appears probable that pecuniary assistance will **have to** be rendered to some of the men to enable them to purchase horses... (1870s; Business letters)

(316) I suppose that Transvaal, Zululand, Tongaland and Natal ought to be a Republic some day or another. [The] Free State would **have to** come in and [the] Cape might enter or hold aloof. (1870s; Social letters)

The interpretation for a median degree of obligation in (315) is supported by the hypothetical nature of the situation that is described (P6). This reading is in turn reinforced by the use of the verb *appears* and by the modal *will*, which is part of the larger verb phrase containing *HAVE to*, as well as by the median-degree probability adjunct *probable* (P2). Much the same applies to (316), e.g. the hypothetical nature of the situation that is discussed (P6), as shown by the use of the modals *ought*, *would* and *might*, that add a layer of tentativeness to the utterance (see § 2.4.1.2; cf. Bybee, 1995:503; 511), as well as by the expression *some day* (the contemporary adverb 'someday') *or another* (P2). The fact that the obligation imposed here is based on an opinion (P6) is flagged by the use of the construction *I suppose*.

As noted previously, the tendency for both high-degree and median-degree uses to co-occur the most by far with objective sources continues into **Period 2**. The next example illustrates the subjective high-degree use of *HAVE to*.

(317) I shall always be pleased to play for you at any time if you get permission from Signor Grimaldi, he tells me I **have to** work very hard before he will consent, so I will work hard... (1920s; Business letters)

Example (317) is an extract from a letter addressed to an official of the 'Juvenile Musical Society', where obligation imposed by the subjective source, *Signor Grimaldi*, is being reported. The source is apparently in an authoritative position (P1), since his *permission* and *consent* are the conditions for the actions of the writer and the addressee of the letter, as here shown by *if* and *before*. The presence of a face-threatening situation, as well as the role of social norms (P5) in the situation further renders the obligation a high degree. There is an indication that the obligation imposed on the writer of the letter, namely to *work very hard*, will be complied with (P5), as seen in the statement *so I will work hard*.

The more frequent objective high-degree instances are exemplified in the following examples.

(318) Then one day, in a skirmish near the Vaal River, Hannes with a few dozen other burghers was cut off from the commando and **had to** surrender. (1940s; Fiction/narrative)

(319) The best years of my life had fled from me, never to return. But how magnificent! It meant that they would **have to** let me go, now. ... All that those years behind walls and bars had done to me was that I had now gone potty... (1940s; Non-fiction)

(320) And we found out that almost all the money Webber had was what he paid on the farm. He was always reading in those green books what he **had to** do. It's lucky that those books are written in English, and that the Boers can't read them. (1940s; Fiction/narrative)

In the extract form a piece of fiction based on the Anglo-Boer war (318), written about four decades after it actually occurred, the obligation is analysed to have a high degree of force, precisely because of the regulatory nature of a battle situation (P6).

This is indicated by the reference to a *skirmish,* the *burghers*[194] and *the commando*[195], as well as by the use of the main verb *surrender*, which shows that a non-compliance with the obligation will have consequences in the context of such a face-threatening situation (P5), perhaps even death. In example (319) conveys the obligation is imposed on what appears to be the prison administration by an external source, i.e. the fact that the convict or prisoner (here the writer of the autobiography [Herman Charles Bosman]) has served his time and is due for release. In a way, this obligation is regulatory, regarding e.g. the duration of the legal sentence, as the prison administration is therefore required by law to release the prisoner (P6), and there is not really any possibility of non-compliance (P5). It is discernible from the broader context of the text, from which the extract in (320) is taken, that *those green books* the narrator refers to are in fact farming guides. Since these 'green books' are the objective sources of the obligation, it renders the obligation itself an instructive nature (P6). It is suggested that not following the instructions may have serious repercussions for farming (P5), for example maladministration or a failed harvest, which positions the obligation on the stronger end of the continuum.

The only use of subjective median-degree *HAVE to* in Period 2 is illustrated in the below example.

(321) I wondered whether I was like my tie in this respect; whether I also, at the end of my imprisonment, retained something inside me that was bright-hued. But I feared not. My tie didn't **have to** say 'Yes, sir' all day long. (1940s; Non-fiction)

In the above example, which is from the same autobiographical text as (319), the narrator is interpreted as the source of the obligation imposed on an inanimate object (P1), his *tie.* Indeed here the 'tie' is actually released from the obligation to say *'Yes sir'* – in a sense something like an anti-obligation – (hence, the negation in *didn't* is not interpreted as a parameter for high-degree obligation), which places the instance

[194] Refers to the Boers (Afrikaners), literally translating to 'citizens'.

[195] Refers to a military unit of the Boer army.

at the very lowest end of the strength continuum. This utterance also contains a sense of irony, which further weakens the obligation.

The next few examples illustrate the co-occurrence of objective sources with median-degree *HAVE to*.

(322) I have good friends in the Dutch Reformed Church, and have for it a deep-seated respect; it is painful for me to **have to** reflect to-day on its action. (1910s; Non-fiction)

(323) Employers may believe they **have to** let evil consequences take care of themselves, but an enlightened State in 1931 has no such excuse... (1930s; Non-fiction)

(324) I **had to** buy a new hat – I was ashamed of he old – that was 7 ½ dollars. (1940s; Social letters)

In (322) it is not exactly clear what the source of the obligation is, but it appears to be external – perhaps the circumstances that prompted the 'reflection' or even something as abstract as the process of writing. This instance is therefore close to a formulaic use, which places it on the lowest end of the strength continuum. The philosophical content (P6) of the text can be seen in the main verb *reflect*, which further flags it for median-degree obligation. This analysis also holds, despite it being *painful for* the writer to fulfil the obligation, which indicates a somewhat emotional tone, and in some ways some surety as to compliance, because no social norms can be detected as forces behind the obligation, and, hence, there is nothing face threatening about it (P5). The obligation in (323) has a median-degree of strength due to its hypothetical content (P6), as expressed in the modal *may*, which adds a layer of possibility, and hence, a sense of tentativeness. It is also slightly unclear what the external source of obligation, to *let evil consequences take care of themselves*, is – but the likeliest source is a 'belief' (as indicated in the verb phrase *may believe*), i.e. an abstract idea. This lowers the force of the obligation, since a 'belief' is essentially a kind of opinion (P6). Also, the writer is of opinion that such a 'belief' is only used as an *excuse*, i.e. he criticises the opinion of the *Employers*, which implies that there was room for non-compliance with the obligation in the first place (P5). The obligation imposed on the speaker in example (324) springs from an objective source, which can be interpreted

as natural circumstances ('the old hat is weathered'), or as social circumstances ('the old hat is unfashionable'). In any case, the source is rather abstract, most likely connected with the speaker's 'shame'. This use is somewhat ambiguous with the dynamic use, since the obligation can also be interpreted to be imposed by the speaker himself, springing from an external need, but the utterance implies that it is not the need itself that is being expressed, but, as noted above, an obligation arising from that need, which sways the instance in the deontic direction, albeit with very little modal strength.

As noted above, both high- and median-degree uses with subjective sources increase in **Period 3**, and more so with high-degree uses, but the objective source continues to remain dominant. The examples below show how the subjective high-degree instances are used. Such instances were mostly found in fiction/narrative, which supports the general rise of deontic meanings of *HAVE to* in this register over the 20th century, as seen in Figure 26.

(325) He was always good for the odd rand or two when they were out of cash, and he was always polite to them. Though, God knows, Vernon didn't always return the compliment. Chrissie used to say to him, "Ag no man, Vernon, yous [sic] got such a bad ettitude [sic]. You don't **have to** be so rude to people..." (1990s; Fiction/narrative [W2F-017])

(326) For the sake of the missing child Simon was aware that he would have to conduct this interview with extreme care. "No guarantee." Fulke admitted after a silence. "They don't act that way. Right now they have all the cards." "Then we **have to** do it my way." He turned his considerable charm on to the grieving mother. "You have my word that I will find Piers and bring him back to you. Triad or no Triad. Will you trust me?" "Yes!" she said, almost hysterical. (1990s; Fiction/narrative [W2F-009])

(327) "Yes, because you are always bothering me with your foolishness. But I'll teach you a lesson." He grabbed her dress and towel and stood at some distance, ignoring her pleas for her clothes. "You'll **have to** come and get them." "But I'm naked..." (1990s; Fiction/narrative [W2F-018])

(328) We had practised apartheid for all our lives, Stompie and I not without soul-searching. In the early 1960s we were living in Durban where I was on the

> editorial staff of Die Nataller, a Nat Party weekly under the editorship of T. J. A. (Theo) Gerdener. He would later become the first Afrikaans-speaking Administrator of Natal before joining John Vorster's Cabinet. His successor as editor was Barend Venter, a very sharp and discerning journalist whose political ideas were so far ahead of his time that he **had to** terminate his duties by "request" of his bosses. (1990s; Non-fiction [W2B-013])

In (325) the subjective source of the obligation is *Chrissie*, which shows her disapproval at the impolite behaviour (the 'bad attitude') of the addressee, *Vernon*, shown by the exclamation *Ag no man*[196], which definitely indicates that the situation has a face-threatening value (P5), especially since there is no irony present in this context. Here the sentence *You don't have to be so rude to people* is basically a substitute for 'don't be so rude to people', so it could be argued that by adding *HAVE to* and making what might have been an imperative clause into a declarative clause, the speaker is relieving the addressee from an 'obligation' to 'be rude to people', for which the source is unclear, i.e. it might point toward median obligation. However, the analysis for high-degree obligation was chosen on the basis that this particular clause is in fact imposing a new obligation, perhaps at the same time as relieving the addressee from another. The negation of the verb phrase further strengthens the analysis for high-degree obligation (P4). The obligation conveyed in example (326) functions on the stronger end of the strength continuum due to, firstly, the nature of the situation that is described. This is a serious, tense, delicate and apparently emotional situation (P5) related to a police investigation into a child's abduction. This is shown in the mention to *extreme care*, which implies a warning (P6), and by the fact that there are 'no guarantees'. The emotional nature of the situation is indicated by adjuncts like *grieving* and *hysterical* (P2). The obligation is imposed by the officer in charge of the case, which indicates authority (P1). It is less certain on whom exactly he is imposing the obligation: it might be on the abductors, on his task team or on the *grieving mother*, or perhaps on all three. Either way, the obligation does not lose its face-threatening quality (P5). In the case where the 'mother' could be the

[196] A South Africanism, which is basically a loan expression in a directly translated form, from the Afrikaans 'ag nee man', which indicates disapproval or disgust, where 'ag' in this context is basically translatable as English 'oh'. (See also example [197].)

receiver of the obligation, the fact that the speaker 'turns on' *his considerable charm*, suggests that he is in control of the situation, which strengthens the idea of authority and of the social implications of the utterance (P5). It can also be deduced that, because the 'mother' obliges 'to do it his way', i.e. to *trust* him, compliance on her account is certain (P5). The obligation expressed in example (327) also has a face-threatening quality (P5), as is shown in the declaration of the speaker to *teach* [the addressee] *a lesson*, which clearly indicates that the intention of the speaker is to shame the addressee. This means that social norms play a role in the situation (P5). When her face is threatened, the addressee protests by pointing out the nature of the shame in the situation – she is *naked* – and by making *pleas for her clothes* to the speaker. This face-threatening quality is further reinforced by the use of the verbs *bothering*, *grabbed* and *ignoring*, and by the noun *foolishness*. The context of the obligation in (328) is one of political tension under the apartheid regime, which indicates the presence of social norms (P5). Here the writer insinuates that the receiver of the obligation, *Barend Venter*, was forced to resign from his job, because he was apparently a threat to the apartheid government, which is indicative of a very face-threatening situation (P5). The forced quality of 'Barend Venter's' compliance (P5) with the obligation, as well as the apparent underhandedness of *his bosses* in the whole process (the subjective sources of the obligation), implied by the use of the quotation marks used by the writer in *"request"* (to show that it was in fact not a 'request', but an order), suggests that there is a legal or regulatory character to the obligation, both under the authority of the bosses (P1), and under the laws against resistance to the government (P6). All of these parameters indicate high-degree obligation.

High-degree *HAVE to* which co-occurs with objective sources in Period 3 can be seen in the examples below.

(329) "I remember how they waited to build the road, because a certain number of houses first had to be built. I even **had to** pay for the pole in front of my house, from which my electricity was supplied." (1990s; News [W2C-003])

(330) Today King Goodwill's palaces are open to an endless stream of visitors. From 1979 until the break with uncle, all his meetings **had to** be sanctioned by the KwaZulu Cabinet. The king recalls: "People were chosen to be with me

when I was seeing someone – that made me very childish. I won't name those people because I am speaking for the sake of peace." Relatives recall how people sent by Buthelezi would sit in on the king's meetings and report back by walkie-talkie on what was said. (1990s; Non-fiction [W2B-001])

(331) Fortunately the law in Utah stipulates that John Taylor – and anybody else facing death by firing squad – **has to** be shot by a five-man team of state law enforcement officers. That should spare Taylor an unnecessarily messy death. (1990s; News [W2E-003])

The source of the obligation in (329) is suggested to be the local municipal authority (P1), seeing that the receiver of the obligation, indicated earlier in the text to be an original resident of an area celebrating its golden anniversary, was required to *pay* for what should have been a *pole* funded by the municipality. The obligation therefore is regulatory in nature (P6), with no real possibility of non-compliance (P5). This analysis is reinforced by the use of the adverb *even*, which conveys an exceeding level of counterexpectancy (P2). In example (330) the obligation has a legal character (P6), i.e. objective source is *the Kwa-Zulu Cabinet*. This is further shown in the verbs *sanctioned* and *report*. The obligation is indeed a threat to the face of *King Goodwill* (P5), the receiver of the obligation, as indicated in his statement that it *made [him] very childish*. The legal nature of the obligation further suggests gravity in the event of non-compliance (P5). The legal nature of the obligation in (331) is very clear from the mentions of *the law in Utah*, *state law enforcement officers* (P6), as well as from verbs such as *stipulates*, which render this an example of high-degree obligation. Phrases such as *death by firing squad*, as well as verbs like *facing* and *shot* further indicate that the obligation is inescapable (P5). The writer also adopts a somewhat emotional or evaluative tone (P5), as indicated by verbs such as *spare*, and adjuncts like *Fortunately*, *unnecessarily* and *messy* (P2). This also contributes to strengthen the force of the obligation, despite the fact that it shows the relief or approval of the writer, precisely because it indicates relief that the law has to be complied with.

Example (38), which is provided in Chapter 3 (§ 3.4.1.1), already illustrates the median-degree use as it co-occurs with the subjective source. The next example also demonstrates this usage.

(332) Chief Kowolski walked the detective to his borrowed car. "Will you keep us informed? This business is right over our heads. I know that. If this department can help in any way you only **have to** ask." He held out his hand. "I have heard Fulke's story and gotten a few ideas." Simon shook hands warmly. "My guess is that this could be far-reaching. You did all you could." (1990s; Fiction/narrative [W2F-009])

The obligation conveyed in the above extract is essentially part of a polite gesture (P5) by the source, which is *Chief Kowolski*. Despite the inference that the 'Chief' of a certain *department* is of higher status than *the detective*, the addressee in this case, the context of the piece of fiction shows that this is a 'Chief' from another department, as indicated in *this department* (i.e. he has no direct authority over the addressee) (P1), which cannot solve the case in question, as shown when the speaker admits *This business is right over our heads* (i.e. he is withdrawing from the case). The polite gesture of the speaker is an offer of *help*, which is not face threatening at all (P5). This is done by the use of some hedges, such as *in any way* and *only* (P5). There are also various other indications that the interlocutors are on friendly terms (P5), e.g. the speaker walks the addressee to his car, they *[shake] hands warmly* (P2), and the addressee shows returned politeness when he says *You did all you could*. There are therefore many parameters to mark median obligation in this case.

The more abundant objective median-degree use is illustrated in the following examples.

(333) He spent a lot of time putting up my roman blinds at last – as I well know the ones in the bathroom look very bland, and I shall **have to** paint (ugh!!!) the loo at last. (1990s; Social letters [W1B-005])

(334) The marsh rose (Orothamnus zeyheri), strawberry everlasting (Syncarpha eximia) and several Mimetes species, for example, may **have to** accommodate the dietary needs of the vlei rat. (1990s; Non-fiction [W2B-024])

(335) But South Africa's constitution-makers, in deciding whether capital punishment is compatible with human rights, will **have to** weigh up if – and how – it should be carried out. (1990s; News [W2E-003])

In example (333) the source of the obligation seems to be related to circumstances, which does not render it clearly subjective. There is also some ambiguity with the dynamic meaning, much like in example (324), since it might appear as if the speaker herself is the internal source. However, *HAVE to* does not necessarily convey a pure need in this case, but rather a statement that is motivated by an evaluation of the circumstances or of the need, rather than by the circumstances itself. There is no clear indication that this evaluation is actually the opinion of the speaker herself, as would have been given in a hedge like 'I think have to paint...', which would have made the example subjective, or whether there is some other abstract force at work such as fashion or aesthetics, as is suggested in the reference to the *bathroom [looking] very bland*. The other possibility is that the pronoun *he* refers to the subjective source, but there is, again, no clear indication that *he* issued a direct obligation. The presence of an external source is indeed suggested by the presence of the modal *shall*, as it collocates with *HAVE to*, much like the use of *will* and *may* add a layer of objectivity to examples (334) and (335) respectively. This in turn also strengthens the analysis for deontic modality, as *shall* is deontic here, in conveying a degree of inevitability. The inevitability itself is however not grounds enough to declare this a case of high-degree obligation, since there is no clear authoritative figure that would enforce the obligation, and no real consequences in the case of non-fulfilment (P5), except that *the loo* would continue to *look very bland*, which is hardly face-threatening in nature (P5). Also, the indication of emotions, e.g. in *(ugh!!!)* and *at last*, in this case does not suggest a high degree of obligation, but rather indicates the unwillingness of the speaker to comply (P5), again flagging the instance for median-degree obligation. Therefore, this is not a prototypical example of objective deontic *HAVE to* with a median-degree of force, but it does tend toward this end of the semantic cline, mostly based on a pragmatic analysis and by the use of *shall* in the verb phrase containing *HAVE to.* Ambivalence with the dynamic meaning is also found in (334), where an internal need is suggested, but, once again, the utterance is based on an evaluation of the necessity of the situation, which is hypothetical (P6) (as conveyed by the use of modal *may*, here adding a layer of possibility), rendering it deontic with a median-degree of strength. The source of this evaluation is however grounded in circumstances, i.e. something external, which renders it objective. The analysis for median obligation is further reinforced by the inanimate subjects of *HAVE to* (plant species) (P1). In (344) the obligation is interpreted as having a median degree of

strength, since the content is rather philosophical, and also based on opinion (P6), but the source of that opinion is not revealed, rendering it objective. This analysis is reinforced by the use of deontic *will* which implies futurity (cf. Collins, 2009a:134), i.e. a hypothetical layer (P6).

4.3.4.2.2 Formulaic *HAVE to*

As seen in Figure 27, the formulaic deontic use of *HAVE to* remains infrequent throughout the chronology of the study, as is also noted for *should* in § 4.3.3.2.2. This use declines by 66% from Period 0 to 1 (from 9,4 to 3,2 normalised instances), but increases again by 137,5% between Period 1 and 2 (to 7,6 normalised instances in Period 2). As briefly noted above, this meaning of *HAVE to* decreases considerably (-89,5%) in Period 3, almost disappearing entirely with only one raw instance detected in the analysis. Table 18 shows the co-occurrence of formulaic deontic *HAVE to* with subjective, objective and indeterminate sources of obligation.

Table 18

Percentages of formulaic deontic uses with subjective, objective or indeterminate sources in *HAVE to*

Period	Subjective	Objective	Indeterminate
0	87,5	0	12,5
1	100	0	0
2	100	0	0
3	100	0	0
Average	96,9	0	3,1

Table 18 shows that formulaic deontic *HAVE to* clearly prefers the subjective source in general, with one raw occurrence of an indeterminate source in Period 0, calculating to 12,5%, being the only 'other' source, albeit unclear, found over the chronological course of the data (see Tables 34.1 and 34.2 in Appendix 5). All of the uses detected in the diachronic corpus have first-person subjects. Such formulaic uses are applied in writing with a similar strategy to that of some of the formulaic uses of

must and *should*, namely to emphasise or introduce an idea to a reader. A comparison with formulaic *must* and *should* is provided in § 4.3.5.

The formulaic uses of *HAVE to* are illustrated below for each of the four periods.

(336) Sir, ...I **have to** request the favour of you to lay them before the Right Honble. Lord Visct. GOODRICH. (1820s; Business letters)

(337) My dear Rouliot
I **have to** acknowledge the receipt of your very interesting letter of Aug. 2nd. (1900s; Social letters)

(338) Dear Sir,
It is with great regret that I **have to** inform you that I resign from your Committee, as I find that I have not sufficient time to fulfill my duties conscientiously. (1920s; Business letters)

(339) However, we **have to** bear in mind the constantly shifting nature of the disciplines within which we operate... (1990s; Non-fiction [W2A-022])

In Period 0, 7 of the 8 raw instances of formulaic *HAVE to* occur in business letters (as illustrated in [336]), and 1 in news, whereas in Period 1, 3 of the 4 raw instances occur in social letters (as illustrated in [337]), and the remaining 1 occurs in non-fiction. Business letters one again contain the bulk (5 uses) of the 9 formulaic uses in Period 2 (as seen in [338]), with the remaining uses found in social letters (2 uses) and in non-fiction (2 uses). In all of these instances of formulaic *HAVE to* in letters, a layer of politeness is noticeable, especially in the older letters, and especially in business letters, where hedging often occurs. The one use that remains in Period 3 occurs in non-fiction, as seen in (339), suggesting that the formulaic hedging that occurred in preceding periods has stopped being used, especially in letters. This apparent movement away from over-formalised styles, as far as these small amounts of formulaic instances can indicate such a movement, is suggestive of the colloquialisation process.

To conclude on the semantic ecology of *HAVE to* in the 20th century, it was found that both deontic and dynamic uses remain dominant, with a narrowing of the

proportional difference between these uses by the 1990s. For deontic *HAVE to*, a high degree of obligation is expressed only slightly more often than a median degree by Period 3, with a trend toward near-equal proportions of these micromeanings (i.e. a micro-level of polysemy) toward the end of the 20th century. Since the same trend was found for *must* and *should*, this suggests a degree of micro-level synonymy between the three major contenders in the obligation and necessity cluster. This differs from the tendency for *HAVE to* in other large native varieties to express a lower degree of obligation in the 20th century. For both the high-degree and median-degree uses, there is shift toward a much more balanced proportion between subjective and objective sources toward the end of the 20th century than in preceding periods. The objective use however remains clearly dominant for median-degree *HAVE to*, whereas the subjective source has become slightly more frequent for high-degree *HAVE to* by the 1990s, even though the objective use remains the more popular option. It was also reported that the tendency for *HAVE to* in other native varieties to be dominantly objective correlates more with the tendency for median-degree than with high-degree *HAVE to* in SAfE by the 1990s, although objective uses are losing ground in SAfE by the end of the 20th century for both high- and median-degree uses. Lastly, the formulaic use of *HAVE to* has all but disappeared by Period 3, which was taken as an indication of the colloquialisation process being at work in this regard. As a summary of the changes in the semantic ecology of *must*, *should* and *HAVE to*, the next section will compare these auxiliaries with each other in terms of the semantic findings.

4.3.5 Comparative summary: the semantic ecology of obligation and necessity

This section will summarise the findings discussed in § 4.3 for the three major contenders of the obligation and necessity cluster in such a way as to aid direct comparison based on both macrosemantic and microsemantic ecology, as well as to locate the SAfE trends in terms of those for other native Englishes. Figure 28 shows the epistemic trends of SAfE *must*, should *and HAVE to* relative to each other.

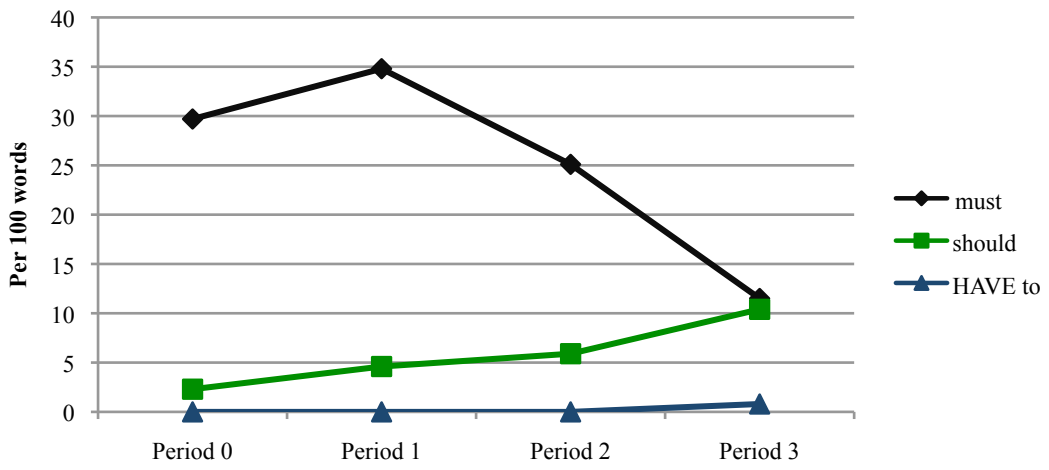

Figure 28

Comparative frequency changes of epistemic modality in *must*, *should* and *HAVE to* in SAfE over time

Figure 28 clearly shows that the trends for epistemic *must* and *should* have moved toward near-converging frequencies in SAfE by the end of the 20th century (compare 11,5 normalised instances for *must* in Period 3 with 10,4 normalised instances for *should* in the same period – i.e. a difference of only a 1,1 normalised instances). This illustrates the apparent trade-off between epistemic *must* (-54,2%) and *should* (+54,2%) during the 20th century well, as noted in § 4.3.2.1.2 and § 4.3.3.1.1. Both *must* and *should* are therefore more or less equally likely to be used in epistemic contexts in contemporary SAfE. This does not entirely agree with the findings for *must* and *should* in other native Englishes, where *must* is much more likely to be used in epistemic contexts than *should* (cf. Collins, 2005; 2009a; Leech *et al*., 2009:115-6), but as far as the gap between the frequency of epistemic *must* and *should* goes, the pattern in SAfE for Period 3 is closest to that of contemporary AmE and AusE, where "*should* is only marginally less popular than *must*, while in BrE and NZE the gap is greater, with epistemic *must* about twice as common as *should* in BrE and about two and a half times in NZE" (Collins, 2005:268). Figure 28 also shows that SAfE *HAVE to* is not a major contender for the epistemic meanings in this semantic cluster (see also § 4.3.4.1.2). This matches the contemporary findings of e.g. Collins (2005; 2009a), Leech *et al*. (2009) and Close and Aarts (2010) for other native Englishes, so the main discrepancy between the trends for SAfE and those of other varieties lies in the relationship between *must* and *should*.

Notwithstanding, the trend of overall decrease for epistemic *must* seen in Figure 28, especially from Period 2 to 3, is not that different from the trends for *must* in other native Englishes, in that epistemic *must* also shows a decline over the 20th century in these varieties (cf. Leech *et al*., 2009:88; Close and Aarts, 2010:179). In view of the overall semantic ecology of the obligation and necessity cluster, as well as the trends for other meanings of *must*, this decrease in frequency was however not taken as evidence that *must* has declined in its use of the epistemic meaning, but rather that *must* is becoming the most popular option for expressing epistemic modality among its rivals, especially *should*, for which the epistemic meaning has also decreased (Leech *et al*., 2009:87). It is clear from Figure 28 that this is much different from the trend of increase for epistemic *should* in SAfE over the 20th century. The reason that the fall of epistemic *must* in other native Englishes has not been linked to the declining use of this meaning, but rather the opposite, is because Leech *et al*. (2009:88-9) describe the fall of epistemic *must* as "shallow" in comparison with its fall in deontic meanings – indeed they argue that the decrease of epistemic *must* "could be due to contamination by the dramatic fall of deontic *must*". By contrast, the deontic meaning of *should* becomes more dominant in other native varieties (Leech *et al*., 2009:86), so the fall of epistemic *should* in other varieties can be taken as evidence of its growing monosemy, with an inclination toward deontic uses. These tendencies within the obligation and necessity cluster do not match the deontic trends for these modals in SAfE, as seen in Figure 29.

Figure 29 illustrates the overall increase in deontic meanings for both *must* and *should*. SAfE *must* has not declined in its expression of deontic meanings over the 20th century, like it has in other native Englishes (cf. Leech *et al*., 2009:88), but has been increasing since Period 1 onwards. This can of course be linked to the finding that SAfE *must* has not strongly declined in its overall frequency either, as it has in other native varieties, but remained rather stable with only a slight decrease over the 20th century, as reported in § 4.2.1.4.1. In other words there is no dramatic decline of this meaning as the by-product of its dramatic decline overall, as in other varieties (cf. Smith, 2003:242; Leech *et al.*, 2009:114-6), but it would rather appear that the relative stability of SAfE *must* over the course of the 20th century is a product of its increasing popularity in deontic contexts over its two rivals, together with its decreasing epistemic frequency (cf. § 4.3.2.1). This suggests a tendency toward a degree of macro-level monosemy for SAfE *must*, in favour of deontic meanings, as

noted in § 4.3.2.1, whereas no such trend toward monosemy was reported for *must* in other native varieties, apart from monosemy on a micro-level, where strong/high-degree deontic uses are becoming more frequent (cf. Smith, 2003:242; Leech *et al.*, 2009:114-6; Millar, 2009:203).

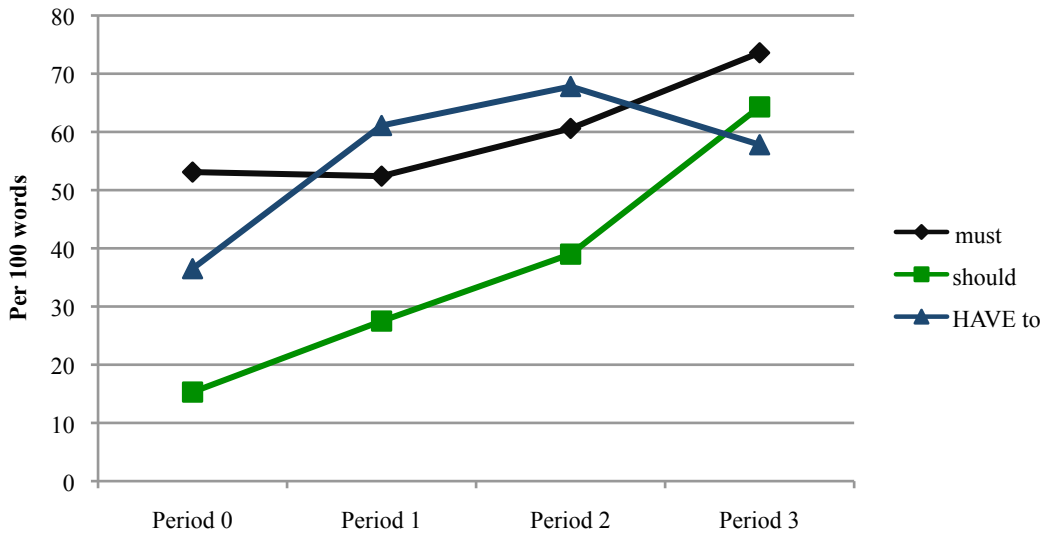

Figure 29

Comparative frequency changes of deontic modality in *must*, *should* and *HAVE to* in SAfE over time

The increase of deontic *should* in SAfE would seem to agree with the degree of dominance that deontic *should* gains in other native Englishes in comparison with its other meanings (cf. Leech *et al.*, 2009:86-7), but in view of the increase of epistemic *should* in SAfE over the 20th century, the degree of deontic dominance is somewhat less in SAfE. What makes the 20th-century deontic trends for *should* and *must* so interesting, is that, although deontic *should* increases more than deontic *must* from Period 2 to 3, *must* remains the dominant option for the expression of deontic meanings in this semantic cluster, whereas *should* is the more dominant deontic (root) option in other native varieties, including other SHEs (cf. Collins, 2005:266). What is striking about Collins' (2005) analysis is that he analyses root meanings, which include dynamic uses. Indeed *should* does not even convey dynamic meanings, while *must* does, so the fact that root *should* surpasses root *must* in other varieties suggests that in truth deontic *should* is really dominant over deontic and dynamic *must*. This contrasts with the SAfE findings, where deontic *must* alone is already more frequent

than deontic *should*. The above-mentioned trend for native Englishes is however also confirmed when only deontic meanings are considered, where, on average, deontic *should* was found to be more than double the frequency of deontic *must* in BrE, AmE and AusE combined (Collins, 2009:35;45). In SAfE, however, the gap between deontic *must* and *should* is not nearly as big by Period 3, with a difference of only 9,3 normalised instances (compare 73,6 normalised instances for *must* and 64,3 for *should*). One effect of this trend is that it means that neither *must* nor *should* have lost their deontic meaning in SAfE at all, whereas it was generally found that the central modals have lost this meaning "to a great extent" since OE, except for *must* (Abraham, 2001:30; cf. § 2.4.2.2).

The decrease in the deontic use of *HAVE to* in SAfE from Period 2 to 3 is not consistent with the tendency for this quasi-modal in other varieties, where it was found to convey even more deontic meanings than *should*, and of course *must*, in other native Englishes (Collins, 2009a). In SAfE *HAVE to* less frequently conveys deontic meanings than both *should* and *must*. Nevertheless, the deontic frequencies of the three auxiliaries in question are closer together by Period 3 than in the three preceding periods. This agrees with the trend for the overall frequencies *must*, *should* and *HAVE to* to move toward a lesser degree of differentiation in Period 3 than in preceding periods, as noted in § 4.2.1.4.1 (see Figure 11), so it would appear as if the deontic trends motivate the overall trends to the largest extent. This suggests that the three auxiliaries move toward more similar degrees of popularity in deontic contexts by the 1990s, than they enjoyed over the 19th and early 20th centuries.

Figure 29 shows that *must* and *HAVE to* have been the closest contenders for the expression of deontic meanings over time, with *should* only really becoming a serious contender in Period 3. After *must* was the more popular option for deontic uses in Period 0, and *HAVE to* moved to become the more popular in Periods 1 and 2, *must* moves into the lead position once again by the late 20th century, i.e. it regains its place as the main option to convey deontic meanings. This is a striking finding indeed, when compared to the decreased "social-semantic popularity of *must*" found in other native varieties over the 20th century (Leech *et al*., 2009:115), since it appears as if SAfE *must* has not decreased in its popularity in comparison with the other contenders, but has rather increased in this respect.

It was noted in various studies that *HAVE to* has become more likely to express general root meanings (deontic and dynamic) over the course of the 20th century, both

in itself and among its contenders in the obligation and necessity cluster (e.g. Collins, 2005:266; Leech *et al.*, 2009:87-9; Close & Aarts, 2010:179). In order to fully assess this for SAfE, the frequency of dynamic instances for *must* and *HAVE to* are compared below.

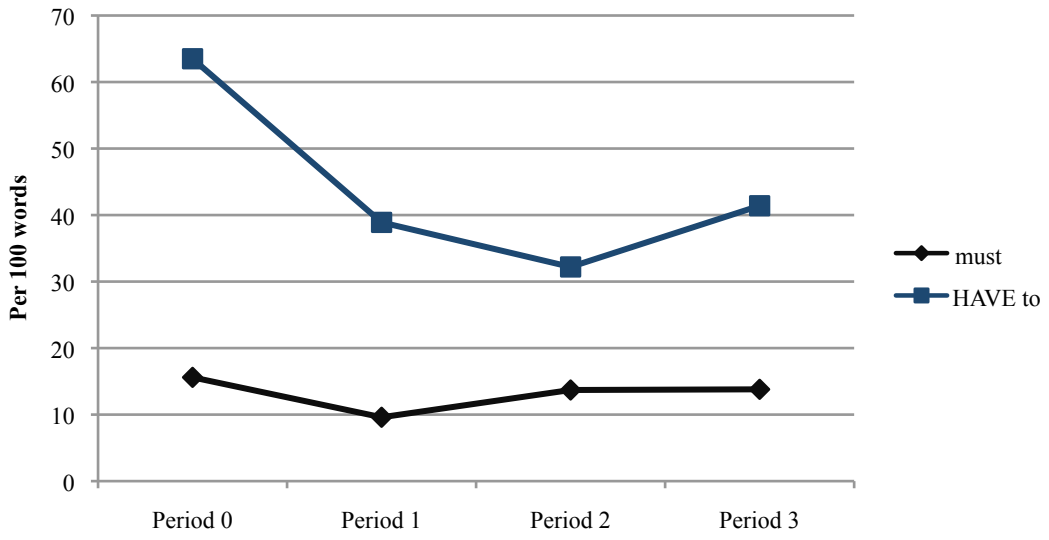

Figure 30

Comparative frequency changes of dynamic modality in *must* and *HAVE to* in SAfE over time

When the normalised instances of deontic and dynamic *must* and *HAVE to* for Period 3 are combined, as to indicate the frequency of root uses in contemporary SAfE, root *HAVE to* is only slightly more frequent than root *must* by 11,8 normalised instances (compare respectively 99,2 and 87,4 normalised instances) – not really enough to confirm a dramatic trend for the dominance of root *HAVE to* by the end of the 20th century. The extent of such dominance is of course influenced by the trend of decline for deontic *HAVE to* in SAfE over the 20th century. Yet dynamic *HAVE to* alone is dominant over dynamic *must* by 27,6 normalised instances by the 1990s (see Figure 30), i.e. it retained its dominance over *must* from Period 0 to 3 in this regard. The increase of the dynamic use of *HAVE to* has not caused the overall increase of this quasi-modal in SAfE, since the simultaneous decrease in deontic meanings had rather a stabilising effect on its 20th-century frequency (see § 4.2.1.4.1). The fact that dynamic *must* remains stable across the 20th century supports the finding that this modal has moved towards a degree of monosemy on a macro-level, since this stable

trend does not impede the effect of the rising deontic and falling epistemic meaning, although it does contribute toward the overall relatively stable trend for *must* from Period 2 to 3.

Ergo, the macrosemantic ecology of the obligation and necessity cluster for contemporary SAfE tends toward the following: *must* and *should* have near-equal propensity to convey epistemic meanings, whereas *must* is the more popular option to convey deontic meanings (just as it was in Period 0), even if the three auxiliaries share more equal levels of popularity by the 1990s than in previous periods, and *HAVE to* is more likely to be used in dynamic contexts.

Figure 31 illustrates the microsemantic shifts of *must*, *should* and *HAVE to* in deontic contexts over the course of the diachronic data.

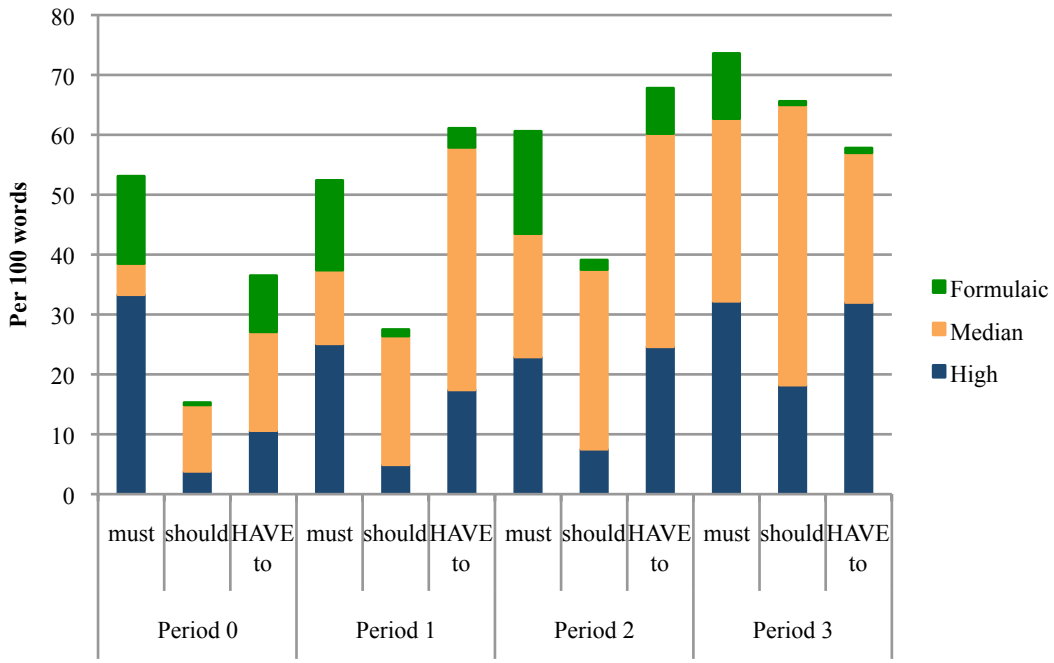

Figure 31

Comparative microsemantic frequency changes of deontic *must*, *should* and *HAVE to* in SAfE over time

When comparing the microsemantic frequencies for the three auxiliaries in question, it becomes clear that *must* is the more popular option to express deontic formulae across all periods. As reported in § 4.3.2.2.2, § 4.3.3.2.2 and § 4.3.2.2.2, the formulaic use of *must*, *should* and *HAVE to* is however in decline toward the end of the 20th

century, indicating a move away from more formal writing formulae, and hence toward a degree of colloquialisation.

Figure 31 further shows that the proportions of high- and median-degree obligation uses even out toward the last period, both for each auxiliary individually, and in terms of the proportional ecology of the semantic cluster. This appears to be the main motivation for the evening out of frequency differentiation in deontic meanings in the 1990s, which, as stated above, in turn motivates the less differentiated pattern for overall frequencies of SAfE *must*, *should* and *HAVE to* by the end of the 20th century (cf. § 4.2.1.4.1).

The popularity of SAfE *must* over its rivals in expressing a high-degree of obligation is clear in Period 0 and 1, but the rising popularity of *HAVE to* in expressing such meanings is also visible from Period 1 onward. In both Period 2 and 3 *must* and *HAVE to* share more equal levels of popularity in this regard. The high-degree obligation meaning was not a popular option for deontic *should* from Period 0 to 2, but in Period 3 this use rises in frequency, but remains lower that that of its rivals. As noted above, the rise in high-degree *should* contributes toward a degree of evening out the proportions between high- and median-degree *should* in the last period. In Period 0 *HAVE to* is the more popular option to convey median-degree obligation, and remains thus in Period 1, but loses some ground to both *should* and *must* in Period 2, and is finally outranked in this respect by both its rivals in Period 3. *Should* is therefore the more popular option for expressing median-degree levels of obligation by the late 20th century, but not to such an extent as to completely overshadow *must* and *HAVE to*. The degree of lessened differentiation between the expressions of micromeanings among these auxiliaries is largely responsible for this.

The increase of median-degree SAfE *should* over the 20th century agrees with the trend for this modal in other native varieties (cf. Leech *et al*., 2009:115; Millar, 2009:203), but the fact that the high-degree use of SAfE *should* has also increased, does not agree with this trend. The increase of high-degree SAfE *HAVE to* from Period 2 onward – even to become slightly more popular than the median-degree use by Period 3, also does not correspond with the finding for other native varieties that this quasi-modal is becoming more face-saving (cf. Leech *et al*., 2009:115). As for SAfE *must*, its growing frequency of median-degree uses over the course of the data is unlike the trends in other native Englishes (cf. § 2.4.2.2). Its near-equal proportions of high- and median-degree uses by the 1990s indicates that this modal has not become

discourse orientated toward use in face-threatening contexts, like in other varieties (cf. Palmer, 1990:69-70; Leech *et al*., 2009:114), and consequently has not declined in overall frequency as a result of this restriction.

Figure 32 provides a look at the late 20th-century proportional makeup of the three auxiliaries in question, according to the source of the obligation as it co-occurs with high- and median-degree uses.

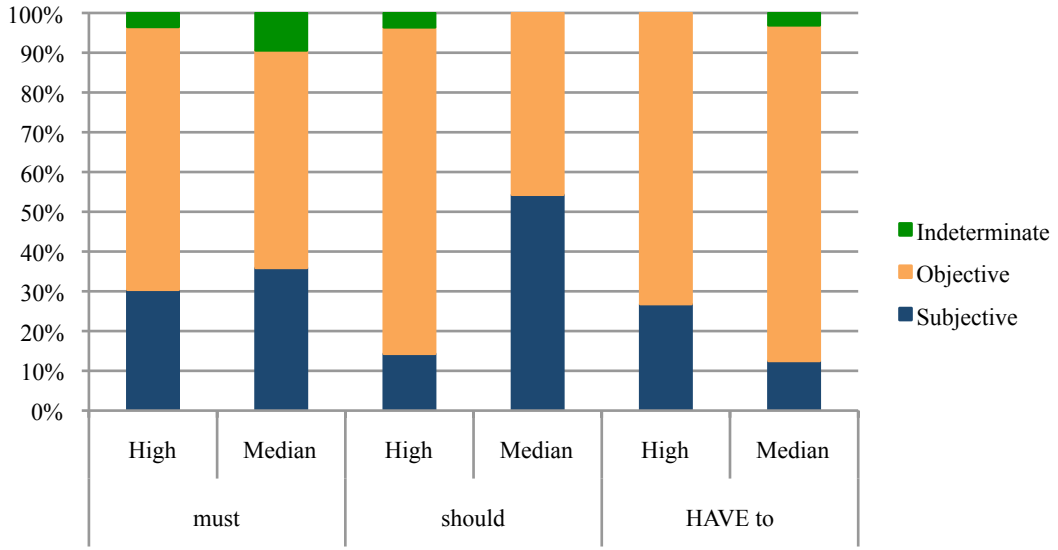

Figure 32

Proportional tendencies for high- and median-degree *must*, *should* and *HAVE to* to occur with objective and subjective sources in SAfE in the 1990s

This figure shows that, for high-degree uses, the options co-occurring most frequently with the subjective source is *must* and *HAVE to*, which is followed by *should*. As an effect, high-degree *should* is the option that co-occurs most frequently with the objective source, followed by *HAVE to* and then *must*. *Should* is also the median-degree option co-occurring most frequently with the subjective source, above *must* and especially *HAVE to*. Median-degree *HAVE to* is however the option that co-occurs most frequently with the objective source, followed by *must* and *should*. In other words, by the 1990s, both high-degree *must* and *HAVE to* often link with the subjective source, whereas high-degree *should* more often links with the objective (cf. § 4.3.2.2.1). For the median-degree uses, *should* is most often linked with the subjective source and least often with the objective, whereas median-degree *HAVE to* most often links with the objective source and least often with the subjective (cf. § 4.3.2.2.1).

Median-degree *must* maintains a middle position in the propensity to be linked with both subjective and objective sources (cf. § 4.3.2.2.1). Since there is no micromeaning of either *must*, *should* or *HAVE to* link exclusively with one source or the other, the pattern of SAfE in this regard agrees with the findings Depraetere and Verhulst's (2008:4) finding (for BrE) that "no particular source is exclusively associated with one particular auxiliary" (cf. § 2.4.1.1).

It was noted in § 4.3.2.2.1 that the average percentages across all periods, as indicators of traditional tendencies, agree with the tendencies for Period 3, which suggests that the traditional uses are remaining popular into the 1990s. This moreover involves the tendency for high-degree *HAVE to* to be associated with objective sources, and the tendency for high-degree *must* to be associated with subjective sources. This suggests a link between the strength and the source of obligation in SAfE. But, because the frequencies of high-degree *must* and *HAVE to* are more similar during Period 3 than on average, it was suggested that the prototypically traditional uses are losing ground, and that high-degree *must* and *HAVE to* may be moving toward a degree of synonymy on this micro-level.

When comparing the formulaic uses of *must*, *should* and *HAVE to*, the tendencies for the pairing of these auxiliaries with objective and subjective sources are different (not illustrated in this section). Formulaic *must*, which is noted above as the modal that is most frequently used in such contexts among its rivals, generally prefers the subjective source, but is also used with objective sources in all of the periods, as seen in Table 10, while *should* almost exclusively prefers objective sources, with the exception of a subjective instance in Period 2, as seen in Table 15, but the overall occurrence of formulaic *should* is lower than formulaic *must*, as seen in e.g. Figure 31. Formulaic *HAVE to*, on the other hand, occurs exclusively with subjective sources, as seen in Table 18. As far as the formulaic uses for *should* and *HAVE to* go, there is a distinguishable link between each auxiliary with one of the sources, pointing toward some degree of micro-level of monosemy, as far as each auxiliary maintaining its own semantic ecology or niche is concerned, but the more polysemous tendency of formulaic *must* on a micro-level however complicates this trend. One trend that is more clear, however, is the movement away from excessive politeness formulae, which is detectable for all three the auxiliaries discussed in this main section, as suggested by the general decline of formulaic uses, especially over the 20th century.

Hence, for the contemporary microsemantic ecology of this semantic cluster, the following is clear: the proportional difference between high- and median-degree uses of *must*, *should* and *HAVE to* has narrowed toward the end of the 20th century, whereas the formulaic uses of these auxiliaries have declined. High-degree *must* is more popular than its rivals in subjective contexts, whereas high-degree *should* is more popular in objective contexts. Median-degree *should* with subjective sources is the more popular option, and median-degree *HAVE to* is more popular in objective contexts.

4.3.6 Conclusion: semantic results

Firstly, the macrosemantic analyses for the obligation and necessity cluster revealed that, in contemporary SAfE, *must* and *should* share near-equal propensity to convey epistemic meanings (with epistemic *should* rising from Period 2 to 3), *must* in itself is more inclined to convey deontic meanings (an indication of monosemy, seeing that epistemic *must* declines and dynamic *must* remains stable), although *should* and *HAVE to* also frequently express deontic meanings, and *HAVE to* tends to convey both dynamic on top of deontic meanings, with a closing gap between the frequencies of these two meanings toward the 1990s (an indication of polysemy). There is also a tendency for *should* toward a degree of stylistic specialisation, in that all its macromeanings have become more frequent in informative or argumentative contexts, e.g. non-fiction, over the 20th century.

Secondly, for the microsemantic interaction between *must* and *should*, it was shown that a degree of synonymy, i.e. a semantic overlap, between the modals can be detected regarding median-degree uses, mostly with objective sources, but also with the subjective source (see examples [183], [184], [186], [194] and [197]). This suggests that SAfE *must* has moved into the traditional semantic domain of *should*, viz. the lower or median-degree obligative use, which in turn is indicative of a polysemous relationship on a microsemantic level between these two modals. Indeed the arguments for simultaneous polysemy of each modal individually and synonymy between the modals, is based on the fact that for both *must* and *should* high- and median-degree uses have reached ratios that are narrower than previously, by the end

of the 20th century, as noted in § 4.3.3.2.1. Therefore, there is strong support for Bowerman's (2004b:477) claim, that was mentioned in § 1.1.2, namely that SAfE *must* "often substitutes for polite should". This is however not the entire story. The same trend toward micro-level polysemy and hence synonymy with *must* and to a degree *should*, was also detected for *HAVE to* in § 4.3.2.2.1. Hence, there is evidence of micro-level semantic overlap in the ecologies of the three major contenders in the obligation and necessity cluster based on modal strength. This account does not correlate with the trend toward micro-level monosemy especially reported for *should* in other native varieties (as the modal that most frequently conveys weaker or median meanings), nor with the trend for *must* to be restricted to high-degree obligative contexts (cf. Myhill, 1995:163). SAfE *HAVE to*, in becoming more polysemous in SAfE, does also not follow the trend toward monosemy in other native varieties, where it was found to progressively convey more meanings that are not face threatening, i.e. weaker meanings. This three-way relationship sheds new light on the manifestation of the democratisation process in SAfE, revealing that the higher frequency of *must* over *should* and *HAVE to* does in fact point toward this process, and not away from it, due to *must* expressing a near-equal ratio of median- and high-degree uses by the end of the 20th century, which is very different from the manifestation of this process in other native varieties, involving the decline of *must* due to the restriction to high-degree obligative contexts.

Finally, in conclusion of the usefulness of the parameters given in 3.4.1.1 for the analysis of high- and median-degree obligation in the case of *must*, *should* and *HAVE to*, all six parameters proved valid means of microsemantic interpretation. It was shown that a final analysis of the most salient micromeaning could not be produced in each case by applying only one parameter, or one part of a parameter, but the productivity of these parameters moreover lies in their application as a set, i.e. it is most useful to the interpretation when all parameters are considered at once. Those parameters that proved the most productive in the entire scope of all the deontic microsemantic analyses of the three auxiliaries in question, were found to be, in order of usefulness, parameter 5 (pragmatic functions), parameter 6 (broad textual analysis), parameter 2 (adjuncts/expressions), parameter 1 (subject and main verb), parameter 4 (polarity/temporality) and, finally, parameter 3 (mood of nearby clauses).

4.4 CONCLUSION

Chapter four has presented, discussed and interpreted both the general quantitative and the macro- and microsemantic results of the corpus analysis, mainly on a diachronic, but also on a synchronic level. The trends arising from the diachronic results, especially for the 20th century, as well as the visible tendencies in the synchronic results, have been compared to the findings for other native varieties of English. It has been found that the SAfE modals decrease more and the quasi-modals increase less than in other native varieties, i.e. these tendencies in SAfE agree with the tendencies for other varieties of English, but the extent of the SAfE trends is different. The colloquialisation process has therefore been found not to have advanced to the same extent as in other native varieties in terms of the overall rise of quasi-modals over the 20th century. The diachronic results have further shown that modal frequencies across all four registers (letters, news, fiction/narrative and non-fiction), as well as quasi-modal frequencies in letters and news, are converging in the 1990s, but this narrowing stylistic gap between registers could not be directly linked to colloquialisation.

Colloquialisation has been found to be more evident on a synchronic level, in terms of the higher quasi-modal frequency in spoken SAfE, and in the trends of some individual modals and quasi-modals, both on a diachronic level, viz. *shall* and *BE going to*, and on a synchronic level, viz. *will* and *BE going to*. Support for colloquialisation taking place has further been found in the decline of archaic or formalised options like *BE to* and *shall* and their restriction to non-fiction and news in SAfE, i.e. distributional fragmentation. Evidence for a trend toward monosemy over the 20th century has been produced for *may* and *can*, in terms of respectively rising epistemic and dynamic possibility meanings, as well as for *shall*, in terms of its restriction to dynamic volition meanings. In addition, *shall* has been reported to become restricted to first-person subjects and *need* to negative polarity contexts, i.e. paradigmatic atrophy.

SAfE modality has been found to be especially unique among other native varieties in terms of both the fact that *must* does not show a significant decline over the 20th century, but rather becomes the highest-frequency modal in the obligation and necessity cluster, as well as the fact that its synchronic spoken frequency is higher

than its written frequency in the 1990s. Synchronic results for contemporary Afrikaans have been compared with the results for contemporary SAfE, which produced quantitative support for the influence of contact with Afrikaans on SAfE regarding the use of *must*. The microsemantic analysis of Afrikaans *moet/moes* and SAfE *must* in spoken, interactive contexts has indeed provided qualitative support for this influence as well. The diachronic semantic results for the three major role players in the obligation and necessity cluster have revealed that SAfE *must* is particularly unique among the other native varieties of English, in that it is the most popular option among its rivals to convey deontic meanings by the 1990s, and just as likely as *should* to be used in epistemic contexts. Indeed the semantic results for *must* have indicated a tendency toward a macro-level of monosemy for this modal in terms of its high deontic frequency by the 20th-century.

Furthermore, the microsemantic results have shown that SAfE *must* is just as likely to be used in median-degree obligation contexts as in high-degree obligation contexts (i.e. a micro-level of polysemy), and that this modal is more likely to occur with objective than subjective sources for both high- and median-degree uses by the end of the 20th century. The gradual increase of the use of *must* in median-degree obligation contexts over the 19th and 20th centuries, and the resultant overall rise of this modal, have been found to be indicative of the democratisation process, which is a very different manifestation of this process than in other native varieties, where democratisation has been found to be visible in the restriction of *must* to high-degree contexts and its consequent overall decline. The 20th-century tendency toward more equal proportions of high- and median-degree uses in *must*, *should* and *HAVE to* has revealed a micro-level of synonymy between these three auxiliaries. Lastly, the decline in formulaic uses of *must*, *should* and *HAVE to* over the 20th century has been found to be indicative of colloquialisation. This chapter has therefore not only provided a detailed diachronic and synchronic description on SAfE modality and compared its patterns with those of other Englishes, but has also explored the reasons behind these patterns. These reasons have been found to be both internal and external to language, i.e. they are linked with both text and context.

CHAPTER 5

CONCLUSION

5.1 INTRODUCTION

Chapter 5 will summarise the quantitative and qualitative findings of the corpus-based study in terms of the research questions, objectives and goals proposed in Chapter 1, as well as conclude on the theoretical implications of these findings for SAfE among the native English varieties of the world. The latter conclusion will centre on the issues of grammaticalisation, the propagation of features and the phase in which SAfE finds itself in terms of Schneider's model of the evolution of New Englishes (2003; 2007), as well as the role of language contact in the development of this variety. After recommendations for further research are made, the final section will show that this thesis has broadened the scope of the description of SAfE beyond that of lexis and phonology, and into the grammatical and intricate semantic domains of a variety that is complex in its context, both locally and in the world.

5.2 SUMMARY OF FINDINGS ACCORDING TO RESEARCH QUESTIONS, OBJECTIVE AND GOALS

This section will summarise the quantitative and qualitative findings of this study, based on the research questions, objectives and goals set out in Chapter 1. All eight the goals set in § 1.5 have been achieved:

1. A historical corpus of written SAfE consisting of letters, news, fiction/narrative and non-fiction has been compiled to represent this variety from the 1820s to the 1990s (§ 3.2.2.1).
2. Each instance of modal and quasi-modal verbs in the historical corpus as well as synchronic corpora of SAfE and Afrikaans has been identified (§ 4.2).
3. The distributional patterns (synchronic) and patterns of change (diachronic) have been analysed in general and in each register, the results of which have been presented in tabular or graphic form and illustrated via examples from the corpus (§ 4.2.1 and § 4.2.2).
4. Each occurrence of *must, should* and *HAVE to* has been analysed in-depth in terms of macro- and micro semantics, namely deontic (high- or median-degree obligation together with the subjective and objective source of the obligation), epistemic and dynamic meanings (§ 4.3).
5. The discerned semantic patterns were compared across registers (§ 4.3.2.1, § 4.3.3.1 and § 4.3.4.1).
6. The findings regarding modal patterns in Afrikaans corpora were compared with those in SAfE, and support for the influence of language contact has been found on both quantitative and qualitative levels (§ 4.2.2.3 and § 4.3.2.2.3).
7. General and modal-specific comparisons of the findings with those of BrE and AmE, as well as AusE and NZE have been done as they arose, both in terms of frequency and semantics (§ 4.2 and § 4.3).
8. The sociohistorical and linguistic implications of the findings and comparisons have been discussed as they arose (e.g. related to colloquialisation, democratisation, language contact, etc.) (§ 4.2 and 4.3), and will continue to be discussed in § 5.3).

In view of the above, the objective to produce an empirical study of SAfE modality based on corpus evidence has been reached. Accordingly, quantitative and qualitative interpretations have been carried out on the modals and quasi-modals and comparisons with the results of other native Englishes have also been drawn. This was done with the aim of contributing a piece of the puzzle that is SAfE grammar.

Six specific research questions were posed in § 1.5 in order to reach the objectives, all of which are restated below, along with direct answers to each of these questions based on the results presented and discussed in Chapter 4.

1

Has the use of modal and quasi-modal verbs changed in the (relatively short) history of SAfE (around 190 years[197])?

Yes. The following changes have occurred in SAfE on a quantitative level (§ 4.2.1.1):

- **The modals have gradually increased from the early 19th to the mid-20th century, but then decreased by 15,3% toward the end of the 20th century[198].**
- **The quasi-modals have shown a slight overall increase from the early 19th to the late 20th century.**

The following changes in frequency have occurred from Period 2 to 3/the 1910s to the 1990s (§ 4.2.1.6) (this period was emphasised for its comparability to modal studies in other native varieties of English). The semantic clusters under which the modals and quasi-modals are grouped are given in brackets:

- **Modals that have increased:**
  - *Can* and *may* (permission, possibility and ability).
- **Quasi-modals that have increased:**
  - *NEED to, BE supposed to* and *(had) better* (obligation and necessity).
  - *BE going to* (prediction and volition).
- **Modals and quasi-modals that have increased, but with a degree of stability setting in:**
  - *Will* and *WANT to* (prediction and volition).

- **Modals that have decreased:**
  - *Could* and *might* (permission, possibility and ability).
  - *Would* and *shall* (prediction and volition).
  - *Need* and *ought* (obligation and necessity).

---

[197] Since the corpus data of this thesis stops in 1999, the study more realistically represents about 170 years of development, but when deeming the 1990s representative of contemporary SAfE, the extension to 190 years in this question is not wholly unrealistic, in terms of the results for Period 3 embodying the current state of the variety.

[198] Much the same situation is present for AmE in the Brown family of corpora, with an increase in modal frequency from the 1930s to 1961, but a decrease from 1961 to 1992 (Mair, 2014).

- **Quasi-modals that have decreased:**
    - *BE able to* (permission, possibility and ability).
    - *BE to* (obligation and necessity).
- **Modals and quasi-modals that have decreased, but with a degree of stability setting in:**
    - *Must* and *should* (obligation and necessity).
    - *HAVE to* and *(HAVE) got to* (obligation and necessity).

A proportional frequency shift was reported between *must, should* and *HAVE to* over the 20th century:

- **The results indicated a movement from more differentiated frequencies during Period 0, 1 and Period 2, to a lesser degree of differentiation among *must, should* and *HAVE to* by Period 3.**

The following general distributional patterns were detected in the synchronic data (ICE-SA) for the 1990s (§ 4.2.2.2):

- **The modals are more frequent than quasi-modals in both speech and writing.**
- **The modals have a slightly higher frequency in speech than in writing with a ratio of 1,08:1.**
- **The quasi-modals have a much higher frequency in speech than in writing with a ratio of 2,67:1.**

The following distributional patterns were detected in the synchronic data regarding individual modals and quasi-modals:

- **Higher frequencies in speech were reported for:**
    - Modals: *can*, *could*, *might*, *must*, *should*, *need*, *ought* and *would.*
    - Quasi-modals: *BE able to*, *HAVE to, (HAVE) got to, NEED to, BE supposed to*, *(had) better BE going to* and *WANT to.*

- **Higher frequencies in writing were reported for:**
    - Modals: *may*, *shall* and *will.*
    - Quasi-modals: *BE to.*

The following synchronic findings (regarding the obligation and necessity cluster) were reported to be the most noteworthy:

- ***Must* is more frequently used in speech than in writing.**
- ***Must* is overall more frequent than *should* (in speech and writing).**
- ***HAVE to* only slightly outranks *must* because of its slightly higher frequency in speech, (*HAVE to* is also the only auxiliary in this semantic cluster to outrank *must* in its spoken frequency).**
- ***Should* is only slightly more frequent in speech than in writing.**
- **No shift in weight between *must*, *should* and *HAVE to* relative to one another in speech or writing was found, since all three options are more frequent in speech.**

2

How have the different registers (social and business letters, news, fiction and other narrative texts, and non-fiction) evolved in terms of modal and quasi-modal use, i.e. what noticeable changes took place over time in the relationship between register and modality?

The register-internal results revealed the following for the modals and quasi-modals (§ 4.2.1.1):

- **The modal frequencies in the four registers show a trend toward convergence in the 1990s, but this could not be directly linked with colloquialisation.**
- **The quasi-modal frequencies in the four registers are internally more unstable than the modal frequencies.**

- **Letters and news more prominently move toward convergence (a narrowing stylistic gap) by the end of the 20th century for:**
    - The modals: the frequency in letters drops and rises in news.
    - And quasi-modals: the frequencies in both letters and news rise.
- **The modal frequency in fiction/narrative and non-fiction remains stable over the 20th century, but modal frequencies in letters and news still converge with those of these registers by the 1990s.**
    - Evidence for colloquialisation in terms of the register-internal patterns for modals is however not strong, based on these frequencies alone.
- **Although quasi-modals decrease in fiction/narrative and non-fiction over the 20th century, quasi-modal frequency in fiction/narrative is closer to that of news and letters, than for non-fiction by the 1990s.**
    - The increase of quasi-modals in letters may indicate a level of colloquialisation in the letter register.
    - News may be undergoing a degree of colloquialisation, since the quasi-modal frequency in news apparently rises to 'meet' the frequency in letters (to the degree that the letter register is more informal).
    - Fiction/narrative has also moved closer to this 'target' point.
    - Given the fact that letters, news and fiction/narrative are all quite balanced for varying formality levels, the 'target' being may be at the 'grey' middle area of the spectrum.
    - Evidence for colloquialisation is not strong, but it is somewhat stronger for the quasi-modals than the modals, especially in news.

These findings therefore seem to reveal more about changes in the registers themselves than about modality per se[199].

For the individual modals and quasi-modals, the following register-internal trends surfaced (§ 4.2.1):

- **Prominent semantic trade-offs in some registers were noted between:**
    - *Can* and *could* (*can* tends to convey more dynamic possibility meanings in especially the non-fiction of the 1990s).

[199] This is also suggested by Mair (2014) for the registers/genres in the extended Brown family of corpora.

- *Can* and *BE able to* (*can* tends to convey more dynamic ability meanings in especially the social letters of the 1990s).

- **Simultaneous register-internal increases and decreases were noted for *will*:**
  - These increases and decreases cancel each other out, and create some stability in the frequency of *will* over the course of the data, and especially over the last two periods. This level of stability is supported by the overall continuity of dynamic and epistemic meanings.

- **Deontic *must* decreases in letters over the 20th century, but increases in the three remaining registers.**
  - This supports the general rising trend of the deontic meaning for *must* over the 20th century, with fiction/narrative being the main carrier of the trend.

- **The letter register is the only register in which all four meanings of *should* decrease (most notably the deontic meaning) by the 1990s.**
  - This supports the general decrease of *should* in letters from Period 2 to 3, which is the main motivation for the significant decline of *should* over the 20th-century.
  - A 20th-century tendency for *should* toward a degree of stylistic specialisation was noted:
    - All the macromeanings of *should* have become more frequent in informative/ argumentative contexts, e.g. non-fiction.

- **By the 1990s, the dynamic meaning of *HAVE to* is more often conveyed in non-fiction and letters, and the deontic meaning in news and fiction/narrative.**
  - This supports the movement towards more equal frequencies of deontic and dynamic meanings for *HAVE to* by the 1990s, which supports the overall stable trend of this quasi-modal over the 20th century.

3

Are there reasons for such changes in (1) embedded in the relevant semantics?

Yes. To summarise for the macro- and microsemantic ecology of obligation and necessity by the late 20th century (§ 4.3.5):

- ***Must* and *should* have near-equal propensity to convey epistemic meanings.**
- ***Must* is the more popular option to convey deontic meanings by the 1990s.**
    - *Must* returns to its early 19th-century role as the dominant carrier of this meaning in the late 20th century.
    - *Must*, *should* and *HAVE to* however share more equal levels of popularity in deontic contexts by the 1990s than in previous periods.
- ***HAVE to* is the more popular contender in dynamic contexts.**
- **The proportional difference between high- and median-degree uses of *must*, *should* and *HAVE to* has narrowed by the 1990s and formulaic uses have declined for all three these contenders.**
- **Among the three major contenders of this semantic cluster:**
    - High-degree *must* is the more popular option in subjective contexts.
    - High-degree *should* is the more popular option in objective contexts.
    - Median-degree *should* is more popular in subjective contexts.
    - Median-degree *HAVE to* is more popular in objective contexts.

These trends for sources of obligation, although generally resembling the traditional view of these modals and quasi-modal in contemporary English, were in fact found to be more complex than this generalisation, in terms of the individual relationships between source and force, as subsequent summaries will show.

The following semantic findings provide explanations for the above-noted finding that the overall frequencies of *must*, *should* and *HAVE to*, although slightly decreasing toward the end of the 20th century, have moved toward a degree of stability, both in term of each of these auxiliaries individually (diachronically), and as a group in the 1990s, in comparison with earlier periods (§ 4.3.2.2.2, § 4.3.3.2.2 and § 4.3.2.2.2):

- **By the late 20th century *must* markedly tends toward deontic uses (macro-level monosemy) that:**
    - Are less face-threatening, in that uses with a median degree of force are just as frequent as those with a high degree of force (micro-level polysemy), which is indicative of democratisation.
    - Are more often objective than subjective for both high- and median-degree uses (i.e. no clear link between modal source and strength for *must*).

- Are less formalised, which is indicative of colloquialisation.

- **By the late 20th century *should* tends toward deontic uses (macro-level monosemy) that:**
  - Are characterised by a tendency toward a smaller proportional difference between high- and median-degree uses (a tendency toward micro-level polysemy).
  - Reveal no distinct move toward weaker obligative meanings.
  - Show a propensity for high-degree uses to co-occur with objective sources.
  - Show a tendency for median-degree uses to co-occur with both objective and subjective sources on near-equal levels, but with a slight tendency toward subjective sources (i.e. some indication of a link between source and strength for *should*).
  - Are less formalised, which is indicative of colloquialisation.

- ***Should* shares a macro-degree of synonymy with *must.***
  - Both tend toward deontic meanings (monosemy).

- ***Should* also shares a micro-degree of synonymy with *must.***
  - Both tend toward a smaller proportional difference between high- and median-degrees of obligation (polysemy).
  - Both tend toward the co-occurrence of median-degree uses with objective and subjective sources with near-equal propensity (supporting Bowerman's (2004b:477) claim of interchangeability between *must* and *should* in SAfE.
  - *Should* however still has a slightly stronger tendency for conveying median-degree meanings than *must* in the 1990s, and the slight tendency toward subjective sources in this meaning.
    - This suggests that *should* has retained a tinge of politeness.

- **By the late 20th century both deontic and dynamic uses remain dominant for *HAVE to*, with a narrowing of the proportional difference between these uses (i.e. macro-level polysemy).**
  - Deontic *HAVE to* shows a trend toward near-equal proportions of high- and median-degree uses (i.e. a micro-level of polysemy), with only a slight inclination toward conveying a high degree of obligation.
  - Both the high- and median-degree uses of *HAVE to* show a shift toward more balanced proportions of subjective and objective sources in the 1990s than in preceding periods, but the objective source remains more popular for both high-

and median-degree uses (i.e. no clear indication of a link between source and strength for *HAVE to*).

- The near disappearance of formulaic *HAVE to* by the 1990s points toward the colloquialisation process.

- ***HAVE to* shares a degree of micro-level synonymy with *must* and *should*.**
  - By the 1990s, all three the major contenders in the obligation and necessity cluster tend toward more balanced proportions between:
    - High- and median-degree uses.
    - Objective and subjective sources as they co-occur with high- and median-degree uses.
  - This supports the overall more stable diachronic patterns of each of these three auxiliaries over the 20th century, as well as the move toward more balanced proportions among the rivals in terms of overall frequency by the 1990s.

4

What influence could social elements such as language contact have on the changes?

The influence of Afrikaans on SAfE has been established on a social level (§ 2.3.3.4), in terms of three factors, viz. migration, contact and identity, with contact being especially prominent. It was reported that the complex and extensive social and linguistic contact situation between the Afrikaners and English-speaking South Africans involves the following: collective and individual relationships, bilingualism, Afrikaans/Dutch input into early SAfE in the 19th century, the maintenance of separate identities, a 'racial affinitive relationship' under apartheid and the military situation during the last half of the 20th century, the continual influx of different strands of British settlers, and Afrikaans teachers teaching English to white communities during the 20th century (cf. e.g. Lanham, 1978:157-8; Branford, 1996:36-9; Giliomee and Mbenga, 2007:206; 433-437; Bekker, 2012:136).

The effect of this contact situation on SAfE modality has been found to manifest on both quantitative and qualitative levels. The synchronic results provided quantitative support for Afrikaans influence on SAfE modality in terms of the following (§ 4.2.2.3):

- **Support was found for a 'dragging' effect or 'pull' of the frequency of *moet/moes* on the frequency of its semantic equivalents, including *must*, *HAVE to*, *(HAVE) got to* and *NEED to*, in terms of overall frequency and register distribution (speech and writing).**
    - Such a quantitative 'pull' is likely to be compounded in the case of *must*, due to its grammatical, phonological and general semantic correspondence with *moet/moes*, which the other semantic equivalents of *moet/moes* generally lack.

A comparison of the microsemantics between deontic *must* in spoken SAfE and deontic *moet/moes* in spoken Afrikaans yielded qualitative support for the role of contact in shaping SAfE modality (§ 4.3.2.2.3):

- **Where they collocate with second-person subjects, both spoken SAfE *must* and spoken Afrikaans *moet/moes* clearly prefer the deontic meaning[200] – specifically the median obligation meaning (with near-equal percentages).**
    - This supports the idea that SAfE *must* follows an Afrikaans-like pattern in its semantics, especially in interactive contexts.
        - This pattern involves the precedence of the weaker deontic force of *must* over the stronger, as it is for *moet/moes.*

[200] Bao (2010) reports that *must* is predominantly deontic and has declined in its epistemic use in Singapore English (SingE) as well, which he attributes to language contact with Chinese (Bao, 2010:1727). This renders SingE *must* a similar case to SAfE *must*, insofar that both have undergone contact-induced change in similar macro-directions. There are however two major differences: firstly, the relevant contact-induced change in SAfE is adstratum transfer into SAfE (a native variety) from Afrikaans, whereas the change in SingE is substratum transfer into SingE (a nonnative variety) from Chinese (2010:1727), and secondly, the substratum language is not related to the superstratum in the case of Chinese and SingE, whereas the adstratum is related to the superstratum in the case of Afrikaans and SAfE. The nature of the contact situations is therefore very different in the case of these two varieties, no doubt giving rise to the fact that the same microsemantic changes that occurred in SAfE *must* did not occur in SingE – i.e. the macrosemantic changes are not motivated by the same means. The similarity of macrosemantic changes between these two varieties may therefore be linguistic coincidence, or it may be part of a larger emergent pattern in the New Englishes, the latter of which necessitates further investigation. Bao (2010:1727) further emphasises "the importance of [using] computerized corpora in the study of contact-induced linguistic change", which has indeed been the strategy of this thesis.

5

Are the changes similar or dissimilar to those that occurred in other native varieties of English?

The overall diachronic trends for modals and quasi-modals in SAfE were found to agree with the overall trends for AmE and BrE over the 20th century, but the extent of the trends differed (§ 4.2.1.5):

- **The modals of SAfE have decreased more than the modals of AmE and BrE by the end of the 20th century:**
    - This is e.g. motivated by the larger decline of *would* than in other varieties.
        - *Would* especially decreased in hypothetical uses in letters.
            - The lessened degree of formality (by means of politeness and tentativeness) links with the process of colloquialisation.
- **The quasi-modals of SAfE have increased less than the quasi-modals of AmE and BrE by the end of the 20th century.**
    - The smaller increase of SAfE quasi-modals over the 20th century suggest that colloquialisation occurs to a lesser extent than in other varieties (to the degree that the rise of quasi-modals is indicative of colloquialisation).

It was found that the diachronic trends of the individual modals and quasi-modals in 20th-century SAfE did not only correspond to either those of AmE or BrE (§ 4.2.1.6). If correspondence was present, there tended to be differences in the extent of agreement. However, in the cases where SAfE did corresponded to other varieties, the more common denominator was AmE, agreeing with the idea of a trend towards Americanisation, so there is some support for American impact on SAfE in terms of modality (cf. Leech *et al.*, 2009:43-4; Trüb, 2008) (§ 4.2.1.5).

- **Cases where SAfE corresponds to both AmE and BrE for the 20th-century trends:**
    - The decline of *ought* and *BE to*.
        - These declines were found to be by-products of their restriction to more formal contexts, indicating colloquialisation in terms of a move away from formal styles (just as for *need*).
    - The decline of *might, shall* and *would* (these modals declined more in SAfE).

        - The decline of e.g. *shall* was reported to be motivated by the shift of epistemic meanings in news and dynamic meanings in letters to *BE going to*, which was linked with colloquialisation.

- **Cases where SAfE corresponds to AmE for the 20th-century trends:**
    - The increase of *BE going to* (it rose less in SAfE).
    - The increase of *NEED to* (it rose with less significance in SAfE).
    - The stability of *HAVE to*.
        - It was noted that the register distribution of *HAVE to* has gained a degree of convergence by the 20th century, which motivates its stable trend.
    - The decline of *need* (it declined more in SAfE).
        - This decline was found to be a by-product of its restriction to more formal contexts, which indicates colloquialisation in terms of a move away from formal styles.
    - The non-significant increase of *BE supposed to.*

- **Cases where SAfE corresponds to neither AmE nor BrE for the 20th-century trends:**
    - More stable trends for *will*, *WANT to*, *(HAVE) got to, should* and especially *must*, as opposed to either significant changes in other varieties (especially in the case of significantly declining *must* in other varieties) or in terms of SAfE frequencies retaining the 'middle ground' between AmE and BrE.
        - It was noted that the register distribution of *must* has gained a degree of convergence by the 20th century (even more so than for *HAVE to*), which motivates its stable trend.
    - The significant increase of *can* (remained more stable in BrE and rose less in AmE).
    - The significant decline of *could* (rose in BrE and AmE).
    - The increase of *may* (declined in BrE and AmE).
    - The slight decline of *BE able to* (increased slightly in BrE and AmE).
    - The non-significant increase of *(had) better* (declined non-significantly in AmE and BrE).

The following general synchronic finding was reported in terms of the apparent progress of colloquialisation in SAfE in comparison to other native English varieties (§ 4.2.2.4):

- **As far as a higher quasi-modal frequency in speech than in writing indicates a growing tendency toward speech-like styles, SAfE may be more advanced than BrE and AusE regarding this tendency, and more conservative than AmE (in terms of the quasi-modal ratio in speech and writing) by the end of the 20th century.**

Noteworthy tendencies in the 1990s for e.g. the modals and quasi-modal *must, should* and *HAVE to* of the obligation and necessity cluster were reported in terms of the tendencies in other native varieties in the late 20th century (cf. Collins, 2005):

- ***Must* is more than twice as frequent as in AmE, BrE, AusE and NZE.**
    - *Must* is more frequent than *should* in contemporary SAfE (*should* is more frequent in AmE, BrE, AusE and NZE).
    - *Must* in SAfE is overall more frequent than most mid-frequency modals and quasi-modals that are more frequent in other varieties.
- ***Must* in SAfE is more frequent in speech than in writing (it is becoming restricted to written contexts in the other native varieties).**
    - Only *HAVE to* in SAfE is slightly more frequent than *must* – overall and in speech (the overall and spoken precedence of *HAVE to* over *must* is much more pronounced in other native varieties).
- **No popularity of *must*, *should* or *HAVE to* over either of the others in terms of a preference for either speech or writing is apparent in the 1990s.**
    - *Must*, *should* and *HAVE to* are all more frequent in speech in SAfE (the restriction of *must* to writing in AmE benefits *should* and *HAVE to* in speech, and the restriction of *must* and *should* to writing in BrE benefits *HAVE to* in speech, which indicates democratisation in these varieties).

The diachronic, semantic results showed that both the macro- and microsemantic ecology of the obligation and necessity group in SAfE by the end of the 20th century is largely different than those of BrE and AmE, but there were some similarities reported as well (§ 4.3.2 - § 4.3.6) (cf. Collins, 2005; 2009a; Leech *et al*., 2009:115-6):

- **Cases where SAfE and other native varieties were found to have similar macrosemantic trends over the 20th century include:**
    - SAfE *HAVE to* is not a major contender for the epistemic meanings in the obligation and necessity cluster.
    - *HAVE to* is the most popular option to convey dynamic meanings in the obligation and necessity cluster.
- **Cases where SAfE and other native varieties were found to have dissimilar macrosemantic trends over the 20th century include:**
    - The decline of epistemic *must* is more dramatic in SAfE than in other native varieties.
    - Epistemic *should* increases in SAfE (it declines in other native varieties).
        - *Must* and *should* are basically equally likely to be used in epistemic contexts in contemporary SAfE (*must* is much more likely to be used in epistemic contexts than *should* in other native varieties).
        - The tendency in SAfE for the 1990s is closest to that of contemporary AmE and AusE, where the gap between epistemic *should* and *must* is smaller, while the gap in BrE and NZE is greater (cf. Collins, 2005:268).
    - Deontic *must* has not declined in SAfE, like in other native Englishes, but has increased steadily since the late 19th century to the late 20th century.
        - There is no dramatic decline of deontic *must* as a by-product of its dramatic decline overall, as in other varieties, but the relative stability of SAfE *must* over the 20th century was reported to be a of its increasing popularity in deontic contexts over its two rivals, together with its decreasing epistemic frequency and its stable dynamic frequency.
        - The degree of macro-level monosemy for SAfE *must*, in favour of deontic meanings, does not agree with the tendency in other native varieties.
    - *Must* is the dominant means of conveying deontic meanings in the obligation and necessity cluster by the 1990s, whereas *should* is the more dominant root option in other native varieties, including other SHEs.
        - Because *should* does not convey dynamic meanings, like *must*, the fact that root *should* surpasses root *must* in other varieties suggests that deontic *should* is dominant over deontic and dynamic *must*, whereas deontic *must* alone is already more frequent than deontic *should* in SAfE.
    - Deontic *HAVE to* decreases in SAfE (deontic and dynamic meanings of *HAVE to* increase in other native varieties).

- SAfE *HAVE to* less frequently conveys deontic meanings than *should* and *must* by the 1990s (*HAVE to* conveys more deontic meanings than *should* and *must* in other native Englishes.

- **Some aspects of the microsemantic ecology of *must*, *should* and *HAVE to* in SAfE agree with those of other native varieties:**
  - Neither *must* nor *should* or *HAVE to* link exclusively with either the objective or subjective source in SAfE, in either of the two micromeanings (high- and median-degree uses), just like in BrE (cf. Depraetere and Verhulst, 2008:4).
  - Although the tendency for *HAVE to* in other native varieties to be dominantly objective agrees with the tendency for median-degree *HAVE to* in SAfE by the 1990s, objective uses have lost ground by this time (for both high- and median-degree uses).

- **The microsemantic ecology of *must*, *should* and *HAVE to* is generally different in SAfE than in other native varieties:**
  - The tendency toward a microsemantic overlap in the ecologies of *must*, *should* and *HAVE to*, i.e. micro-level polysemy within each of these auxiliaries, based on modal strength (more balanced proportions between median- and high-degree uses) and modal source (median-degree uses of *must* and *should* tend to co-occur with objective and subjective sources with near-equal propensity), does not agree with the trend toward micro-level monosemy reported in other native varieties, where:
    - *Should* more frequently conveys median-degree strength.
    - *Must* tends to be restricted to conveying a high-degree strength.
    - *HAVE to* progressively conveys non-face-threatening (median) meanings.
  - On average, from the early 19th to late 20th centuries, there is no clear tendency for high-degree uses of *must* in SAfE to have a subjective source of obligation or for median-degree *must* to have an objective source of obligation more frequently, as generally suggested for native varieties of English (cf. Collins, 2009a:38).
  - The micro-level of synonymy reported between *must*, *should* and *HAVE to* in SAfE disagrees with the tendency in other native varieties for modals and quasi-modals to each find their ecological 'niche' (cf. Leech *et al*., 2009).

- **Democratisation presents differently in SAfE during the 20th century:**
  - This process was found to be evident in the extension of *must* into median-degree obligation contexts and its consequent overall increase, whereas the restriction of *must* to high-degree obligation contexts and its consequent overall decline indicated this process in other native varieties.

6

Why might SAfE have developed in this way in relation to other native Englishes?

Considering the development of *must* in SAfE, this question essentially refers back to question 4 and the role of language contact. In this thesis, it has been argued that SAfE is a unique variety among the SHEs, PCEs and WEs, in terms of:

- **The extended formation period of SAfE, which will be considered in § 5.3.**
- **The complex contact situation between SAfE and a related West-Germanic language, Afrikaans, in which:**
    - The contact proceeded on a large scale and over an extended period of time, as well as under some complex social conditions in the second half of the 20th century.
    - The structural, phonological and broad semantic correspondence between the Afrikaans modal *moet* (and its past form *moes*) (which is often used in median-degree contexts) with English *must*, has facilitated the gradual development of median-degree obligation meanings in near-equal proportions with high-degree obligation meanings in SAfE *must.*

In the attempt to arrive at answers to the research questions given above, this thesis has essentially traced the trek of modality within written SAfE, (with some comparisons with spoken SAfE and written and spoken Afrikaans), from the year when the first wave of settlers set foot on South African soil, up to its contemporary state, across text (language-internal aspects) and context (language-external aspects). The general theoretical implications of these findings for SAfE are discussed in the next section.

5.3 THEORETICAL IMPLICATIONS FOR THE DEVELOPMENT OF SAfE

This section addresses the implications of the findings for SAfE in terms of, firstly, grammaticalisation, and secondly, the diffusion/propagation of features, the nature and role of language contact in the development of SAfE and the stages in the development of this variety.

5.3.1 Grammaticalisation in SAfE

It was noted in § 2.2.1.1.1 that the extension into new contexts of use (context generalisation) is one of the traditional mechanisms in the grammaticalisation of linguistic expressions (Heine, 2003:578-9). This mechanism relates to pragmatics, by which a linguistic expression can gain "properties characteristic of their uses in new contexts... sometimes to the extent that their meaning may show little resemblance to the original meaning" (2003:579), which links it with semantics. This process generally occurs by means of an 'overlap', i.e. ambiguity between two usage patterns in a linguistic expression (which can be the usage pattern of another expression), which finally leads to the loss of one of these usage patterns, and in some cases even to one of the expressions (usually the 'older' forms) (Heine, 1993:48-53; Harris, 1987:184). This process happens gradually (but could be apparent within one generation of users) (Leech, 2003:223), and the early stages of this process are often indicated by shifts in frequency (Hopper and Traugott, 1993:59).

As far as the findings of this study go, there is support for the conclusion that particularly *must* has extended into new contexts of use in SAfE, in that it moved from primarily conveying high-degree obligation in the 1820s-1860s to gradually conveying more and more median-degree obligation during the 1870s-1900s and the 1910s to 1950s, and finally conveying median-degree and high-degree obligation in near-equal proportions by the 1990s. The reported usefulness of Parameter 5 (pragmatic functions) in the analysis of these micromeanings (§ 4.3.6) supports this conclusion, in that it strengthens the idea that the most salient development of *must* involves the expansion of its contextual functions. The micro-level polysemy in *must* indicates a level of ambiguity between the two patterns of use. This, together with the

tendency for the examples of high- and median-degree *must* to be non-prototypical by the 1990s (they both co-occur more with objective sources, whereas before they showed different preferences in terms of source) – i.e. an overlap in the micromeanings, suggests that grammaticalisation, in its traditional sense, has proceeded in SAfE *must* over the course of 190 years. Since these shifts toward more equal proportions of micromeanings largely occurred between Period 2 (1910s-1950s) and 3 (1990s), the grammaticalisation process can be said to be noticeable within slightly more than one generation (where thirty years represent one generation) in this case, since the data for Period 2 ends in 1959 and the data for Period 3 begins in 1990.

Furthermore, the degree of micro-level synonymy reported between *must* and *should*, as well as *HAVE to* (both *should* and *HAVE to* had similar microsemantic tendencies than *must*, regarding more balanced proportions of degrees of strength than previously, and a preference for the objective source by the 1990s), suggests that a similar kind of grammaticalisation has proceeded over time for *should* and *HAVE to* as well. This level of support for functional change within all three major contenders in the obligation and necessity cluster, contrasts with Myhill's (1995:159) finding for AmE that "there was no functional change in the language" since the 19th century, but that "new forms simply replaced old forms with the same general functions", e.g. *HAVE to* replaced *must*. In SAfE, it is clear that *must*, *should* and *HAVE to* have not replaced each other since the 19th century – they instead have less differentiated frequencies by the end of the 20th century. This does however suggest that, since the three major contenders in this cluster overlap functionally, they may indeed be in competition with each other. It remains to be seen whether the older patterns of use, both for the three rivals individually and across the group as a whole, will be lost in the future, with the newly grammaticalised forms as the survivors.

The general movement towards overall less differentiated frequencies by the 1990s than previously for *must*, *should* and *HAVE to*, and especially the high frequency of *must* among its rivals (Figure 11), support the argument for grammaticalisation, in that the narrowing of the overall frequencies suggests overlapping contexts of use (ambiguity), and that *must*, having extended its functions to the largest extent, appears to be emerging as the more popular form, i.e. the most probable survivor for the foreseeable future. This however, as with all predictions about the future, cannot be taken for granted, especially since the grammaticalisation process appears to be in an earlier rather than later stage. Indeed, the crossover between the frequencies of *must*

and *should* (in favour of *must*) somewhere around the 1960s or 1970s (Figure 10), supports the argument for an earlier stage of grammaticalisation, since the gap between the point of crossover and the endpoint in the 1990s represents one generation or even less.

Apart from the more traditional view on grammaticalisation, the view of Boye and Harder (2012:8-9) was considered in § 2.2.1.1.1, regarding e.g. the mutual interdependency of usage and convention (applying to linguistic properties and social meaning, like positively vs. negatively charged words) in the process of grammaticalisation. In the results of this thesis, this kind of interdependency is suggested in the relationship between the overall more stable frequency of e.g. *must* over the 20th century than previously (as far as it is indicative of usage), and its microsemantic tendency to be used in contexts conveying both a high and a median degree of obligation. It therefore appears that, by convention, the social meaning of SAfE *must* is not as negatively charged anymore by the 1990s as it had been previously (especially in the early 19th century). In other words, the shift in convention over the 20th century for *must* to be nearly equally likely to be used in face-threatening and non-face-threatening situations may have expanded the likelihood of usage events (the opposite of 'constraint', which is mentioned by Boye & Harder [2012]). This ultimately leads to a more stable 20th-century trend in SAfE than for other native Englishes during this period, where *must*, by convention, tends to be more confined to face-threatening contexts, and, as an effect, has declined in its usage. Furthermore, despite the idea of the 'loss' of one form and the rise of another described above, the view of Boye & Harder (2012) that grammaticalisation more likely involves the functional 'enrichment' of one form with another is supported by the tendency toward the degree of micro-level synonymy between *must*, *should* and *HAVE to*, as well as by their less differentiated frequencies of deontic uses by the 1990s.

Whether we consider the more traditional or the newer views on grammaticalisation, it appears as though SAfE shows clear symptoms of this process especially in the past century, in terms of changing contexts of use (pragmatics) and the meanings arising within these contexts (semantics), as well as changing conventions, which reflects in changing frequencies.

5.3.2 The development of SAfE: propagation of features, stages of development and language contact

This section will assess the implications of the results for SAfE in terms its own development (considering the propagation of features and the factor of language contact), as well as in terms of its place within frameworks for English varieties, especially that of Schneider (2003; 2007). It was explained in § 2.2.1.2.1 that, in the evolution of a new variety of English, feature innovation adds a potentially new feature to the feature pool, but that the process of diffusion or propagation involves changes in the probability of occurrence of features (Croft, 2000:3-6). Language contact was noted to be essential in the introduction of new features (as part of Schneider's [2003:244] Phase One), while the propagation of a new feature was noted to be driven by social factors like identity, i.e. identity creates an 'enabling frame' for the adaptation of variants and the spread of these variants across the social ecological strands proposed by Schneider (2003) (see § 2.3.2.4; cf. Van Rooy, 2010; Croft, 2000).

The gradual increase of *must* in median-degree obligation contexts (see Figure 20) can be regarded as evidence for the gradual spread of this new feature over the course of the 19th and 20th centuries, which is enabled and driven by the interaction between the two major white South African identities (white, English-speaking South Africans, and white, Afrikaans-speaking South Africans). The propagation of increasingly less face-threatening *must*, as a feature of SAfE, is therefore driven by contact between identities, which arose from contact between two related languages in the first place, i.e. a connection "between two systems [made] by bilingual speakers" (Van Rooy, 2010:9). The innovation of this feature is however not as simple as a direct borrowing. Since deontic *must* did contain some median-degree uses in Period 0 of this study (Figure 20), i.e. the 1820s-1860s, it cannot really be said that this use *arose* solely due to contact with Afrikaans, since it was already present, albeit very infrequent, in proto-SAfE (the ratio for high to median uses was reported to be 6,4:1 during this period [§ 4.3.2.2.1). The more likely scenario is that the potential for this use to propagate has been *unlocked* by contact with Afrikaans. For such cases e.g. Lass and Wright (1986:201) refer to 'diffused innovation', where, in a contact situation, an "apparently 'original' feature of one [language] turns up in the other" via borrowing or transfer. In other words, Afrikaans contact, as the catalyst or

source in this scenario, provided the linguistic resources (e.g. the cognate form of *must*, *moet*, with a different nuance in meaning), which created the momentum for the spread of a less face-threatening SAfE *must* to commence (slowly at first in Period 1 [1870s-1900s]), and supported the continuity of the spread (with an acceleration of the spread in Period 2 [1910s-1950s]) – finally ending with a ratio of 1,1:1 between high and median uses in Period 3 [1990s]) (§ 4.3.2.2.1). This is supported by the overall pattern for deontic *must*, where the frequency between Period 0 and 1 remain relatively stable, but rises in Period 2 and continues to rise in Period 3, due to the rise of both high- and median-degree uses (see Figures 20, 29 and 31). This suggests that the propagation of less face-threatening *must* in SAfE had only really gained the above-mentioned momentum in the first half of the 20th century (1910s-1950s).

The median-degree use is, of course, not the prototypical use of English *must* (which is on the high-degree, subjective end of the scale [§ 2.4.1.1]), especially when paired with the objective source (as it more often is in SAfE by the 1990s), but belongs to the periphery; hence, the non-prototypical use of deontic *must* has diffused into SAfE to such an extent as to become as frequently used as the prototypical, high-degree use, but with the high-degree use also pairing more frequently with the objective source by the 1990s (Table 8). This suggests that the prototype for SAfE *must* in particular has moved to a point on the continuum (much) closer to the periphery than previously by the end of the 20th century. In a way, this is itself an innovation in this variety, considering the opposite trend for all other native Englishes (and even the nonnative Englishes, including BSAfE, as discussed below [see also Van Rooy & Wasserman, 2014]), even if the innovation is not wholly without roots in the original input system. The tendency that spoken SAfE median-degree *must* is slightly more frequent than spoken Afrikaans median-degree *moet/moes* (see Table 12) may indicate a case where the feature propagation has gained momentum to such an extent that the usage of the feature in the target (SAfE) has overtaken that of the catalyst of the propagation (Afrikaans).

The momentum for the propagation of an increasingly popular and less face-threatening *must* over the 20th century seems to have had a 'pull' on its closest competitor, *should*, in terms of frequency from Period 2 to 3, which is no doubt supported by the extent of micro-level synonymy between these modals (Figure 29). Although this kind of synonymy had permeated *HAVE to* by the 1990s too, this quasi-modal gets left behind in terms of its deontic frequency (Figure 29), with *must* and

should overtaking it at the end of the 20th century. The propagation of less face-threatening deontic *must* has therefore impacted on both its rivals regarding deontic frequencies, but with a different outcome in each case, and indeed established *must* as the more popular option for conveying deontic meanings by Period 3.

In consideration of Baxter *et al*. (2009) and Blythe and Croft's (2012) proposed mechanisms of propagation (§ 2.2.1.2.1), two of these seem to be at work in the case of deontic *must* in SAfE. The mechanism of 'replicator selection' ("the differential weighting of the competing variants in a change" [Blythe & Croft, 2012:269]) is suggested where the differential weighting between high- and median-degree uses in SAfE has narrowed to a point of near symmetry by the 1990s (with much the same process occurring for *should*), but some cross-linguistic symmetry also exists between the differential weighting for *moet/moes* between high- and median-degree uses in Afrikaans (Table 12) (cf. Blythe & Croft, 2012:272). Furthermore, more directly on the level of contact, the mechanism of 'weighted interaction selection' (based on an influential social group propagating a feature) is suggested for SAfE *must* in the qualitative (and quantitative) evidence discussed above. These suggestions are strengthened by what looks like the beginning of an S-curve for deontic *must* in Figure 29 – in indication of these mechanisms becoming activated between Period 1 and 2, and continuing the propagation into Period 3.

In this study, the Afrikaner nation has been established as such an influential social group in the propagating of a feature (see above), not only in terms of language-internal/textual (quantitative and qualitative) patterns, but also on a language-external/contextual (sociohistorical) level. It was noted in § 1.1.1 that different arguments exist regarding the influence of Afrikaans on SAfE in terms of specific grammatical constructions, but the general influence of this language on SAfE has been credited. Lass and Wright (1986:201) for instance emphasised how complex it is to determine the effects of language contact on the transfer of a feature, since endogenous sources for innovation can be just as likely.

Considering the argument of and Mesthrie (2002) for the '*busy* + progressive' construction in SAfE not to be attributable to Afrikaans influence on a qualitative level, but rather to an endogenous development reliant on existing English-internal options (whereas Lass and Wright [1986] argued that Afrikaans may have lifted the semantic restriction in English) (§ 1.1.1), the case of SAfE *must* appears to be different than that of the '*busy* + progressive' construction by Mesthrie's account,

since qualitative support for Afrikaans influence has indeed been produced in this study (§ 4.3.2.2.3). In a sense, the argument for SAfE *must* in this thesis is more in line with the earlier argument of Lass and Wright (1986) for the '*busy* + progressive' construction, in that Afrikaans can be said to have played a crucial role in the lifting of the semantic restriction of *must* over the 19th and especially the 20th centuries, even if the English-internal 'option' was there in the first place, since this 'option' was very limited in earlier stages (both in frequency and in terms of a preference for the objective source in Period 0), but gained unprecedented momentum in later periods in terms of frequency (Figure 20) and more balanced proportions between objective and subjective sources, particularly by Period 3 (Table 9).

To answer the question whether a rival language like Afrikaans, having been stigmatised by English-speaking South Africans, and having come into contact with SAfE under such conditions including acts of discrimination, exclusion and conflict (mostly in the 19th and early 20th centuries) could have influenced SAfE, the linguistic results of this study, as well as the sociohistorical considerations of this thesis point toward a 'yes'. The relationship between the Afrikaners and English-speaking South Africans has been more complicated than the negative conditions mentioned above. Such factors such as intermarriage, bilingualism, contact due to urbanisation or small rural environments, Afrikaans forming part of the input of two of the waves of British settlement, the widespread phenomenon of Afrikaans-speaking English teachers, the accommodation of English speakers to Afrikaans L2 English[201] (cf. e.g. Lanham, 1978:156; Mesthrie & West, 1995:115), etc. point toward a prolonged and close contact situation – even an artificial one when considering the 'racial affinity' between these groups under apartheid laws, as well as inside the South African Army (cf. § 4.2.2.3). Jeffery and Van Rooy (2004:279) indeed note that "[m]utual influence is inevitable when languages have enjoyed (or, sometimes, endured) contact as long and close as that between Afrikaans and English in South Africa". Such influence on SAfE has therefore been established in this thesis in terms of modality (especially

[201] This aspect of the relationship between Afrikaners and English-speaking South Africans appears to link with Trudgill's (2008; 2010) idea of 'behavioural coordination', which is not understood as a social act (i.e. it does not involve identity), but involves accommodation in face-to-face interaction due to an innate biological drive (§ 2.2.1.2.1; cf. Van Rooy, 2010:5). This kind of accommodation is, however, as shown above, only one small aspect of the relationship between the above-mentioned groups, and the other aspects of the relationship mostly involve social situations/factors. Furthermore, Trudgill (2010:192) applies this concept mainly to dialect contact, and not to language contact as such, so the 'link' supposed here should be considered with some caution.

regarding *must*), which supports the findings of Jeffery and Van Rooy (2004) for Afrikaans influencing aspects of SAfE grammar.

In conclusion of the role of contact in shaping SAfE, it is useful to consider whether the findings for *must* match the criteria of Thomason (2007) and Levey and Poplack (2010) for a contact-induced change (as noted in § 2.3.2.4). Since no other native or nonnative variety of English follows the same semantic (especially microsemantic) trends of SAfE *must*, this change in SAfE matches the criterion of Thomason (2007:42) in that it "would have been less likely to occur outside [this] particular contact situation". The findings also match the criteria posed by Levey and Poplack (2010:398), including the specification that the candidate for contact-induced change, *must*, in being present through interlingual coincidence (in the cognate form *moet*) in the source variety, "is not conditioned in the same way as in the source" (*moet* more often conveys median-degree meanings than traditional English *must*), as well as the specification that the candidate (*must*) "can ...be shown to parallel in some non-trivial way the behaviour of a counterpart feature in the source" (*must* and *moet* share phonological, structural and broad semantic correspondence due to the relatedness between English and Afrikaans). According to these criteria, therefore, the case for *must* being influenced by contact with Afrikaans is strong.

When examining the effect of the semantic development of *must* on the development of SAfE in terms of the model and its application to SAfE proposed by Schneider (2003; 2007), and considering the extended formation of the variety (cf. Bekker, 2012), the following conclusions can be drawn. If Schneider's Phase Two for SAfE begins during the mid-1930s and transitions into Phase Three around the late 1970s or early 1980s (drawing on Bekker's (2012) ideas – see § 2.3.3.3), then an exonormative 'pull' would be exerted on SAfE during this phase. Such an argument is indeed supported by the findings for the microsemantics of *must*, where such a 'pull' seems to be keeping the increase of deontic *must* more modest between Period 1 and 2 (i.e. up to the 1950s), influenced by the increase of median uses, than between Period 2 and 3 (up to the 1990s), where deontic *must* increases with acceleration, influenced by the increase of both high and median uses. Furthermore, the publication of Beeton and Dorner's prescriptive work in 1975 agrees with the idea that the 1970s formed part of the transition between the second and third phase, since their prescriptions are largely based on British norms (characteristic of Phase Two), and is indicative of a complaint tradition arising in South Africa (characteristic of Phase Three). Seeing that

the overall frequency patterns of *must* and *should* intersect somewhere around the 1960s/70s (with *must* overtaking *should* as the more frequent option in their semantic cluster) (Figure 10) this supports the idea that Phase Three may have ensued during this era (i.e. the latter part of 20th century). Additionally, the fact that the deontic trends of *must* and *HAVE to* intersect more or less around the same time, with *must* emerging as the more popular deontic option (Figure 29) is a further reinforcing factor. The balancing out of the microsemantic proportions of *must* by the 1990s (Figure 20) seems to suggest that 'nativization' (Phase Three) had already begun by this time, but that is very likely still in progress.

The massive scale of social remodelling that occurred in 1994 was shown earlier in this thesis to be less impressive in terms of identity remodelling of the entire population (§ 2.3.3.3), so, based on SAfE modality, it can be argued that Phase Three is not yet complete for the STL strand itself by the end of the 20th century.

This is supported by the comparative findings for SAfE and BSAfE (the IDG strand) reported in Van Rooy and Wasserman (2014), concerning both the overall diachronic patterns of the modals and quasi-modals and the individual diachronic patterns of the modals *must* and *should* over the last half of the 20th century. It was found that "[t]he direction of change [for combined modals and quasi-modals] in BSAfE is exactly opposite from the direction of change in [SAfE]" (modals increase and quasi-modals only increase as part of the general trend of "conveying modality overtly more often") (Van Rooy & Wasserman, 2014:63). The gradual development of micro-level synonymy between SAfE *must* and *should* reported in this thesis is also "not shared by BSAfE, where the meanings of these modals remained remarkably unchanged from the middle to the end of the twentieth century" (i.e. no qualitative convergence), although *must* and *should* in BSAfE also "remain in widespread use, and do not decline dramatically in frequency", like in SAfE (i.e some quantitative correspondence) (2014:63). The same motivation for the stability of *must* in SAfE cannot however be attributed to BSAfE, so the quantitative correspondence does not necessarily show convergence of use. Convergence between the STL and IDG strands is indeed the criterion set by Schneider (2003; 2007) "for the evolution of a postcolonial variety of English to endonormative stabilization" (Phase Four) of his model (Van Rooy and Wasserman, 2014:63). Since no such convergence takes place in the 20th century, the argument that SAfE still lingers in the 'nativization' phase is strengthened in the bigger scope of South African Englishes in general.

To summarise the implications of the findings in terms of Schneider's model (in comparison with the application by Schneider [2007]) (see e.g. § 3.2.2.1):

- **Period 0 together with Period 1 in the corpus (the 1820s-1860s and the 1870s-1900s) represents almost the whole duration of the extended foundation phase (Phase One).**
    - Schneider (2007:175-176) considers Phase One to span 1806 to the 1822/1870s.
    - According to the findings for SAfE modality, Phase One extends another 30 or 50 years after the time that Schneider (2007) dates its termination.
    - This phase includes all three waves of British settlement, the discovery of diamonds and gold, the Anglo-Boer War and the formation of the Union government (since it actually extends beyond the 1910s).
- **Period 2 in the corpus (the 1910s-1950s) represents the transition into 'exonormative stabilization' (Phase Two), since the onset for Phase Two probably occurs more or less the middle of Period 2 (the mid-1930s).**
    - Schneider (2007:178) considers Phase Two to begin between 1822 and the 1870s and to last up until 1910.
    - According to the findings for SAfE modality, Phase Two only really begins 20 years or so after the time that Schneider (2007) dates its termination.
    - This phase begins about twenty years after Union formation, and extends into the age of the formation of the Republic of South Africa, during which apartheid laws were formally introduced (1960s-70s).
- **Period 3 in the corpus (1990s) represents early inroads into the 'nativization' phase (Phase Three), since the onset of this phase can be predicted to have been around the late 1970s, but perhaps more probably around the early 1980s.**
    - Schneider (2007:247) considers Phase Three to span 1910 to 1994, and hence Phase Four ('endonormative stabilization') to have its onset in 1994.
    - According to the findings for SAfE modality, Phase Three only really begins a decade or so before the time that Schneider (2007) dates its termination, and no evidence for the onset of Phase Four was found in the 1990s data.
    - This phase largely develops during times of political unrest (including e.g. terror attacks), various protests against the apartheid regime ending in violence, as well as the era of the South African Army (§ 2.3.3.4.1).
    - It therefore appears as though substantive nativisation from Afrikaans to SAfE could have occurred even without any comprehensive alignment of social identities in this case. However, the unique 'racial affinitive' relationship between

speakers of Afrikaans and English in South Africa, especially during the last half of the 20th century, could have affected the alignment of 'racial identities' – preparing the conditions for nativisation.

According to this account of SAfE modality, a so-called 'event X' is yet to take place in South Africa, and, since no linguistic convergence based on comprehensive identity revision between the STL and IDG strands has occurred, even after 1994 with the formation of the democratic government (cf. Schneider, 2003; 2007). SAfE data covering the 21st century might provide more insight into this subject in the future.

5.4 RECOMMENDATIONS FOR FURTHER RESEARCH

The following extensions of the research topic of this thesis could cast more light on the trends for modality, grammar, register variation, etc. both within SAfE itself, and in terms of comparing these trends to those of other native and nonnative varieties of English.

- **The expansion of the historical corpus of SAfE.**
  - The compilation of newer data with which to update the findings for the 21st century and hence trace the more recent trends.
    - This will enable comparison with the more contemporary findings for AmE and BrE e.g. regarding modality and register variation (as in e.g. Leech, 2011; Mair, 2014).
  - The compilation of data for the 1960s to 1980s, with which to fill the only chronological gap in the corpus.
    - This could provide more clarity as to the onset of Schneider's (2003) third phase in SAfE.
- **The diachronic analysis of Afrikaans modality.**
  - A diachronic corpus of Afrikaans has recently been compiled by J. Kirsten of the North-West University, covering the early 20th to early 21st centuries.
    - This new resource could help compare the development of SAfE *must* with that of Afrikaans *moet/moes* for instance, to investigate whether these cognate modals have developed in a similar way over the 20th century.

- **A multi-dimensional study of SAfE registers.**
  - With the goal of tracing the evolution of the various registers in the historical corpus.
    - This will reveal more about the nature and progression of the colloquialisation process in SAfE.
- **The expansion of the linguistic focus to include other aspects of the verb phrase in SAfE.**
  - These include e.g. tense, aspect and evidentiality, as well as other kinds of modal expressions.

Much more insight into the grammar of SAfE is clearly to be gained, but by focusing on modality in this study, as the first big attempt toward describing the grammar of SAfE, a wider platform, from which to launch its full description and comparison with other English varieties, has been created.

5.5 CONCLUSION

The final conclusion will show that this thesis has broadened the description of SAfE beyond that of lexis and phonology, and into the grammatical and intricate semantic domains of a variety that is complex in both its local and international context.

This study has described the diachronic development of modality in SAfE from the early 19th century up to its contemporary state (1820s to 1990s) in the registers of letters, news, fiction/narrative and non-fiction, on the basis of the theoretical framework of sociohistorical linguistics and the empirical approach of corpus linguistics. Quantitative and qualitative analyses have been conducted for modal and quasi-modal verbs, by means of the historical corpus of SAfE and ICE-SA (with the addition of Afrikaans corpora for comparison). The study therefore explored both general frequency changes and macro- and microsemantic changes. The findings have been compared with the trends for modality in other native Englishes, such as AmE, BrE and AusE, with the focus on the full description of SAfE modality, as opposed to the traditional focus on linguistic anomaly. This is a step towards the full description of SAfE grammar, which will aid the comparison of SAfE with other

native varieties of English, in terms of variation in the verb phrase and beyond, as well as provide a clearer picture regarding the nature of the development of a new English variety in a postcolonial setting as complex as the South African context.

The findings of this study indicate that the trends of modality in SAfE correspond to those of other native varieties in some cases, and deviate from them in others. SAfE was found to be especially unique among other varieties in its quantitative and qualitative trends for *must*, and in the manifestation of the democratisation process. Considering the extended formation period of SAfE and the modal evidence found in terms of the progress of Schneider's (2003) phases (§ 5.3.2), the findings of this thesis support Bekker's (2012:143) argument that "SAfE is in fact the youngest of the colonial varieties of English", especially in the Southern Hemisphere. In this thesis I have analysed what is deemed one of the most intricate linguistic areas of study, namely modality, within one of the most complex social settings in the native English-speaking world, namely the 19th- and 20th-century context of SAfE. It was shown that the social structure of English-speaking South Africans has altered greatly over the past two centuries, which has brought about contact-induced linguistic change. In essence, the trek of modality in South African English reflects the trek of the variety as such within the complex setting of post-colonial and indeed post-apartheid South Africa: both have approached the winding path of change with a kind of stubborn resilience to outside expectations, rendering SAfE modality and hence SAfE as a New English, unique in the world.

APPENDIX 1
WORD COUNTS FOR THE HISTORICAL CORPUS OF SAfE

Table 19: Word counts for the Historical Corpus of SAfE (Period 0, 1, 2 and Period 3 [a selection from written ICE-SA], per decade and register

Decade	News	Fiction/Narrative	Letters		Non-fiction	SUBTOTAL
			Private/social letters	Formal/business letters		
1820s	22121	12366	5388	29063	6976	75914
1830s	8325	13079	2742	0	0	24146
1840s	11530	11229	537	1909	7584	32789
1850s	4564	10790	0	424	0	15778
1860s	26093	11101	7424	826	0	45444
Per 0 Total	**72633**	**58565**	**16091**	**32222**	**14560**	**194071**
1870s	16341	22810	6003	3819	12107	61080
1880s	18201	23604	7073	1145	4054	54077
1890s	4308	7928	9844	1617	6977	30674
1900s	3261	7284	9706	1883	3221	25355
Per 1 Total	**42111**	**61626**	**32626**	**8464**	**26359**	**171186**
1910s	486	0	1219	4506	7775	13986
1920s	8330	4363	3343	4027	3318	23381
1930s	16813	9435	3878	6034	3071	39231
1940s	1852	8975	4802	2365	13897	31891
1950s	4392	7795	10824	635	9030	32676
Per 2 Total	**31873**	**30568**	**24066**	**17567**	**37091**	**141165**
SUBTOTAL	**146617**	**150759**	**72783**	**58253**	**78010**	**506422**
1990s[1]	39455	35346	8206	6209	64132	153348
TOTAL	**186072**	**186105**	**80989**	**64462**	**142142**	**659770**

[1] A selection from the written component of ICE-SA.

APPENDIX 2
DIACHRONIC AND SYNCHRONIC RAW AND NORMALISED FREQUENCIES OF MODALS AND QUASI-MODALS IN SAfE.

Table 20.1: Raw diachronic frequencies of modals in SAfE

Period	Register	*can*	*could*	*may*	*might*	*must*	*shall*	*should*	*need*	*ought*	*will*	*would*	TOTAL Modals
0	Letters	111	51	104	30	46	73	53	4	6	258	121	857
0	News	144	68	123	42	93	68	114	8	16	296	196	1168
0	Fiction/narrative	38	157	32	20	46	18	80	0	8	87	121	607
0	Non-fiction	23	22	30	15	7	7	14	1	2	22	28	171
	TOTAL	**316**	**298**	**289**	**107**	**192**	**166**	**261**	**13**	**32**	**663**	**466**	**2803**
1	Letters	142	58	68	22	72	52	100	5	9	234	123	885
1	News	75	49	90	56	46	17	97	2	8	161	160	761
1	Fiction/narrative	47	113	33	34	32	9	55	1	3	57	124	508
1	Non-fiction	54	26	47	15	37	29	31	9	3	85	62	398
	TOTAL	**318**	**246**	**238**	**127**	**187**	**107**	**283**	**17**	**23**	**537**	**469**	**2552**
2	Letters	131	74	90	20	71	51	89	5	4	255	175	965
2	News	23	33	18	13	32	6	23	1	1	114	92	356
2	Fiction/narrative	32	88	9	25	32	16	23	2	2	53	112	394
2	Non-fiction	66	57	41	22	40	7	52	1	3	65	95	449
	TOTAL	**252**	**252**	**158**	**80**	**175**	**80**	**187**	**9**	**10**	**487**	**474**	**2164**
3	Letters	35	11	16	0	14	7	14	0	0	61	31	189
3	News	66	44	38	15	32	0	41	0	1	209	97	543
3	Fiction/narrative	76	96	6	13	68	2	32	0	0	64	99	456
3	Non-fiction	209	51	139	17	60	4	67	2	0	199	56	804
	TOTAL	**386**	**202**	**199**	**45**	**174**	**13**	**154**	**2**	**1**	**533**	**283**	**1992**

Table 20.2: Raw diachronic frequencies of quasi-modals in SAfE

Period	Register	*BE able to*	*HAVE to*	*(HAVE) got to*	*NEED to*	*BE supposed to*	*BE going to*	*WANT to*	*(had) better*	*BE to*	TOTAL Quasi-modals
0	Letters	17	16	0	0	1	4	3	1	35	77
0	News	20	22	0	1	3	11	2	1	39	99
0	Fiction/narrative	15	45	0	0	0	8	6	4	22	100
0	Non-fiction	5	2	0	0	0	0	0	0	23	30
	TOTAL	**57**	**85**	**0**	**1**	**4**	**23**	**11**	**6**	**119**	**306**
1	Letters	23	28	1	2	5	9	20	1	34	123
1	News	15	26	0	2	0	0	2	1	41	87
1	Fiction/narrative	6	54	6	0	1	6	6	2	31	112
1	Non-fiction	6	18	0	2	1	5	0	0	9	41
	TOTAL	**50**	**126**	**7**	**6**	**7**	**20**	**28**	**4**	**115**	**363**
2	Letters	24	34	0	3	0	5	17	1	13	97
2	News	12	25	0	0	1	0	6	0	13	57
2	Fiction/narrative	7	24	4	1	1	22	17	1	12	89
2	Non-fiction	15	35	1	4	1	4	6	0	17	83
	TOTAL	**58**	**118**	**5**	**8**	**3**	**31**	**46**	**2**	**55**	**326**
3	Letters	6	17	0	1	0	6	8	4	2	44
3	News	12	43	0	19	1	13	20	0	11	119
3	Fiction/narrative	6	26	2	4	3	17	30	0	7	95
3	Non-fiction	31	42	0	17	2	3	13	0	10	118
	TOTAL	**55**	**128**	**2**	**41**	**6**	**39**	**71**	**4**	**30**	**376**

Table 21.1: Normalised diachronic frequencies of modals in SAfE per 100 000 words

Period	Register	*can*	*could*	*may*	*might*	*must*	*shall*	*should*	*need*	*ought*	*will*	*would*	TOTAL Modals
0	Letters	229,8	105,6	215,3	62,1	95,2	151,1	109,7	8,3	12,4	534	250,5	1773,8
0	News	198,3	93,6	169,3	57,8	128	93,6	198,3	11	22	407,5	269,8	1608,1
0	Fiction/narrative	64,9	268,1	54,6	34,2	78,5	30,7	136,6	0	13,7	148,6	206,6	1036,5
0	Non-fiction	158	151,1	206	103	48,1	48,1	96,2	6,9	13,7	151,1	192,3	1174,5
	TOTAL	**162,8**	**153,6**	**148,9**	**55,1**	**98,9**	**85,5**	**134,5**	**6,7**	**16,5**	**341,6**	**240,1**	**1444,3**
1	Letters	345,6	141,2	165,49	53,5	175,2	126,6	243,4	12,2	21,9	569,5	299,3	2153,8
1	News	178,1	116,4	213,7	133	109,2	40,4	230,3	4,7	19	382,3	379,9	1807,1
1	Fiction/narrative	76,3	215,8	53,5	55,2	51,9	14,6	89,2	1,6	4,9	92,5	201,2	824,3
1	Non-fiction	204,9	98,6	178,3	56,9	140,4	110	117,6	34,1	11,4	322,5	235,2	1509,9
	TOTAL	**185,8**	**143,7**	**139**	**74,2**	**109,2**	**62,5**	**165,3**	**9,9**	**13,4**	**313,7**	**274**	**1490,8**
2	Letters	314,7	177,7	216,2	48	170,5	122,5	213,8	12	9,6	612,5	420,3	2317,9
2	News	72,2	103,5	56,5	40,8	100,4	18,8	72,2	3,1	3,1	357,7	288,6	1116,9
2	Fiction/narrative	104,7	287,9	29,4	81,8	104,7	52,3	75,2	6,5	6,5	173,4	366,4	1288,9
2	Non-fiction	177,9	153,7	110,5	59,3	107,8	18,9	140,2	2,7	8,1	175,2	256,1	1210,5
	TOTAL	**178,5**	**178,5**	**111,9**	**56,7**	**124**	**56,7**	**132,5**	**6,4**	**7,1**	**345**	**335,8**	**1533**
3	Letters	242,8	76,3	111	0	97,1	48,6	97,1	0	0	423,2	215,1	1311,1
3	News	167,3	111,5	96,3	38	81,1	0	103,9	0	2,5	529,7	245,8	1376,3
3	Fiction/narrative	215	271,6	17	36,8	192,4	5,7	90,5	0	0	181,1	280,1	1290,1
3	Non-fiction	325,9	79,5	216,7	26,5	93,6	6,2	104,5	3,1	0	310,3	87,3	1253,7
	TOTAL	**251,7**	**131,7**	**129,8**	**29,3**	**113,5**	**8,5**	**100,4**	**1,3**	**0,7**	**347,6**	**184,5**	**1299**

Table 21.2: Normalised diachronic frequencies of quasi-modals in SAfE per 100 000 words

Period	Register	*BE able to*	*HAVE to*	*(HAVE) got to*	*NEED to*	*BE supposed to*	*BE going to*	*WANT to*	*(had) better*	*BE to*	TOTAL Quasi-modals
0	Letters	35,2	33,1	0	0	2,1	8,3	6,2	2,1	72,5	159,4
0	News	27,5	30,3	0	1,4	4,1	15,1	2,8	1,4	53,7	136,3
0	Fiction/narrative	25,6	76,8	0	0	0	13,7	10,2	6,8	37,6	170,8
0	Non-fiction	34,3	13,7	0	0	0	0	0	0	158	206
	TOTAL	**29,4**	**43,8**	**0**	**0,5**	**2,1**	**11,9**	**5,7**	**3,1**	**61,3**	**157,7**
1	Letters	56	68,1	2,4	4,9	12,2	21,9	48,7	2,4	82,7	299,3
1	News	35,6	61,7	0	4,7	0	0	4,7	2,8	97,4	206,6
1	Fiction/narrative	9,7	87,6	9,7	0	1,6	9,7	9,7	3,2	50,3	181,7
1	Non-fiction	22,8	68,3	0	7,6	3,8	19	0	0	34,1	155,5
	TOTAL	**29,2**	**73,6**	**4,1**	**3,5**	**4,1**	**11,7**	**16,4**	**2,3**	**67,2**	**212,1**
2	Letters	57,6	81,7	0	7,2	0	12	40,8	2,4	31,2	233
2	News	37,6	78,4	0	0	3,1	0	18,8	0	40,8	178,8
2	Fiction/narrative	22,9	78,5	13,1	3,2	3,2	72	55,6	3,2	39,3	291,2
2	Non-fiction	40,4	94,4	2,7	10,8	2,7	10,8	16,2	0	45,8	223,8
	TOTAL	**41,1**	**83,6**	**3,5**	**5,7**	**2,1**	**22**	**32,6**	**1,4**	**39**	**230,9**
3	Letters	41,6	117,9	0	6,9	0	41,6	55,5	27,7	13,9	305,2
3	News	30,4	109	0	48,2	2,5	32,9	50,7	0	27,9	301,6
3	Fiction/narrative	17	73,6	5,7	11,3	8,5	48,1	84,9	0	19,8	268,8
3	Non-fiction	48,3	65,5	0	26,5	3,1	4,7	20,3	0	15,6	184
	TOTAL	**35,9**	**83,5**	**1,3**	**26,7**	**3,9**	**25,4**	**46,3**	**2,6**	**19,6**	**245,2**

Table 22: Combined diachronic raw and normalised (per 100 000 words) frequencies of auxiliaries in SAfE

Period	Register	Raw	Normed
0	Letters	934	1933,2
0	News	1267	1744,4
0	Fiction/narrative	707	1207,2
0	Non-fiction	201	1380,5
	TOTAL	**3109**	**1602**
1	Letters	1008	2453,2
1	News	848	2013,7
1	Fiction/narrative	620	1006,1
1	Non-fiction	439	1665,5
	TOTAL	**2915**	**1702,9**
2	Letters	1062	2550,9
2	News	413	1295,8
2	Fiction/narrative	483	1580,1
2	Non-fiction	532	1434,3
	TOTAL	**2490**	**1763,9**
3	Letters	233	1616,4
3	News	662	1677,9
3	Fiction/narrative	551	1558,9
3	Non-fiction	922	1437,7
	TOTAL	**2368**	**1544,2**

Table 23.1: Raw synchronic frequencies of modals in SAfE

Register	*can*	*could*	*may*	*might*	*must*	*shall*	*should*	*need*	*ought*	*will*	*would*	TOTAL Modals
Written	386	202	199	45	174	13	154	2	1	533	283	1992
Spoken	1702	584	188	161	530	25	423	10	10	1034	1028	5695
TOTAL	2088	786	387	206	704	38	577	12	11	1567	1311	7687

Table 23.2: Raw synchronic frequencies of quasi-modals in SAfE

Register	*BE able to*	*HAVE to*	*(HAVE) got to*	*NEED to*	*BE supposed to*	*BE going to*	*WANT to*	*(had) better*	*BE to*	TOTAL Quasi-modals
Written	55	128	2	41	6	39	71	4	30	376
Spoken	162	646	251	151	44	766	595	19	25	2659
TOTAL	217	774	253	192	50	805	666	23	55	3035

Table 24.1: Normalised synchronic frequencies of modals in SAfE per 100 000 words

Register	*can*	*could*	*may*	*might*	*must*	*shall*	*should*	*need*	*ought*	*will*	*would*	TOTAL Modals
Written	251,7	131,7	129,8	29,3	113,5	8,5	100,4	1,3	0,7	347,6	184,5	1299
Spoken	418,4	143,6	46,2	39,6	130,3	6,1	104	2,5	2,5	254,2	252,7	1399,9
TOTAL	372,8	140,3	69,1	36,8	125,7	6,8	103	2,1	2	279,7	234	1372,3

Table 24.2: Normalised synchronic frequencies of quasi-modals in SAfE per 100 000 words

Register	*BE able to*	*HAVE to*	*(HAVE) got to*	*NEED to*	*BE supposed to*	*BE going to*	*WANT to*	*(had) better*	*BE to*	TOTAL Quasi-modals
Written	35,9	83,5	1,3	26,7	3,9	25,4	46,3	2,6	19,6	245,2
Spoken	39,8	158,8	61,7	37,1	10,8	188,3	146,3	4,7	6,1	653,6
TOTAL	38,7	138,2	45,2	34,3	8,9	143,7	118,9	4,1	9,8	541,8

Table 25: Combined synchronic raw and normalised (per 100 000 words) frequencies of auxiliaries in SAfE

Register	Raw	Normed
Written	2368	1544,2
Spoken	8354	2053,5
TOTAL	10722	1914,1

APPENDIX 3

LOG LIKELIHOOD SCORES OF MODAL AND QUASI-MODAL CHANGE IN SAfE BETWEEN PERIODS TO ASSESS STATISTICAL SIGNIFICANCE

Table 26.1: Log likelihood scores[2] of modals over time[3]

Period	*can*	*could*	*may*	*might*	*must*	*need*	*ought*	*shall*	*should*	*will*	*would*	Total
0-1	-1.52	+0.27	+0.35	-2.81	-0.48	-0.53	+0.53	+3.56	-3.01	+1.20	-2.25	-0.75
1-2	+0.13	*-3.80*	+2.91	+2.21	-0.97	+1.01	+1.83	+0.30	+3.44	-1.46	*-6.31*	-0.58
2-3	*-12.42*	*+7.13*	-1.34	*+9.28*	+0.42	*+3.96*	*+5.06*	*+38.92*	*+4.69*	-0.01	*+44.40*	*+19.36*
0-3	*-19.24*	+1.69	+1.29	*+8.18*	-1.06	*+5.06*	*+17.23*	*+72.16*	*+5.23*	-0.05	*+7.14*	*+7.67*

[2] According to Rayson's online calculator (http://ucrel.lancs.ac.uk/llwizard.html) "[t]he log-likelihood value itself is always a positive number", but the log likelihood calculator "compares relative frequencies" between two periods "in order to insert an indicator for '+' overuse and '-' underuse" of the earlier period relative to the later period. The '+' and '-' signs in Tables 26.1 and 26.2 are therefore not indicative of an increase or decrease respectively, but rather of a higher or lower frequency in the preceding period. Therefore, the higher the log likelihood score, "the more significant is the difference between two frequency scores". A log likelihood score "of 3.8 or higher is significant at the level of $p < 0.05$" and a score "of 6.6 or higher is significant at $p < 0.01$". The following list is provided on Rayson's website (as also mentioned in Section 3.3.1: 95th percentile, 5% level, $p < 0.05$; critical value = 3.84; 99th percentile, 1% level, $p < 0.01$, critical value = 6.63; 99.9th percentile, 0.1% level, $p < 0.001$, critical value = 10.83; 99.99th percentile, 0.01% level, $p < 0.0001$, critical value = 15.13.

[3] The statistically significant scores are italicised.

Table 26.2: Log likelihood scores of quasi-modals over time

Period	*BE able to*	*BE going to*	*BE supposed to*	*BE to*	*(had) better*	*HAVE to*	*(HAVE) got to*	*NEED to*	*WANT to*	**Total**
0-1	+0.00	+0.00	-0.68	-0.28	+0.68	*-7.71*	*-5.55*	-1.93	*-4.72*	*-7.91*
1-2	-2.07	-2.99	+0.68	*+7.48*	+0.34	-0.63	+0.00	-0.40	*-6.02*	-0.82
2-3	+0.32	-0.19	-0.68	*+6.23*	-1.05	+0.00	+1.93	*-14.45*	-2.15	-0.41
0-3	-0.76	*-4.67*	-0.68	*+21.74*	+0.00	*-12.71*	-1.39	*-30.19*	*-34.89*	*-18.93*

APPENDIX 4
SYNCHRONIC COMPARATIVE FREQUENCIES OF *MUST, SHOULD* AND *HAVE TO* IN SAfE AND OTHER NATIVE VARIETIES

1. The data used for Figure 15 (Collins, 2009a; Collins, 2005):

Table 27: Combined spoken and written frequencies of *must, should* and *HAVE to* in ICE-corpora across native English varieties per million words

	SAfE	AusE	NZE	BrE	AmE
must	1256,7	613	714	675	402
should	1030	1141	1577	1124	850
HAVE to	1381,7	1311	1182	1244	1385

2. The data used for Figure 16 (Mair & Leech, 2006; Leech *et al.*, 2009):

Table 28: The spoken and written frequencies of *must, should* and *HAVE to* per million words in ICE-SA compared to F-LOB and Frown, and the BNC and LCSAE

Register	SAfE			BrE			AmE		
	must	***should***	***HAVE to***	***must***	***should***	***HAVE to***	***must***	***should***	***HAVE to***
written	1135	1004	835	814	1147	825	668	787	643
spoken	1303	1040	1588	717	1045	1616	297	1189	2589

APPENDIX 5
RAW AND NORMALISED DIACHRONIC FREQUENCIES FOR THE MACRO- AND MICROSEMANTIC ANALYSES OF *MUST*, *SHOULD* AND *HAVE TO* IN SAfE

Table 29.1: Raw macrosemantic frequencies of SAfE *must, should* and *HAVE to* over time

	Period	Epistemic	Deontic	Dynamic	Indeterminate	Quasi-subjunctive	Preterite	TOTAL
must	0	57	102	30	3			192
	1	65	98	18	6			187
	2	44	106	24	1			175
	3	20	128	24	2			174
	TOTAL	**186**	**434**	**96**	**12**			**728**
should	0	6	40		1	115	99	261
	1	13	78		4	95	93	283
	2	11	73		1	58	44	187
	3	14	101		3	30	6	154
	TOTAL	**44**	**292**		**9**	**298**	**242**	**885**
HAVE to	0	0	31	54	0			85
	1	0	77	49	0			126
	2	0	80	38	0			118
	3	1	74	53	0			128
	TOTAL	**1**	**262**	**194**	**0**			**457**

Table 29.2: Normalised macrosemantic frequencies of SAfE *must, should* and *HAVE to* per 100 words (%) over time

	Period	Epistemic	Deontic	Dynamic	Indeterminate	Quasi-subjunctive	Preterite
must	0	29,7	53,1	15,6	1,6		
	1	34,8	52,4	9,6	3,2		
	2	25,1	60,6	13,7	0,6		
	3	11,5	73,6	13,8	1,1		
	TOTAL	**25,5**	**59,6**	**13,2**	**1,6**		
should	0	2,3	15,3		0,4	44,1	37,9
	1	4,6	27,5		1,4	33,6	32,9
	2	5,9	39		0,5	31	23,5
	3	9,1	65,6		1,9	19,5	3,9
	TOTAL	**5**	**33**		**1**	**33,7**	**27,3**
HAVE to	0	0	36,5	63,5	0		
	1	0	61,1	38,9	0		
	2	0	67,8	32,2	0		
	3	0,8	57,8	41,4	0		
	TOTAL	**0,2**	**57,3**	**42,5**	**0**		

Table 30.1: Raw macrosemantic frequencies of SAfE *must* per register over the 20th century

	Deontic		Dynamic		Epistemic		TOTAL
	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3	
Letters	45	10	11	2	15	2	**85**
News	23	30	3	1	6	1	**64**
Fiction/narrative	14	37	3	15	14	15	**98**
Non-fiction	24	51	7	6	9	2	**99**
TOTAL	**106**	**128**	**24**	**24**	**44**	**20**	**346**

Table 30.2: Normalised macrosemantic frequencies of SAfE *must* per register per 100 words (%) over the 20th century

	Deontic		Dynamic		Epistemic	
	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3
Letters	42,5	7,8	45,8	8,3	34,1	10
News	21,7	23,4	12,5	4,2	13,6	5
Fiction/narrative	13,2	28,9	12,5	62,5	31,8	75
Non-fiction	22,6	39,9	29,2	25	20,5	10
TOTAL	**30,7**	**37**	**6,9**	**6,9**	**12,7**	**5,8**

Table 31.1: Raw macrosemantic frequencies of SAfE *should* per register over the 20th century

	Deontic		Epistemic		Quasi-subjunctive		Preterite		TOTAL
	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3	
Letters	35	7	3	2	25	4	26	1	**103**
News	5	25	3	6	9	8	5	2	**63**
Fiction/narrative	10	27	1	0	8	2	4	1	**53**
Non-fiction	23	42	4	6	16	16	9	2	**118**
TOTAL	**73**	**101**	**11**	**14**	**58**	**30**	**44**	**6**	**337**

Table 31.2: Normalised macrosemantic frequencies of SAfE *should* per register per 100 words (%) over the 20th century

	Deontic		Epistemic		Quasi-subjunctive		Preterite	
	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3
Letters	47,9	6,9	27,3	14,2	43,1	13,3	59,1	16,7
News	6,8	24,8	27,3	42,9	15,5	26,7	11,4	33,3
Fiction/narrative	13,8	26,7	9,1	0	13,8	6,7	9,1	16,7
Non-fiction	31,5	41,6	36,3	42,9	27,6	53,3	20,4	33,3
TOTAL	**21,7**	**30**	**3,3**	**4,1**	**17,2**	**8,9**	**13**	**1,8**

Table 32.1: Raw macrosemantic frequencies of SAfE *HAVE to* per register over the 20th century

	Deontic		Dynamic		Epistemic		TOTAL
	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3	
Letters	25	8	9	11	0	0	**53**
News	6	21	13	9	0	0	**49**
Fiction/narrative	13	23	11	15	0	0	**62**
Non-fiction	36	22	5	18	0	1	**82**
TOTAL	**80**	**74**	**38**	**53**	**0**	**1**	**246**

Table 32.2: Normalised macrosemantic frequencies of SAfE *HAVE to* per register per 100 words (%) over the 20th century

	Deontic		Dynamic		Epistemic	
	Period 2	Period 3	Period 2	Period 3	Period 2	Period 3
Letters	31,3	10,8	23,7	20,7	0	0
News	7,5	28,4	34,2	17	0	0
Fiction/narrative	16,2	31,1	28,9	28,3	0	0
Non-fiction	45	29,7	13,2	34	0	100
TOTAL	**32,5**	**30,1**	**15,5**	**21,5**	**0**	**0,4**

Table 33.1: Raw microsemantic frequencies of deontic SAfE *must, should* and *HAVE to* over time

	Period	High	Median	Formulaic
must	0	64	10	28
	1	47	23	28
	2	40	36	30
	3	56	53	19
	TOTAL	**207**	**122**	**105**
should	0	10	29	1
	1	14	61	3
	2	14	56	3
	3	28	72	1
	TOTAL	**66**	**218**	**8**
HAVE to	0	9	14	8
	1	22	51	4
	2	29	42	9
	3	41	32	1
	TOTAL	**101**	**139**	**22**

Table 33.2: Normalised microsemantic frequencies of deontic SAfE *must, should* and *HAVE to* per 100 words (%) over time

	Period	High	Median	Formulaic
must	0	33,3	5,2	14,6
	1	25,1	12,3	15
	2	22,9	20,6	17,1
	3	32,2	30,5	10,9
	TOTAL	**28,4**	**16,8**	**14,4**
should	0	3,8	11,1	0,4
	1	4,9	21,5	1,1
	2	7,5	30	1,6
	3	18,2	46,8	0,6
	TOTAL	**7,5**	**24,6**	**0,9**
HAVE to	0	10,6	16,5	9,4
	1	17,4	40,5	3,2
	2	24,6	35,6	7,6
	3	32	25	0,8
	TOTAL	**22,1**	**30,4**	**4,8**

Table 34.1: Raw frequencies of subjective, objective and indeterminate sources with high-degree, median-degree and formulaic *must, should* and *HAVE to* in SAfE over time

must	**Period**	**High**				**Median**				**Formulaic**			
		Subjective	Objective	Indeterminate	**TOTAL**	Subjective	Objective	Indeterminate	**TOTAL**	Subjective	Objective	Indeterminate	**TOTAL**
	0	41	23	0	**64**	3	7	0	**10**	16	9	3	**28**
	1	18	29	0	**47**	14	9	0	**23**	19	7	2	**28**
	2	14	23	3	**40**	18	18	0	**36**	23	3	4	**30**
	3	17	37	2	**56**	19	29	5	**53**	10	5	4	**19**
should	**Period**	**High**				**Median**				**Formulaic**			
		Subjective	Objective	Indeterminate	**TOTAL**	Subjective	Objective	Indeterminate	**TOTAL**	Subjective	Objective	Indeterminate	**TOTAL**
	0	2	8	0	**10**	7	22	0	**29**	0	1	0	**1**
	1	4	10	0	**14**	22	39	0	**61**	0	3	0	**3**
	2	7	7	0	**14**	32	24	0	**56**	1	2	0	**3**
	3	4	23	1	**28**	38	32	0	**70**	0	1	0	**1**
HAVE to	**Period**	**High**				**Median**				**Formulaic**			
		Subjective	Objective	Indeterminate	**TOTAL**	Subjective	Objective	Indeterminate	**TOTAL**	Subjective	Objective	Indeterminate	**TOTAL**
	0	0	9	0	**9**	0	13	1	**14**	7	0	1	**8**
	1	2	19	1	**22**	5	45	1	**51**	4	0	0	**4**
	2	2	27	0	**29**	1	38	3	**42**	9	0	0	**9**
	3	11	30	0	**41**	4	27	1	**32**	1	0	0	**1**

Table 34.2: Normalised frequencies of subjective, objective and indeterminate sources with high-degree, median-degree and formulaic *must, should* and *HAVE to* per 100 words (%) in SAfE over time

must	**Period**	**High**			**Median**			**Formulaic**		
		Subjective	Objective	Indeterminate	Subjective	Objective	Indeterminate	Subjective	Objective	Indeterminate
	0	64	36	0	30	70	0	57,2	32,1	10,7
	1	38,3	61,7	0	60,9	39,1	0	67,9	25	7,1
	2	35	57,5	7,5	50	50	0	76,7	10	13,3
	3	30,4	66,1	3,5	35,9	54,7	9,4	52,6	26,3	21,1
should	**Period**	**High**			**Median**			**Formulaic**		
		Subjective	Objective	Indeterminate	Subjective	Objective	Indeterminate	Subjective	Objective	Indeterminate
	0	20	80	0	24,1	75,9	0	0	100	0
	1	28,6	71,4	0	36,1	63,9	0	0	100	0
	2	50	50	0	57,1	42,9	0	33,3	66,7	0
	3	14,3	82,1	3,6	54,3	45,7	0	0	100	0
HAVE to	**Period**	**High**			**Median**			**Formulaic**		
		Subjective	Objective	Indeterminate	Subjective	Objective	Indeterminate	Subjective	Objective	Indeterminate
	0	0	100	0	0	92,9	7,1	87,5	0	12,5
	1	9,1	86,4	4,5	9,8	88,2	2	100	0	0
	2	6,9	93,1	0	2,4	90,5	7,1	100	0	0
	3	26,8	73,2	0	12,5	84,4	3,1	100	0	0

Table 35.1: Raw frequencies of the constructions of quasi-subjunctive *should* in SAfE over time

Period	Mandative	Adversative	Purposive	Emotive	Conditional	TOTAL
0	57	2	0	24	32	**115**
1	45	5	0	25	20	**95**
2	24	0	2	14	18	**58**
3	18	0	0	0	12	**30**
TOTAL	**144**	**7**	**2**	**63**	**82**	**298**

Table 35.2: Normalised frequencies of the constructions of quasi-subjunctive *should* per 100 words (%) in SAfE over time

Period	Mandative	Adversative	Purposive	Emotive	Conditional
0	21,8	0,8	0	9,2	12,3
1	15,9	1,8	0	8,8	7,1
2	12,8	0	1,1	7,5	9,6
3	11,7	0	0	0	7,8
TOTAL	**48,3**	**2,3**	**0,7**	**21,1**	**27,5**

APPENDIX 6
RAW AND NORMALISED SYNCHRONIC FREQUENCIES FOR *MUST* IN SAfE AND *MOET/MOES* IN AFRIKAANS ACCORDING TO SUBJECT

Table 36: Raw numbers and percentages of person of subject with SAfE *must* and Afrikaans *moet/moes* in the spoken register

	1st person		2nd person		3rd person		TOTAL	
	Raw	*%*	*Raw*	*%*	*Raw*	*%*	*Raw*	*%*
SAfE *must*	171	32,3	104	19,6	255	48,1	530	100
Afrikaans *moet/moes*	109	28,7	143	37,6	128	33,7	380	100

Table 37: Raw and normalised macrosemantic frequencies for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register

	Deontic		Dynamic		Epistemic		TOTAL	
	Raw	*%*	*Raw*	*%*	*Raw*	*%*	*Raw*	*%*
SAfE *must*	87	83,7	5	4,8	12	11,5	104	100
Afrikaans *moet/moes*	125	87,4	16	11,2	2	1,4	143	100

Table 38: Raw and normalised microsemantic deontic frequencies for contemporary SAfE *must* and contemporary Afrikaans *moet/moes* with 2nd-person subjects in the spoken register

	High		Median		Formulaic		TOTAL	
	Raw	*%*	*Raw*	*%*	*Raw*	*%*	*Raw*	*%*
SAfE *must*	16	18,4	67	77	4	4,6	87	100
Afrikaans *moet/moes*	31	24,8	88	70,4	6	4,8	125	100

BIBLIOGRAPHY

Aarts, Bas. 2004. Modelling linguistic gradience. *Studies in Language*, 28(1): 1-49.

Aarts, Bas, Gerald Nelson and Sean Wallis. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam/Philadelphia: John Benjamins. 342 p.

Abraham, Werner. 1998. The aspectual source of the epistemic-root distinction of modal verbs in German. (*In* Bechert, Johannes and Winfried Boeder, *eds*. *Sprache in Raum und Zeit. 2. Beiträge zur empirischen Sprachwissenschaft.* Tübingen: Gunter Narr, 231-250.)

Abraham, Werner. 2001. How far does semantic bleaching go? About grammaticalization that does not terminate in functional categories. (*In* Faarlund, Jan Terie, *ed*. *Grammatical Relations in Change.* Amsterdam: John Benjamins, 15-64.)

Abraham, Werner. 2002. Modal verbs: epistemics in German and English. (*In* Barbiers, Lambertus C. J., Fritz H. Beukema, Wim Van Der Wurff, *eds*. *Modality and Its Interaction With the Verbal System*. Amsterdam: John Benjamins, 19-50.)

Abraham, Werner and Elisabeth Leiss. 2012. Introduction: theory of mind elements across languages. Traces of Bühler's legacy in modern linguistics. (*In* Abraham, Werner and Elisabeth Leiss, *eds*. *Modality and Theory of Mind Elements Across Languages*. Berlin, Mouton De Gruyter, 1-36.)

Aijmer, Karin. 1985. The semantic development of *will*. (*In* Fisiak, Jacek, *ed.* *Historical Semantics: Historical Word-Formation*. Berlin: de Gruyter, 11-22.)

Aijmer, Karin. 2004. The semantic path from modality to aspect: *Be able to* in a cross-linguistic perspective. (*In* Lindquist, H. and Mair, C., *eds.* *Corpus*

Approaches to Grammaticalization in English. Amsterdam: John Benjamins, 57-78).

Alexander, L. G. 1988. *Longman English Grammar.* London: Longman.

Allen, Cynthia L. 1995. *Case Marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: Clarendon.

Arnauld, Antoine and Claude Lancelot. 1660. *Grammaire générale et raisonnée.* Paris: Pierre Le Petit.

Atkinson, Quentin D. 2011. Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 332: 346-349.

Balasubramanian, Chandrika. 2009. *Register Variation in Indian English*. Amsterdam: John Benjamins. 284 p.

Banfield, Ann. 1982. *Unspeakable Sentences: Narration and Representation in the Language of fiction*. London: Routledge. 340 p.

Bao, Zhiming. 2010. *Must* in Singapore English. *Lingua*, 120(7): 1727-1737.

Barton, David and Nigel Hall. 2000. Introduction. (*In* Barton, David and Nigel Hall, *eds*. *Letter Writing As a Social Practice.* Amsterdam: John Benjamins, 1-14.)

Baugh, Albert, C. and Thomas Cable. 1978. *A History of the English Language.* Tenth Edition. London & New York: Routledge. 438 p.

Baxter, Gareth, Richard Blythe, William Croft and Alan McKane. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change*, 21: 257-96.

Beeton, D. R. and Helen Dorner. 1975. *A Dictionary of English Usage in Southern Africa.* Cape Town: Oxford University Press. 196 p.

Bekker, Ian. 2009. *The Vowels of South African English.* Potchefstroom: NWU. (Thesis – Ph.D).

Bekker, Ian. 2012. South African English as a *late* 19th-century extraterritorial variety. *English World-Wide*, 33(2): 127-146.

Bekker, Ian. 2013. The formation of South African English. *English Today*, 29(1): 3-9.

Besalke, Brandi, *ed.* 2012. *A Dictionary of the English Language: A Digital Edition of the 1755 Classic by Samuel Johnson*. Page View, Page 1342. *must*, verb imperfect. http://johnsonsdictionaryonline.com/?page_id=7070&i=1342 Date of access: 19 April 2014.

Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, Douglas. 1992. The multi-dimensional approach to linguistic analyses of genre variation: an overview of methodology and findings. *Computers and the Humanities*, 26(5/6): 331-345.

Biber, Douglas. 1995. *Dimensions of Register Variation: A Cross-Linguistic Comparison.* Cambridge: Cambridge University Press. 428 p.

Biber, Douglas. 2003. Compressed noun-phrase sturctures in newspaper discourse: the competing demands of popularization vs. economy. (*In* Aitchinson, Jean and Diana M. Lewis, *eds*. *New Media Langauge*. London: Routledge, 169-181.)

Biber, Douglas, Susan Conrad and Randi Reppen. 1998. *Corpus linguistics: Investigating Language Structure and Use.* Cambridge: Cambridge University Press.

Biber, Douglas and Edward Finegan. 1989. Drift and the evolution of English style: a history of three genres. *Language*, 65(3): 487-517.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finnegan. 1999. *The Longman Grammar of Spoken and Written English*. Harlow: Pearson.

Bible. 1911. *The Holy Bible: a Facsimile in a reduced size of the Authorised Version published in the year 1611, with an introduction by A. W. Pollard, and illustrative documents*. Oxford: Oxford University Press.

Blythe, Richard and William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language*, 88 (2): 269-304.

Bonnot de Condillac, Étienne. 1793. *Essai sur l'origine des connoissances humaines.* Paris: Chez les Libraires Associés.

Bopp, Franz. 1833. *Vergleichende Grammatik des Sanskrit, Send, Armenischen, Griechischen, Lateinischen, Litauischen, Altslavischen, Gothischen und Deutschen: Enthaltend die Laut-Lehre, Wurzel-Vergleichung und Casus-Bildung.* Vol. 1. Berlin: Dümmler. 288 p.

Bopp, Franz and Karl Josef H. Windischmann. 1816. *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache: Ramajan und Muhabharat und eigenen Abschnitten aus den Vedas.* Frankfurt am Main: Andreae. 312 p.

Botha, Yolande. 2013. Corpus evidence of antideletion in Black South African English noun phrases. *English Today,* 29(1): 16-21.

Bowerman, Sean. 2004a. White South African English: phonology. (*In* Schneider, Edgar W., Kate Burrage, Bernd Kortmann, Rajend Mesthrie and Clive Upton, *eds. A Handbook of Varieties of English: Phonology.* Berlin: Mouton de Gruyter, 931-942.)

Bowerman, Sean. 2004b. White South African English: Morphology and Syntax. (*In* Kortmann, Bernd, Edgar W. Schneider, Kate Burrage, Rajend Mesthrie and Clive Upton, *eds. A Handbook of Varieties of English: Morphology and Syntax*. Berlin: Mouton de Gruyter, 472-478.)

Boye, Kasper and Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language*, 88(1): 1-44.

Bragg, Melvyn. 2003. *The Adventure of English: The Biography of a Language*. London: Hodder & Stoughton. 354 p.

Branford, Jean and William Branford. 1991. *A Dictionary of South African English*. Oxford: Oxford University Press. 412 p.

Branford, William. 1996. English in South African society: a preliminary overview. (*In* de Klerk, Vivian, *ed. Focus on South Africa*. Amsterdam: Benjamins, 35-51.)

Brinton, Laurel J. 1991. The origin and development of quasimodal *have to* in English. Paper presented at the 10th International Conference on Historical Linguistics, Amsterdam, ms. 59 p.

Brown, Penelope and Stephen Levinson. 1987. *Politeness: Some Universals in Language Usage.* Cambridge: Cambridge University Press.

Butler, Guy, *ed.* 1974. *The 1820 Settlers: An Illustrated Commentary*. Cape Town: Human & Rousseau.

Bybee, Joan. 1995. The semantic development of past tense modals in English. (*In* Bybee, Joan and Suzanne Fleischman, *eds*. *Modality in Grammar and Discourse*. Amsterdam: John Benjamins, 503-518.)

Bybee, Joan. 2006. From usage to grammar: the mind's response to repetition. *Language*, 82(4): 711-733.

Bybee, Joan and Suzanne Fleischman. 1995. Modality in grammar and discourse: an introductory essay. (*In* Bybee, Joan and Suzanne Fleischman, *eds*. *Modality in Grammar and Discourse*. Amsterdam: John Benjamins, 1-14.)

Bybee, Joan, William Pagluica and Revere Perkins. 1991. Back to the future. (*In* Traugott, Elizabeth and Bernd Heine, *eds*. *Approaches to Grammaticalization*. *Vol. 2.* Amsterdam: John Benjamins, 17-58.)

Bybee, Joan, William Pagluica and Revere Perkins. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World.* Chicago: University of Chicago Press. 420 p.

Carroll, Lewis. 1897. *Through the Looking Glass: And what Alice Found There.* Philadelphia: Henry Altemus.

Charnock, Ross. 2009. When *may* means *must*: deontic modality in English statute construction. (*In* Salkie, Raphael, Pierre Busuttil and Johan van der Auwera, *eds*. *Modality in English: Theory and Description.* Berlin and New York: Mouton de Gruyter, 177-198.)

Cheshire, Jenny, *ed*. 1991. *English around the World: sociolinguistic perspectives*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1957. *Syntactic Structures*. Gravenhage: Mouton & Co.'s.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, Massachusetts: MIT Press. 251 p.

Chomsky, Noam. 1966. *Cartesian Linguistics: a Chapter in the History of*

Rationalist Thought. New York: Harper & Row.

Clackson, James. 2007. *Indo-European Linguistics: An Introduction*. Cambridge: Cambridge University Press. 260 p.

Close, Joanne and Bas Aarts. 2010. Current change in the modal system of English: A case study of *must*, *have to* and *have got to*. (*In* Lenker, Ursula, Judith Huber and Robert Mailhammer, *eds*. *The History of English Verbal and Nominal Constructions*. Vol. I of *English Historical Linguistics 2008: Selected Papers from the Fifteenth International Conference on English Historical Linguistics* (*ICEHL 15), Munich 24-30 August 2008*. Amsterdam: John Benjamins, 165-181.)

Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries*. Beckenham: Croom Helm.

Cochrane, James R. 1987. *Servants of Power: the Role of the English-speaking Churches in South Africa, 1903-1930*. Braamfontein: Ravan Press.

Coetzee-Van Rooy, Susan. 2013. Afrikaans in contact with English: endangered language or a case of exceptional bilingualism? *International Journal for the Sociology of Language*, 224: 179-207.

Coetzee-Van Rooy, Susan and Bertus Van Rooy. 2005. South African English: labels, comprehensibility and status. *World Englishes*, 24(1): 1-19.

Cohen, David, *ed.* 1974. *Explaining Linguistic Phenomena*. New York: Wiley.

Collins, Michael and Salikoko S. Mufwene. 2005. What we mean when we say 'creole': an interview with Salikoko S. Mufwene. *Callaloo*, 28(2): 425-462.

Collins, Peter. 2005. The modals and quasi-modals of obligation and necessity in Australian English and other Englishes. *English World-Wide*, 26(3): 249-273.

Collins, Peter. 2009a. *Modals and Quasi-modals in English*. Amsterdam: Rodopi.

Collins, Peter. 2009b. Modals and quasi-modals in World Englishes. *World Englishes,* 28(3): 281-292.

Conboy, Martin. 2010. *The Language of Newspapers: Socio-Historical Perspectives.* London: Continuum. 176 p.

Concise Oxford English Dictionary. 2006. 11th ed. (revised). Oxford: Oxford University Press. 1708 p.

Conrad, Susan and Douglas Biber. 2001. *Variation in English: Multi-Dimensional Studies*. London: Longman. 255 p.

Conradie, C. Jac. 1976. Die modale werkwoord in Nederlands en Afrikaans. *Taalfasette*, 20(4): 62-68.

Conradie, C. Jac. 1980. Aspects of semantic change in Dutch modal auxiliaries. (*In* Rupp, H. and H. G. Roloff, *eds*. *Akten des VI. Internatationalen Germanisten-Kongresses Basel*. Bern: Peter Lang, 488-493.)

Conradie, C. Jac. 1987. Semantic change in modal auxiliaries as a result of speech act embedding. (*In* Harris, Martin and Paolo Ramat, *eds. Historical Development of Auxiliaries.* Berlin: De Gruyter, 171-180.)

Cort, Alison and David Denison. 2005. The category modal – a moving target? Paper presented at the First International Conference on the Linguistics of Contemporary English, University of Edinburgh.

Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach.* Harlow: Person Education.

Croft, William. 2002. The Darwinization of Linguistics. *Selection*, 3(1): 75-91.

Croft, William. 2003. *Typology and Universals*. Cambridge: Cambridge University Press. 341 p.

Croft, William. 2009. Evolutionary models and functional-typological theories of language change. (*In* Van Kemenade, Ans and Los Bettelou, *eds*. *The Handbook of the History of English.* Chichester: Wiley & Sons, 68-94.)

Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. Sixth Edition. Oxford: Blackwell. 529 p.

De Haan, Ferdinand. 2006. Typological approaches to modality. (*In* Frawley, William J., Erin Eschenroeder, Sarah Mills and Thao Nguyen, *eds*. *The Expression of Modality*. Berlin: de Gruyter, 27-70.)

De Haan, Ferdinand and Björn Hansen. 2009. Introduction. (*In* Hansen, Björn and Ferdinand de Haan, *eds*. *Modals in the Languages of Europe.* Berlin/New York: Mouton de Gruyter, 1-8.)

De Klerk, Vivian. 2006. *Corpus Linguistics and World Englishes: An Analysis of Xhosa English*. London: Continuum.

Denison, David. 1993. *English Historical Syntax*. Oxford: Oxford University Press.

Denison, David. 2001. Gradience and linguistic change. (*In* Brinton, J., *ed. Historical Linguistics 1999. Selected Papers from the 14th International Conference on Historical Linguistics.* Vancouver, 9-13 August 1999. Amsterdam: John Benjamins, 119-144.)

Denison, Norman. 1998. Language change in progress: variation as it happens. (*In* Coulmas, Florian, *ed*. *The Handbook of Sociolinguistics*. Oxford: Wiley-Blackwell, 65-80.)

Depraetere, Ilse and Susan Reed. 2006. Mood and modality in English. (*In* Aarts, Bas and April McMahon, *eds*. *The Handbook of English Linguistics.* London: Blackwell, 269-290.)

Depraetere, Ilse and An Verhulst. 2008. Source of modality: a reassessment. *English Language and Linguistics*, 12(1): 1-25.

Deumert, Ana. 2005. The unbearable lightness of being bilingual: English-Afrikaans language contact in South Africa. *Language Sciences*, 27: 113–135.

Deutscher, Guy. 2005. *The Unfolding of Language: An Evolutionary Tour of Mankind's Greatest Invention.* New York: Metropolitan.

De Wet, Febe, Alta de Waal and Gerhard B. van Huyssteen. 2011. *Developing a broadband automatic speech recognition system for Afrikaans.* Paper delivered at the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), Florence, Italy, 27-31 August 2011.

Ehrman, Madeline. 1966. *The Meanings of the Modals in Present-Day American English*. The Hague: Mouton & Co.

Enfield, N. J. 2003. *Linguistic Epidemiology: Semantics and Grammar of Language*

Contact in Mainland Southeast Asia. London: Routledge.

Facchinetti, Roberta, Manfred Krug and Frank Palmer. 2003. Preface. (*In* Facchinetti, Roberta; Manfred Krug and Frank Palmer, *eds*. *Modality in Contemporary English*. Berlin: Mouton de Gruyter, v-xiii.)

Fairclough, Norman. 1992. *Discourse and social change.* Cambridge: Polity Press.

Fleischman, Suzanne. 1982. *The future in thought and language: Diachronic evidence from Romance.* Cambridge: Cambridge University Press.

Fontaine, André and William A. Glavin. 1987. *The Art of Writing Nonfiction.* Second Edition. New York: Syracuse University Press. 200 p.

Funk, Isaac K, *ed*. 1898. *Standard Dictionary of the English Language*. *Vol. II: M-Z.* New York and London: Funk & Wagnalls Company. 1168 p.

Garson, James. 2009. Modal Logic. (*In* Zalta, Edward N., *ed*. *The Stanford Encyclopedia of Philosophy.* Winter 2009 Ed. http://plato.stanford.edu/archives/win2009/entries/logic-modal. Date of access: 22 Sept. 2012.)

Geeraerts, D. 1985. Cognitive restrictions on the structure of semantic change. (*In* Fisiak, J., *ed.* *Historical Semantics*. Berlin: Mouton de Gruyter, 127-53.)

Geisler, Christer. 2002. Investigating register variation in nineteenth-century English: a multi-dimensional comparison. (*In* Reppen, Randi, Susan M. Fitzmaurice and Douglas Biber, *eds*. *Using Corpora to Explore Linguistic Variation.* Amsterdam: John Benjamins, 249-272.)

Giliomee, Hermann and Mbenga, Bernard. 2007. *Nuwe geskiedenis van Suid-Afrika*. Cape Town: Tafelberg.

Givón, Talmy. 1971. Historical syntax and synchronic morphology. *CLS*, 7: 394-415. Chicago: Chicago Linguistic Society.

Givón, Talmy. 1973. The time-axis phenomenon. *Language*, 49(4): 890-925.

Givón, Talmy. 1984. Prolegomena to discourse-pragmatics. *Journal of Pragmatics,* 8(4): 489-516.

Goffman, Erving. 1967. *Interaction ritual: essays on face to face behavior*. New York: Garden City.

Goossens, Louis. 1987. The auxiliarization of the English modals: a functional grammar view. (*In* Harris, Martin and Paolo Ramat, *eds. Historical Development of Auxiliaries.* Berlin: De Gruyter, 111-144.)

Gordimer, Nadine. 1966. *The Late Bourgeois World*. New York: Viking Press.

Greenberg, Joseph H. 1979. Rethinking linguistics diachronically. *Language*, 55(2): 275-290.

Gunther, Irene and Leslie T. Sharpe. 1994. *Editing Fact and Fiction: A Concise Guide to Book Editing.* Cambridge: Cambridge University Press. 277 p.

Gupta, Anthea F. 1997. Colonisation, migration, and functions of English. (*In* Schneider, Edgar W., *ed*. *Englishes Around the World.* Vol. 1. Amsterdam, John Benjamins, 47-58.)

Hall, Nigel. 2000. The Materiality of Letter Writing. A nineteenth century perspective. (*In* Barton, David and Nigel Hall, *eds*. *Letter Writing As a Social Practice.* Amsterdam: John Benjamins, 83-130).

Halliday, M. A. K. 1970. Functional diversity in language as seen from a consideration of modality and mood in English. *Foundations of Language*, 6(3): 322-361.

Halliday, M. A. K. 1994. *Introduction to Functional Grammar*. Second Edition. London: Arnold.

Halliday, M. A. K., and Matthiessen, Christian. 2004. *An Introduction to Functional Grammar.* Third Edition. London: Arnold. 689 p.

Harris, Martin. 1986. English *ought to*. (*In* Kastovsky, Dieter and Aleksander Szwedek, *eds*. *Linguistics across historical and geographical boundaries*. *Vol 1: Linguistic Theory and Historical Linguistics*. Berlin, Mouton, 347-358.)

Harris, Martin. 1987. Syntactic and semantic change within the modal systems of English and Afrikaans. (*In* Harris, Martin and Paolo Ramat, *eds. Historical Development of Auxiliaries*. Berlin: De Gruyter, 181-189.)

Hauptfleisch, T. 1979. *Language Loyalty in South Africa. Volume 3: Motivations to language use: Opinions and attitudes of white adults in urban areas* (Report no. TLK/L-10). Pretoria: Human Sciences Research Council.

Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford: Oxford University Press.

Heine, Bernd. 2003. Grammaticalization. (*In* Joseph, Brian and Janda, Richard, *eds*. *The Handbook of Historical Linguistics.* Oxford: Wiley-Blackwell, 575-677.)

Hickey, Raymond. 2007. *Irish English: History and Present-Day Forms*. Cambridge: Cambridge University Press. 504 p.

Hock, Hans Heinrich. 1991. *Principles of Historical Linguistics*. Second Edition. Berlin: Mouton de Gruyter. 745 p.

Hockly, Harold Edward. 1966. *The Story of the British Settlers of 1820 in South Africa*: *A Brief History for Use in Schools and Other Educational Institutions.* Cape Town & Johannesburg: Juta & Company.

Hodson, T. C. 1939. Socio-linguistics in India. *Man in India,* 19: 94-8.

Hofmeyr, George, *ed*. 2002. *NG Kerk 350: Eenhonderd bakens in die geskiedenis van die Nederduitse Gereformeerde Kerk 1652-2002*. Wellington: Lux Verbi. 252 p.

Hooper, A. G. 1951. English in South Africa. (*In* Clark, John W. and Eric Partridge, *eds*. *British and American English Since 1900.* London: Andrew Dakers Limited, 80-84.)

Hopper, Paul. 1996. Some Recent Trends in Grammaticalization. *Annual Review of Anthropology*, 25: 217-236.

Hopper, Paul. 1998. Emergent grammar. (*In* Michael Tomasello, *ed*. *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*. Mahwah, NJ: Erlbaum, 155-75.)

Hopper, Paul and Traugott, Elizabeth. 2003. *Grammaticalization*. Second Edition. Cambridge: Cambridge University Press.

Horne Tooke, John. 1857. *The diversions of Purley.* London: W. Tegg & co. 739 p.

Huddleston, Rodney. 2002. The verb. (*In* Rodney Huddleston and Geoffrey K. Pullum, *eds*. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press, 71-212.)

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Hudson, Richard A. 1996. *Sociolinguistics.* Cambridge: Cambridge University Press. 179 p.

Hughes, Arthur and Peter Trudgill. 1979. *English Accents and Dialects: An Introduction to Social and Regional Varieties of British English*. Michigan: University Park Press. 90 p.

Hughes, Geoffrey. 2000. *A History of English Words.* Malden: Wiley-Blackwell. 430 p.

Hull, David. 2001. *Science and Selection: Essays on Biological Evolution and the Philosophy of Science.* Cambridge: Cambridge University Press.

Hundt, Marianne. 1998. *New Zealand English Grammar: Fact or Fiction?* Amsterdam: Benjamins.

Jacobson, S. 1980. Issues in the study of syntactic variation. (*In* Jacobson, S., *ed*. *Papers from the Scandinavian Symposium on Syntactic Variation, Stockholm, May 18-19, 1979.* Stockholm: Almqvist and Wicksell, 23-26).

Jeffery, Chris. 2003. On compiling a corpus of South African English. *Southern African Linguistics and Applied Language Studies*, 21(4): 341–344.

Jeffery, Chris and Bertus van Rooy. 2004. Emphasiser *now* in colloquial South African English. *World Englishes*, 23(2): 269-280.

Jespersen, Otto. H. 1924. *The philosophy of grammar.* London: Allen & Unwin.

Jespersen, Otto. H. 1961. *A Modern English Grammar on Historical Principles: Part II: Syntax.* VII. London: Allen & Unwin.

Johnson-Laird, P. N. 1978. The meaning of modality. *Cognitive Science*, 2: 17-26.

Joos, M. 1964. *The English Verb: Forms and Meanings*. Madison, WIS: The University of Wisconsin Press.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: the English language in the outer circle. (*In* Quirk, R. and Widdowson, H. G., *eds*. *English in the World*. Cambridge: Cambridge University Press.)

Kachru, Braj B. 1991. Liberation linguistics and the Quirk Concern. *English Today*, 7: 3-13.

Kachru, Braj B. 1992. *The Other Tongue. English Across Cultures.* Urbana: University of Illinois Press.

Kachru, Braj B. 1996. World Englishes: agony and ecstasy. *Journal of Aesthetic Education,* 30(2):135-155. (Special Issue: Distinguished Humanities Lectures II.)

Katamba, Francis. 2005. *English Words: Structure, History, Usage*. London: Routledge. 322 p.

Kawaguchi, Yuji, Wolfgang Viereck and Makoto Minegishi. 2011. Introduction. (*In* Kawaguchi, Yuji, Makoto Minegishi and Wolfgang Viereck, *eds*. *Corpus-based Analysis and Diachronic Linguistics.* Amsterdam/Philadelphia: John Benjamins, 7-20.)

Kerswill, Paul. 2010. Contact and new varieties. (*In* Hickey, Raymond, *ed. The Handbook of Language Contact*. Oxford: Wiley-Blackwell, 230- 251.)

Kielkiewicz-Janowiak, Agnieszka. 1992. *A Socio-historical Study in Address: Polish and English*. Frankfurt am Main: Peter Lang.

Kroch, Anthony S. 1978. Toward a theory of social dialect variation. *Language in Society*, 7: 17-36.

Krug, Manfred G. 2000. *Emerging English Modals: A Corpus-Based Study of Grammaticalization.* The Hague: Mouton de Gruyter.

Labov, William. 1963. The social motivation of a sound change. *Word*, 19: 273-309.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics, 2006. Second Edition. Cambridge: Cambridge University Press.

Labov, William. 1972. Some Principles of Linguistic Methodology. *Language in Society*, 1:97-120.

Labov, William. 1973. The boundaries of words and their meanings. (*In* C.-J. Bailey and R. Shuy, *eds*. *New Ways of Analyzing Variation in English*. Washington,

DC: Georgetown University Press, 340-373. Reprinted in Aarts. B. *et al*., *eds*. *Fuzzy Grammar.* Oxford: Oxford University Press, 67-90.)

Labov, William. 1994. *Principles of linguistic change: internal factors.* Vol. 1 of *Principles of Linguistic Change*. Oxford: Wiley-Blackwell. 672 p.

Labov, William. 2001. *Principles of linguistic change: social factors*. Vol. 2 of *Principles of Linguistic Change*. Oxford: Wiley-Blackwell. 592 p.

Labov, William. 2007. Transmission and diffusion. *Language,* 83(2): 344-387.

Labov, William. 2010. *Principles of Linguistic Change, Cognitive and Cultural Factors*. Vol. 3 of *Principles of Linguistic Change*. Oxford: Wiley-Blackwell. 424 p.

Lakoff, George. 1972. Hedges: a study in meaning criteria and the logic of fuzzy concepts. *CLS*, 8: 183-228.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind.* Chicago: University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Vol. 1: Theoretical Prerequisites*. Stanford, CT: Stanford University Press.

Langacker, Ronald W. 2008. *Cognitive Grammar*. Oxford: Oxford University Press.

Lanham, Len W. 1967. *The Pronunciation of South African English: A Phonetic-Phonemic Introduction.* A. A. Balkema, Cape Town.

Lanham, Len W. 1978. South African English. (*In* Lanham, Len W. and K. P. Prinsloo, *eds. Language and Communication Studies in South Africa*. Cape Town: Oxford University Press.)

Lanham, Len W. 1982. English in South Africa. (*In* Bailey, R. W. and Manfred Görlach, *eds*. *English as a World Language*. Ann Arbor: University of Michigan Press, 324-52.)

Lanham, Len W. 1996. A history of English in South Africa. (*In* Vivian de Klerk, *ed. Focus on South Africa*. Amsterdam: Benjamins, 19-34.)

Lanham, Len W. and MacDonald, Carol A. 1979. *The Standard in South African History and its Social History*. Heidelberg: Groos.

Larreya, Paul. 2003. Irrealis, past time reference and modality. (*In* Facchinetti, Roberta; Manfred Krug and Frank Palmer, *eds*. *Modality in Contemporary English*. Berlin: Mouton de Gruyter, 21-46.)

Larreya, Paul. 2009. Towards a typology of modality in language. (*In* Salkie, Raphael, Pierre Busuttil and Johan van der Auwera, *eds. Modality in English: Theory and Description.* Berlin and New York: Mouton de Gruyter, 9-30.)

Larreya, Paul and Claude Rivière. 2005. *Grammaire explicative de l'anglais*. 3rd ed. France: Pearson Education.

Lass, Roger. 1987. *The Shape of English: Structure and History.* London & Melbourne: J. M. Dent & Sons Limited. 384 p.

Lass, Roger. 2002. South African English. (*In* Rajend Mesthrie, *ed. Language in South Africa.* Cambridge: Cambridge University Press, 104-126.)

Lass, Roger and Wright, Susan. 1986. Endogeny vs. contact: 'Afrikaans influence' on South African English. *English World-Wide*, 7(2): 201–223.

Leech, Geoffrey. 1991. The state of the art in corpus linguistics. (*In* Altenberb, B. and K. Aijmer, *eds*. *English Corpus Linguistics: Studies in Honour of Jan Svartvik.* London/New York: Longman, 8-29.)

Leech, Geoffrey. 2003. Modality on the move: the English modal auxiliaries 1961-1992. (*In* Roberta Facchinetti, Manfred Krug and Frank Palmer, *eds*. *Modality in Contemporary English.* Berlin: Mouton de Gruyter, 223-240.)

Leech, Geoffrey. 2011. The modals ARE declining: Reply to Neil Millar's "Modal verbs in TIME: Frequency changes 1923–2006". *International Journal of Corpus Linguistics*, 16(4): 547-564.

Leech, Geoffrey and Jennifer Coates. 1980. Semantic indeterminancy and the modals. (*In* Greenbaum, S., Leech, G. and Svartvik, J., *eds*. *Studies in English Linguistics for Randolph Quirk.* London: Longman, 79-90.)

Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press. 341 p.

Leech, Geoffrey and Nicolas Smith. 2006. Recent grammatical change in written English 1961-1992: some preliminary findings of a comparison of American with British English. (*In* Antoinette Renouf and Andrew Kehoe, *eds. The changing face of corpus linguistics.* New York: Rodopi, 173-200.)

Leech, Geoffrey and Nicolas Smith. 2009. Change and constancy in linguistic change: How grammatical usage in written English evolved in the period 1931-1991. (*In* Antoinette Renouf and Andrew Kehoe, *eds*. *Corpus Linguistics: Refinements and reassessments.* New York: Rodopi, 185-204.)

Lehmann, Winfred, P. 1973. *Historical Linguistics: An Introduction.* Second Edition. New York: Holt, Rinehart and Winston, Inc. 273 p.

Le Page, Robert B. 1998. The evolution of a sociolinguistic theory of language. (*In* Coulmas, Florian, *ed*. *The Handbook of Sociolinguistics*. Oxford: Wiley-Blackwell, 15-32.)

Levey, Stephen and Shana Poplack. 2010. Contact-induced grammatical change: A cautionary tale. (*In* Auer, Peter and Jürgen Erich Schmidt, *eds*. *Language and Space: An International Handbook of Linguistic Variation. Theories and Methods*. Berlin: Mouton De Gruyter, 391-418.)

Lichtenberk, F. 1991. Semantic Change and Heterosemy in Grammaticalization. *Language*, 67(3): 475-509.

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge: MIT Press.

Lindquist, Hans and Mair, Christian. 2004. Introduction. (*In* Lindquist, H. and Mair, C., *eds. Corpus Approaches to Grammaticalization in English*. Amsterdam: John Benjamins, ix-xiv).

Lyons, John. 1977. *Semantics*. Vol. 2. Cambridge: Cambridge University Press.

Mair, Christian. 1995. Changing patterns of complementation, and concomitant grammaticalization, of the verb *help* in present-day British English. (*In* Aarts, B. and Meyer, C. F., *eds. The Verb in Contemporary English: Theory and Description.* Cambridge: Cambridge University Press, 258-272.)

Mair, Christian. 1997. Parallel corpora: a real-time approach to the study of language change in progress. (*In* Ljung, Magnus, *ed.* *Corpus-based studies in English.* Amsterdam: Rudopi, 195-209.)

Mair, Christian. 2004. Corpus Linguistics and grammaticalization theory: Statistics, frequencies and beyond. (*In* Lindquist H. and Mair, C., *eds.* *Corpus Approaches to Grammaticalization in English*. Amsterdam: John Benjamins, 121-150.)

Mair, Christian. 2006. *Twentieth-Century English: History, Variation, and Standardization.* Cambridge: Cambridge University Press. 244 p.

Mair, Christian. 2014. Fluctuating statistics, volatile genres, and structural change: modality in the extended Brown family of corpora. Paper delivered at ICAME 35 (*Corpus Linguistics, Context and Culture)*, Nottingham, England, 30 April - 4 May 2014.

Mair, Christian and Geoffrey Leech. 2006. Current changes in English syntax. (*In* Aarts, Bas and April McMahon, *eds*. *The Handbook of English Linguistics.* Oxford: Blackwell, 318-342.)

Marckwardt, Albert. 1958. *American English*. New York: Oxford University Press.

Martinet, André. 1953. Preface to U. Weinreich. *Languages in Contact*. The Hague: Mouton.

Matthews, Peter Hugoe. 2007. *The Concise Oxford Dictionary of Linguistics.* Second Edition. Oxford: Oxford University Press. 443 p.

McCormick, Kay. 2002. *Language in Cape Town's District Six.* Oxford: Oxford University Press.

McCallum-Bayliss, H. 1985. *The Modal Verbs: Universal/Univocal lexical items*. Ph.D. thesis, Georgetown University. Ann Arbor: University Microfilms International.

McWhorter, John H. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars.* Oxford: Oxford University Press.

Mesthrie, Rajend. 1992. *English in Language Shift.* Johannesburg: University of the Witwatersrand Press.

Mesthrie, Rajend. 1993. English in South Africa. *English Today*, 9(1): 27-33.

Mesthrie, Rajend. 1996. Imugint excusations: Missionary English in the nineteenth century Cape Colony, South Africa. *World Englishes*, 15(2): 139-157.

Mesthrie, Rajend. 2002. Endogeny versus contact revisited: aspectual *busy* in South African English. *Language Sciences*, 24: 345-358 .

Mesthrie, Rajend. 2003. The world Englishes paradigm and contact linguistics: refurbishing the foundations. *World Englishes*, 22(4): 449-461.

Mesthrie, Rajend. 2009. Deracialising the GOOSE vowel in South African English: Accelerated linguistic change amongst young, middle-class, females in post-apartheid South Africa. (*In* Thomas Hoffman and Lucia Siebers, *eds*. *World Englishes: Problems – Properties – Prospects.* Amsterdam: Benjamins: 3-18.)

Mesthrie, Rajend and Paula West. 1995. Towards a grammar of Proto-South African English. *English World-Wide*, 16: 105-133.

Millar, Neil. 2009. Modal verbs in TIME: frequency changes 1923-2006. *International Journal of Corpus Linguistics*, 14(2): 191-220.

Milroy, James. 1992. *Linguistic variation and change: On the historical sociolinguistics of English*. Oxford: Blackwell.

Milroy, James and Lesley Milroy. 1998. Varieties and variation. (*In* Coulmas, Florian, *ed*. *The Handbook of Sociolinguistics*. Oxford: Wiley-Blackwell, 47-64.)

Mitchell, A. G. and A. Delbridge. 1965. *The Pronunciation of English in Australia*. Sydney: Angus & Robertson.

Mortelmans, Tanja, Kasper Boye and Johan van der Auwera. 2009. Modals in the Germanic languages. (*In* Hansen, Björn and Ferdinand de Haan, *eds*. *Modals in the Languages of Europe.* Berlin/New York: Mouton de Gruyter, 11-70.)

Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.

Mufwene, Salikoko S. 2008. *Language Evolution: Contact, Competition and Change.* London: Continuum. 354 p.

Myhill, John. 1995. Change and continuity in the function of American English modals. *Linguistics*, 33: 157-211.

Nelson, Gerald. 2002. *International Corpus of English: Markup Manual for Written Texts.* http://ice-corpora.net/ice/manuals.htm. Date of access: 10 Feb. 2013.

Nevalainen, Terttu. 1996. Introduction. (*In* Nevalainen, Terttu and Helena Raumolin-Brunberg, *eds*. *Sociolinguistics and language history: studies based on the Corpus of early English correspondence*. Amsterdam: Rudopi.)

Nevalainen, Terttu. 2009. Historical Sociolinguistics and Language Change. (*In* van Kemenade, Ans and Bettelou Los, *eds.* *Handbook of the History of English*. Oxford: Wiley-Blackwell, 558-588.)

Nevalainen, Terttu and Helena Raumolin-Brunberg. 2003. *Historical sociolinguistics: language change in Tudor and Stuart England.* London: Longman. 266p.

Nevalainen, Tertuu and Helena Raumolin-Brunberg. 2012. Historical sociolinguistics: origins, motivations, and paradigms. (*In* Hernández-Campoy, Juan M. and Conde-Silvestre, J. Camilo, *eds*. *The Handbook of Historical Sociolinguistics*. New-York: Wiley-Blackwell.)

Newmeyer, Frederick. 1986. *Linguistic Theory in America*. Second Edition. New York: Academic Press.

Nida, Eugene. 1949. *Morphology*. Second Edition. Ann Arbor: University of Michigan Press.

Nuyts, Jan. 2006. Modality: overview and linguistic issues. (*In* William Frawley, Erin Eschenroeder, Sarah Mills, Thao Nguyen, *eds*. *The Expression of Modality*. Berlin: de Gruyter, 1-27.)

Onions, C. T, *ed.* 1966. *The Oxford Dictionary of English Etymology*. Oxford: Clarendon Press. 1024 p.

Online Etymology Dictionary. 2013. *must* (v). http://www.etymonline.com/index.php?allowed_in_frame=0&search=must&searchmode=none Date of access: 9 May 2013.

Oxford Dictionary of Current English. 1919. *must,* v. https://archive.org/details/con00ciseoxforddicfowlrich Date of access: 19 April 2014.

Oxford English Dictionary (OED) Online. 2012a. *mote*, v.1. http://www.oed.com/view/Entry/124229?rskey= hXmANU&result=7 Date of access: 5 June 2012.

Oxford English Dictionary (OED) Online. 2012b. *must*, v.1. http://www.oed.com/view/Entry/124229?rskey= hXmANU&result=7 Date of access: 5 June 2012.

Oxford English Dictionary (OED) Online. 2013. *need*, v.1. http://www.oed.com.nwulib.nwu.ac.za/view/Entry/125754?rskey=uMhbXN&result=4#eid Date of access: 9 May 2013.

Pahta, Päivi, Minna Nevala, Arja Nurmi and Minna Palander-Collin. 2010. Language practices in the construction of social roles in Late Modern English. (*In* Pahta, Päivi, Minna Nevala, Arja Nurmi and Minna Palander-Collin, *eds*. *Social Roles and Language Practices in Late Modern English.* Amsterdam, John Benjamins, 1-27.)

Palmer, Frank. R. 1990. *Modality and the English modals*. Second Edition. London: Longman.

Palmer, Frank. R. 2001. *Mood and Modality*. Second Edition (Original Edition, 1986). Cambridge: Cambridge University Press.

Paul, Hermann. 1909. *Prinzipien der Sprachgeschichte.* Fourth Edition. Halle: Niemeyer.

Percival, W. Keith. 1976. The Notion of Usage in Vaugelas and in the Port Royal Grammar. (*In* Parret, Herman, *ed. History of Linguistic Thought and Contemporary Linguistics.* Berlin: de Gruyter, 374-382.)

Perkins, M. 1983. *Modal Expressions in English.* London: Frances Pinter.

Pharos Afrikaans-Engels & English-Afrikaans Woordeboek/Dictionary. 2005. Cape Town: Pharos Dictionaries. 1552 p.

Plank, Frans. 1984. The modals story retold. *Studies in Language*, 8: 305-364.

Ponelis, Fritz A. 1979. *Afrikaanse Sintaksis*. Pretoria: J. L. van Schaik.

Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.

Putnam, Hilary. 2001. *Enlightenment and pragmatism*. Amsterdam: Koninklijke van Gorcum.

Pyles, Thomas. 1965. The role of historical linguistics. *College English*, 26(4): 292-298.

Quirk, Randolph. 1990. Language varieties and standard language. *English Today*, 6: 3-10.

Quirk, Randolph, Sydney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

Radford, Andrew. 2003. *Syntax: A Minimalist Introduction*. Cambridge: Cambridge University Press. 283 p.

Raidt, Edith H. 1991. *Afrikaans en Sy Europese Verlede*. Cape Town: Nasou Beperk. 271 p.

Rauch. I. 1990. Evidence of language change. (*In* Polomé, Edgar, *ed*. *Research Guide on Language Change. Issue 48 of Trends in Linguistics: Studies and monographs*. Berlin: de Gruyter, 36-70.)

Rayson, Paul, Damon Berridge and Brian Francis. 2004. Extending the Cochran rule for the comparison of word frequencies between corpora. (*In Purnelle G., Fairon C., Dister A., eds. Vol. II: Le poids des mots: Proceedings of the 7th International Conference on Statistical analysis of textual data (JADT 2004.* 10-12 March 2004. Louvain-la-Neuve, Belgium: Presses universitaires de Louvain, 926-936).

Rayson, Paul and Roger Garside. 2000. Comparing corpora using frequency profiling. (*In* Proceedings of the *Workshop on Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000)*. 1-8 October 2000. Hong Kong, 1-6.)

Recanati, François. 2004. *Literal Meaning*. Cambridge: Cambridge University Press.

Reppen, Randi, Susan M. Fitzmaurice and Douglas Biber. 2002. *Using Corpora to*

Explore Linguistic Variation. Amsterdam: John Benjamins. 274 p.

Roberge, Paul T. 2002. Afrikaans: considering origins. (*In* Mesthrie, Rajend, *ed*. *Language in South Africa*. Cambridge: Cambridge University Press, 79-103.)

Roberts, Ian and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press. 288 p.

Romaine, Suzanne. 1982. *Socio-historical Linguistics: its Status and Methodology*. Cambridge: Cambridge University Press. 315p.

Romaine, Suzanne. 1988. Historical sociolinguistics: problems and methodology. (*In* Ammon, U., Dittmanr, N. and Mattheier, K.J., *eds*. *Sociolinguistics: An International Handbook of the Science of Language and Society*, Vol. 3. Berlin: de Gruyter, 1696-1703.)

Rosch, Eleanor. 1975a. Universals and cultural specifics in human categorization. (*In* Brislin, R. W., S. Bochner and W. Lonner, *eds*. *Cross-cultural Perspectives on Learning*. New York: Wiley, 177-206.)

Rosch, Eleanor. 1975b. Cognitive representations of semantic categories. *Journal of Experimental Psychology*, 104(3): 192-233.

Rosch, Eleanor and Carolyn B Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4): 573-605.

Rossouw, Ronel and Bertus van Rooy. 2012. Modality in South African English. *English World-Wide*, 33(1): 1-26.

Russell, David R. 2002. *Writing in the Academic Disciplines: A Curricular History*. Second Edition. Illinois: Southern Illinois University Press. 410 p.

Salkie, Raphael. 2009. Degrees of modality. (*In* Salkie, Raphael, Pierre Busuttil and Johan van der Auwera, *eds. Modality in English: Theory and Description.* Berlin and New York: Mouton de Gruyter, 79-104.)

Sapir, Edward. 1929. The status of linguistics as a science. *Language*, 5(4): 207-214.

Saussure, de, Ferdinand. 2011 (1916). *Course in General Linguistics*. Translated from the French by Wade Baskin. Edited by Meisel, Perry and Haun Saussy. New York City: Colombia University Press. 261 p.

Schneider, Edgar W. 2003. The dynamics of New Englishes: from identity construction to dialect birth. *Language,* 79(2): 233-281.

Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.

Silva, Penny. 1996. *A Dictionary of South African English on Historical Principles.* Oxford: Oxford University Press. (In Association with the Dictionary Unit for South African English: Rhodes University, Grahamstown.) 825 p.

Silverstein, Michael. 1981. Metaforces of power in traditional oratory. Paper presented at Yale University.

Smith, Nicolas. 2003. Changes in modals and semi-modals of strong obligation and epistemic necessity in recent British English. (*In* Roberta Facchinetti, Manfred Krug and Frank Palmer, *eds*. *Modality in Contemporary English.* Berlin: Mouton de Gruyter, 241-266.)

Southworth, Franklin. 1990. Synchronic manifestations of linguistic change. (*In* Polomé, Edgar, C., *ed. Research Guide on Language Change. Issue 48 of Trends in Linguistics: Studies and monographs*. Berlin: de Gruyter, 25-35.)

Statistics South Africa. 2003. *Census 2001: Census in Brief.* Pretoria: Statistics South Africa.

Statistics South Africa. 2012. *Census 2011: Census in Brief.* Pretoria: Statistics South Africa.

Steyn, Jaap C. 1980. *Tuiste in eie taal*. Cape Town: Tafelberg.

Stockwell, Peter. 2007. *Sociolinguistics: A Resource Book for Students*. Second Edition. London: Routledge. 254 p.

Sweetser, Eve, E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge: Cambridge University Press.

Tagliamonte, Sali. 2004. *Have to, gotta, must*: grammaticalisation, variation and specialization in English deontic modality. (*In* Lindquist H. and Mair, C., *eds. Corpus Approaches to Grammaticalization in English*. Amsterdam: John Benjamins, 33-55.)

Taylor, John R. 1995. *Linguistic Categorization: Prototypes in Linguistic Theory*.

Second Edition. Oxford: Clarendon Press. 312 p.

Terblanche, Elizabeth D. 2011. *Modelling narrativity in East African English.* Vanderbijlpark: NWU. (Dissertation – M.A.).

Thomason, Sarah G. 2007. Language contact and deliberate change. *Journal of Language Contact*, 1: 41-62.

Tieken-Boon van Ostade, Ingrid. 1987. *The Auxiliary* Do *in Eighteenth-Century English. A Sociohistorical-linguistic Approach*. Dordrecht: Foris.

Timotijevic, Jelena. 2009. Another look at modals and subjectivity. (*In* Salkie, Raphael, Pierre Busuttil and Johan van der Auwera, *eds. Modality in English: Theory and Description.* Berlin and New York: Mouton de Gruyter, 105-122.)

Titlestad, Peter. 1996. English, the constitution and South Africa's language future. (*In* Vivian de Klerk, *ed. Focus on South Africa*. Amsterdam: Benjamins, 163-173.)

Tomasello, Michael. 2000. The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, 4(4): 156-163.

Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge (MA): Harvard University Press. 408 p.

Tomasello, Michael. 2008. *Origins of Human Communication*. Cambridge (MA): MIT Press. 393 p.

Trask, R. L. (edited by Peter Stockwell). 2007. *Language and Linguistics: The Key Concepts*. Second Edition. Oxon: Routledge.

Traugott, Elizabeth Closs. 1972. *A History of English Syntax; a Transformational Approach to the History of English Sentence Structure*. New York: Holt, Reinhart and Winston.

Traugott, Elizabeth Closs. 1976. Review of 'Variation and linguistic theory' by Bailey, Charles-James. *Language*, 52: 502-6.

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language,* 65(1): 31-55.

Traugott, Elizabeth Closs. 1996. Semantic change: an overview. *Glot International*,

2: 3-7.

Traugott, Elizabeth Closs. 1997. Subjectification and the development of epistemic meaning: The case of *promise* and *threaten*. (*In* Swan, Toril and Olaf Jansen Westvik, *eds*. *Modality in Germanic Languages: Historical and Comparative Perspectives.* Berlin: de Gruyter: 185-210.)

Traugott, Elizabeth Closs. 2006. Historical aspects of modality. (*In* William Frawley, Erin Eschenroeder, Sarah Mills, Thao Nguyen, *eds*. *The Expression of Modality*. Berlin: de Gruyter, 107-139.)

Traugott, Elizabeth Closs and Graeme Trousdale. 2010. Gradience, gradualness and grammaticalization: how do they intersect? (*In* Traugott, Elizabeth Closs and Graeme Trousdale, *eds*. *Gradience, Gradualness and Grammaticalization*. Amsterdam: John Benjamins, 19-44).

Tregidgo, Philip. 1982. MUST and MAY: demand and permission. *Lingua*, 56: 75-92.

Trüb, Regina. 2008. *American English Impact on South African English: An Empirical Analysis of its Manifestations and Attitudes Towards it.* Regensburg: Universität Regensburg. (Thesis – Ph.D).

Trudgill, Peter. 2001. Sociohistorical linguistics and dialect survival: a note on another Nova Scotian enclave. (*In* M. Ljung *ed.* *Linguistic structure and variation: a Festschrift for Gunnel Melchers.* Stockholm: Stockholm University Press, 193-211.)

Trudgill, Peter. 2004. *New-dialect formation: the inevitability of colonial Englishes.* Edinburgh: Edinburgh University Press.

Trudgill, Peter. 2008. Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language and Society*, 37: 241-54.

Trudgill, Peter. 2010. *Investigations in Sociohistorical Linguistics: Stories of Colonisation and Contact*. Cambridge: Cambridge University Press.

Van der Auwera, Johan. 2008. In defence of classical semantic maps. *Theoretical Linguistics*, 34(1): 39-46.

Van der Auwera, Johan and Alfonso Zamorano Aguilar. (in press) The history of mood and modality. (*In* Nuyts, Jan and Johan van der Auwera, *eds*. *The Oxford Handbook of Mood and Modality*. Oxford: Oxford University Press.)

Van der Auwera, Johan and Andreas Ammann. 2005. Overlap between situational and epistemic modal marking. (*In* Haspelmath, Martin, Matthew Dryer, David Gil and Bernard Comrie, *eds*. *The World Atlas of Language Structures*. Oxford: Oxford University Press, 310-314.)

Van der Auwera, Johan, Petar Kehayov and Alice Vittrant. 2009. Acquisitive modals. (*In* Hogeweg, Lotte, Helen de Hoop and Andrej Malchukov, *eds*. *Cross-linguistic Semantics of Tense, Aspect and Modality.* Amsterdam: John Benjamins, 271-302.)

Van der Auwera, Johan, Dirk Noël and Astrid De Wit. 2012. The diverging *need (to)*'s of Asian Englishes. (*In* Hundt, Marianne and Ulrike Gut, *eds*. *Mapping Unity and Diversity World-Wide: Corpus-Based Studies of New Englishes.* Amsterdam: John Benjamins, 55-75.)

Van der Auwera, Johan and Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology*, 2: 79-124.

Van der Gaaf, W. 1931. *Beon* and *habban* connected with an inflected infinitive. *English Studies*, 13: 176-188.

Van Kemenade, Ans. 1992. Structural factors in the history of English modals. (*In* Rissanen, M., O. Ihalainen, T. Nevaiainen and I. Taavitsainen, *eds*. *History of English: New Methods and Interpretations in Historical Linguistics.* Berlin/New York: Mouton de Gruyter, 287-309.)

Van Rooy, Bertus. 2006. The extension of the progressive aspect in Black South African English. *World Englishes*, *25*: 37-64.

Van Rooy, Bertus. 2008. Beyond deficit terminology: tense and aspect in Black South African English. *World Englishes*, 27: 335-58.

Van Rooy, Bertus. 2010. Social and linguistic perspectives on variability in World Englishes. *World Englishes*, 29(1): 3-20.

Van Rooy, Bertus. 2011. A principled distinction between error and conventionalised innovation in African Englishes. (*In* Mukherjee, J. and M.

Hundt, *eds. Exploring Second-Language Varieties of English and Learner Englishes: Bridging a Paradigm Gap*. Amsterdam: John Benjamins, 191-209.)

Van Rooy, Bertus. 2014. Convergence and endonormativity at Phase Four of the Dynamic Model. (*In* Buschfeld, Sarah, Thomas Hoffman, Magnus Huber and Alexander Kautzsch, *eds. The Evolution of Englishes: The Dynamic Model and Beyond*. Amsterdam: John Benjamins, 21-38.)

Van Rooy, Bertus. (in press) English in South Africa. (*In* Filppula, M., I. Klemola and D. Sharma, *eds. The Oxford Handbook of World Englishes.* Oxford: Oxford University Press.)

Van Rooy, Bertus and Lize Terblanche. 2010. Complexity in word-formation processes in New Varieties of South African English. *Southern African Linguistics and Applied Language Studies,* 28(4): 357–374.

Van Rooy, Bertus and Ronel Wasserman. 2014. Do the Modals of Black and White South African English Converge? *Journal of English Linguistics*, 42(1): 51-67.

Vaugelas, Claude Favre de. 1647. *Remarques sur la Langue Françioise.* Paris: La veuve J. Camusat et P. Le Petit.

Verstraete, Jean-Christophe. 2001. Subjective and objective modality: interpersonal and ideational functions in the English modal auxiliary system. *Journal of Pragmatics*, 33: 1505-1528.

Verstraete, Jean-Christophe. 2005. Scalar quantity implicatures and the interpretation of modality problems in the deontic domain. *Journal of Pragmatics*, 37: 1401-1418.

Verstraete, Jean-Christophe. 2007. *Rethinking the Coordinate-subordinate Dichotomy: Interpersonal Grammar and the Analysis of Adverbial Clauses in English*. Berlin: Mouton de Gruyter.

Visser, Frederikus Th. 1963-73. *An Historical Syntax of the English Language, 3 Vols.* Leiden: Brill.

Warner, A. R. 1982. *Complementation in Middle English and the Methodology of Historical Syntax*. London and Canberra: Croom Helm.

Warner, A. R. 1993. *English Auxiliaries: Structure and History*. Cambridge: Cambridge University Press.

Wasserman, Ronel and Bertus van Rooy. 2014. The development of modals of obligation and necessity in White South African English through contact with Afrikaans. *Journal of English Linguistics*, 42(1): 31-50.

Watkins, Calvert. 1989. New parameters in historical linguistics, philology, and culture history. *Language*, 65(4): 783-799.

Webster's Complete Dictionary of the English Language. 1886. *must,* v. https://archive.org/details/websterscomplete00webs Date of access: 19 April 2014.

Westin, Ingrid. 2002. *Language Change in English Newspaper Editorials.* Amsterdam: Rodopi. 204 p.

Westney, P. 1995. *Modals and Periphrastics in English: an Investigation into the Semantic Correspondence between certain English Modal Verbs and their Periphrastic Equivalents*. Tübingen: Max Niemeyer Verlag.

Wittgenstein, L. 2001 (1953). *Philosophical Investigations (Philosophische Untersuchungen)*. Oxford: Blackwell.

Wright, Joseph. 1896-1905. *English Dialect Dictionary.* Oxford: Oxford University Press.

Wright, G. H. 1951. *An Essay in Modal Logic.* Amsterdam: North-Holland.

Zadeh, L. A. 1972. A fuzzy-set-theoretic interpretation of linguistic hedges. *Journal of Cybernetics*, 2(3): 4-34.

Zelinsky, Wilbur. 1992 (1973). *The Cultural Geography of the United States.* Second Edition. Englewood Cliffs, NJ: Prentice-Hall.