

**The role of editorial intervention in ongoing
language variation and change in South African
and Australian English**

M Law

 orcid.org/0000-0001-7525-3553

Thesis accepted for the degree Doctor of Philosophy in
Linguistics and Literary Theory at the North-West University

Promoter: Prof AJ van Rooy

Graduation: July 2019

Student number: 13090658

Statement of originality

This thesis is being submitted to Macquarie University and North-West University in accordance with the Cotutelle agreement dated 17 December 2015.

To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.



(Signed):

Date: 31 December 2018

Melanie Ann Law

Abstract and keywords

The role of editorial intervention in written published texts has been noted by a number of researchers in areas of World Englishes as well as language variation and change. The widely held view of these scholars is that editorial intervention is a primarily normative activity in which editors fulfil a gatekeeper function by consciously removing innovative features and limiting variation in the texts they edit. As a result of this view, it is often argued that editorial intervention is an obstacle to the processes of change, and within the context of World Englishes, the progression of varieties towards endonormativity. There is no doubt that this view is partly true: part of editors' work is to consciously match the text that they edit with usage sanctioned in the norm-providing sources (or overt norms) that their employers/clients require them to use. However, editors' removal of the features associated with ongoing change must be a matter of degree, since corpus-based investigations of written published texts have shown that features and usage patterns associated with ongoing change do actually occur in texts that have undergone editorial intervention, and that sometimes these types of texts are highly receptive to these features.

This study proceeds from the view that because most written published texts undergo some form of editorial intervention, the language of published written texts, which forms part of the input that language users are exposed to, cannot be solely attributed to the author of the text, and it is therefore necessary to investigate the contributions of editors to these texts. Furthermore, this study argues that editorial work forms an important part of processes of language variation and change, particularly in World Englishes where investigations of editorial intervention in written published texts might provide some information on the acceptability of these features within a variety. Thus, this study focuses on the interaction of overt norms and covert norms in conditioning editors' acceptability judgements of such features, which may shed light on how varieties progress towards endonormativity.

To do this, this study investigates how editors working in different varieties of English, particularly Australian English and South African English, respond to the presence of a particular linguistic feature associated with ongoing change in written published registers. The feature this study is concerned with is genitive alternation, and the interest of the study is in how editorial work influences the patterns of variation (and potentially change) of this feature across the two varieties investigated.

Three research questions guide this study. The first question is theoretical and explores how broad trends of language change, the different contextual forces of the varieties of English, register effects and editorial practice interact to potentially influence language variation and change. The second question zooms in on editors and editorial work and enquires into the sociolinguistic profiles of editors (providing information on their covert norms) and the norm-providing sources (as a reflection of the overt norm) used by editors of English texts in Australia and South Africa. Against this background, the study then focuses on how editors of English texts in the two varieties respond to the presence of one feature, namely genitive

alternation. The genitive alternation is selected as the feature to investigate in the current study for several reasons: there is evidence of ongoing change in the use of this feature in present-day English; this change can be linked to broader processes of language change, namely colloquialisation and densification, which are known to play out differentially in different varieties of English and across different written published registers; and some normative advice exists for the feature (reflecting overt norms), but it is not so saliently marked that it would be specifically targeted by editors (providing room for covert norms to condition editorial choices). Furthermore, genitive alternation is a good example of a linguistic variable in the variationist sense, and is known to be conditioned by both language-internal and language-external factors. It can therefore be investigated by drawing on variationist methods and state-of-the-art statistical techniques. The third research question draws together the theoretical and empirical dimensions of the study to reflect on how editorial intervention in different varieties of English and across different registers interacts with broader processes of language change, specific processes of language change in varieties of English, and register effects, to influence different opportunities for and constraints on the processes of dissemination and conventionalisation in a variety's progression towards endonormativity.

To answer the first research question, the study adopts a usage-based view of language that integrates cognitive and social factors to account for language structure, use and ongoing change in World Englishes. It makes an innovative theoretical contribution by positioning editorial influence as an important mechanism in the processes of variation and change in different written published registers in different varieties of English. To answer the second research question, the study surveys editors of English texts in Australia and South Africa in order to gather information on their sociolinguistic profiles (as a reflection of their covert norms), and the kinds of norm-providing sources that they use to guide their editorial choices (as a reflection of their overt norms). To investigate the actual changes that editors make to the texts they edit and their treatment of genitive constructions across five written published registers, a corpus-based approach is adopted in which a register-differentiated, parallel corpus of unedited texts and their edited counterparts representing each of the two varieties and five written published registers (academic, creative, instructional, popular and reportage) is constructed.

The findings of the empirical investigation show that editors of English texts in Australia and South Africa are mostly older females who are highly educated and, who in sociolinguistic terms, are therefore most likely more conservative language users who prefer standard, prestige forms of the language (as sanctioned in overt norm-providing sources). However, the contrast between the more homogenous linguistic landscape in Australia and the heterogenous linguistic landscape in South Africa is reflected both in editors' language profiles and choice of norm-providing sources: the Australian editors are all first-language users of the variety and draw on norm-providing sources for Australian English while the South African editors are mostly English–Afrikaans bilinguals who are either first- or second-language users of two sub-varieties in the country, White South African English and Afrikaans English. The Australian editors therefore represent a group of editors for whom the normative environment is homogenous and in which overt norms

are closely aligned to covert norms, while the South African editors represent a group of editors who work in a much more diffuse normative environment and in which there is a high degree of heterogeneity in covert norms and overt norms.

The findings of the corpus-based investigation show the importance of covert norms in influencing editors' acceptability judgements of features associated with ongoing change in different varieties of English, but demonstrate that this is not simply a matter of a broad, national usage pattern. Instead, the findings demonstrate that stylistically distinct usage patterns of features associated with ongoing change may arise at the level of a variety more broadly, but may also arise in contexts where multiple sub-varieties interact with each other in the same context. Furthermore, these stylistically distinct patterns of features associated with ongoing change are informed by the unique interaction of many factors in different varieties of English, including differing stylistic preferences across sub-varieties, substrate influence, the amount, type and duration of contact among the language users of the different sub-varieties, the differing strength of linguistic factors, the progression of the sub-variety along the stages of the Dynamic Model and the kinds of overt norms available. In other words, the findings show that while overt norms are more saliently represented in editors' mental representations of language than they are for other language users, editors are also frequently exposed to changing and differing usage patterns through the process of reading and evaluating unedited writing, and as a result they largely accept the distinctive usages of authors. This shows that endonormativity may be achieved at the level of a variety more broadly, as is the case for Australian English, or it may be achieved at the level of individual (but interacting) sub-varieties, as is shown to be the case for South African English.

The study therefore demonstrates how editorial intervention forms part of the processes of ongoing language variation and change in different varieties of English and across different written registers, and how, at least in the case of the feature investigated in this study, editorial intervention helps to accept distinctive stylistic usages of the feature in different varieties of English, legitimising their use and contributing to their further dissemination.

Keywords: editorial intervention, normativity, language variation and change, World Englishes, register studies, genitive alternation, Australian English, South African English, colloquialisation, densification

Acknowledgments

I owe the realisation of this thesis to many people located on opposite sides of the Indian Ocean, all of whom have contributed, in one way or another, to making my PhD journey an extraordinary one. There are too many to mention by name, but some I must single out.

It has been a special privilege to have been guided through this study by two of the kindest and most knowledgeable supervisors a PhD candidate could wish for. This thesis would not have been possible without my two principal supervisors, **Dr Haidee Kruger** and **Prof. Bertus van Rooy**. Thank you for the countless hours you have spent reading these pages, providing feedback, and patiently listening to and answering my many questions. All remaining errors are my own.

I have also been fortunate enough to have been supervised by **Emeritus Prof. Pam Peters**, who generously shared her extensive knowledge of World Englishes and the Australian publishing and editing landscape with me.

During my candidature, I was fortunate enough to spend 2017 working on this thesis in the Department of Linguistics at Macquarie University. My thanks go to the **all members of staff and PhD candidates in the Department** who welcomed me into their fold and whose words of kindness and encouragement are greatly appreciated.

My time at Macquarie University was made possible because I was granted a year's research leave from my employer, North-West University. I would therefore like to thank **Prof. Susan Coetzee Van Rooy** and **Associate Prof. Gustav Butler** for making this possible and also for their interest in the study and encouragement throughout its completion.

I am very grateful to my colleagues at the School of Languages on the North-West University's Vaal Triangle Campus for the support they have shown me. Special thanks are due to **Dr Esté Hefer-Jordaan**, **Dr Ella Wehrmeyer**, **Mari-Leigh Pienaar**, **Wendy Barrow**, **Karien Redelinghuys**, **Keabetswe Motlodi** and **Marist Partridge**, who assumed my work duties during my time in Australia.

Financial assistance for the completion of this thesis was made available through the awarding of an International Macquarie Research Excellence Scholarship (iMQRES) by the **Macquarie University Higher Degree Research Office** and a bursary by the research focus area **Understanding and Processing Language in Complex Settings (UPSET)** at the North-West University's Vaal Triangle Campus.

Thanks are also due to all the **editors in Australia and South Africa** who completed the questionnaire used in this study and all the **editors, authors and contributors** who generously supplied the texts used for the corpus constructed in this study. This research would not have been possible without their generosity and willingness to participate, and for this I am extremely grateful.

My family and friends have been a source of constant support and love, not only for the duration of my candidature but in all areas of my life. Special thanks go to my parents, **Gary** and **Hilary** and my sisters, **Colleen** and **Bronwen**.

And last, but most certainly not least, to **Daive**, who has been the greatest cheerleader of all: Thank you for your love and patience, for all the sacrifices you have made to ensure that I had the space and time to complete this thesis and for your never-ending and unfaltering encouragement.

Contents

List of figures	xiii
List of tables.....	xv
Chapter 1 Introduction, contextualisation and problem statement	1
1.1 Introduction and rationale.....	1
1.2 Contextualisation.....	4
1.3 Research questions	12
1.4 Research objectives	13
1.5 Methodology	13
1.6 Chapter outline	15
Chapter 2 Theoretical framework: Editorial intervention, language variation and language change in written varieties of English	17
2.1 Introduction	17
2.2 Theoretical approaches to language structure and use	20
2.2.1 The cognitive organisation of language.....	24
2.2.2 The social organisation of language	29
2.3 Language change: Evolutionary theories	32
2.4 Variation in the world's Englishes	41
2.5 Register and language variation	45
2.6 Editorial intervention, norms and (endo)normativity	48
2.6.1 Norms and normativity	48
2.6.2 The nature of editorial intervention and the development of norms in PCEs.....	49
2.7 Conclusion.....	53
Chapter 3 Varieties, registers and grammatical feature investigated	55
3.1 Introduction	55
3.2 Varieties under investigation.....	55
3.2.1 Australian English	56
3.2.2 South African English.....	59
3.2.3 AusE and SAfE: Endonormativity and editing in sum	66
3.3 Registers investigated in this study	67
3.4 Feature under investigation: The genitive alternation	72
3.4.1 The constructions.....	75
3.4.2 Historical changes in the genitive alternation.....	78

3.4.3	Factors influencing genitive alternation.....	79
3.4.3.1	Linguistic factors conditioning the genitive alternation.....	80
3.4.3.2	Language-external factors conditioning the genitive alternation.....	85
3.4.3.2.1	Genitive alternation across varieties.....	85
3.4.3.2.2	Genitive alternation across modality.....	89
3.4.3.2.3	Genitive alternation across written registers.....	90
3.5	Motivation for the selection of the varieties, registers and feature investigated in this thesis.....	92
3.6	Conclusion.....	97
Chapter 4	Methodology: Combining survey and corpus research to investigate the effects of editing in written varieties of English.....	99
4.1	Introduction.....	99
4.2	Questionnaire.....	100
4.2.1	Questionnaire design.....	100
4.2.2	Piloting of the questionnaire.....	101
4.2.3	Participants and sample.....	102
4.2.4	Processing and analysis of questionnaire data.....	103
4.3	Corpus-based investigation.....	108
4.3.1	Corpus design.....	108
4.3.2	Corpus construction and size.....	111
4.3.3	Data extraction and annotation.....	117
4.3.4	Data analysis and interpretation.....	122
4.4	Ethics application and considerations.....	125
4.5	Conclusion.....	126
Chapter 5	Presentation and analysis of the empirical data: Evidence for the role of editorial intervention in ongoing language change in AusE and SAfE and across written published registers.....	127
5.1	Introduction.....	127
5.2	Presentation and analysis of questionnaire data.....	127
5.2.1	Demographic profile of respondents.....	128
5.2.2	Respondents' language profiles.....	139
5.2.3	Norm-providing sources.....	148
5.2.3.1	Overview of responses.....	149
5.2.3.2	Australian respondents.....	153
5.2.3.3	South African respondents.....	160
5.2.3.4	Comparison of norm-providing sources.....	168
5.3	Presentation and analysis of corpus findings.....	172
5.3.1	Distributional analysis of genitive constructions in AusE and the sub-varieties of SAfE.....	172

5.3.1.1	Distributional analysis: Proportional frequency.....	173
5.3.1.2	Distributional analysis: Normalised frequency	178
5.3.1.3	Statistical evaluation of the interaction between variety and register	180
5.3.2	The influence of editorial intervention on genitive alternation across varieties and registers...181	
5.3.2.1	Distributional and statistical analysis of the influence of editorial intervention on the proportional frequencies of genitive variants across variety and register	182
5.3.2.2	Distributional analysis of the influence of editorial intervention on the normalised frequencies of genitive variants across variety and register.....	186
5.3.2.3	The types of editorial changes that influence the shifting frequencies of the genitive constructions across variety and register.....	189
5.3.3	Statistical evaluation of the interaction of linguistic and language-external factors in genitive alternation in AusE and SAfE.....	202
5.3.3.1	Random forest analysis of the factors influencing genitive alternation across unedited and edited writing	203
5.3.3.1.1	Contrastive random forest analysis for the effect of editorial intervention across varieties	203
5.3.3.1.2	Contrastive random forest analysis for the effect of editorial intervention in AusE.....	206
5.3.3.1.3	Contrastive random forest analysis for the effect of editorial intervention in WSAfE.....	208
5.3.3.1.4	Contrastive random forest analysis for the effect of editorial intervention in AfrE	210
5.3.3.1.5	Contrastive random forest analysis for the effect of editorial intervention in BSAfE.....	212
5.3.3.1.6	Summary of findings for the contrastive random forest analysis.....	214
5.3.3.2	Conditional inference tree analysis of the interaction between linguistic and language-external factors in influencing genitive alternation.....	217
5.3.3.2.1	Conditional inference tree analysis of the interaction of linguistic and language-external factors in AusE.....	218
5.3.3.2.2	Conditional inference tree analysis of the interaction of linguistic and language-external factors in WSAfE.....	222
5.3.3.2.3	Conditional inference tree analysis of the interaction of linguistic and language-external factors in AfrE.....	225
5.3.3.2.4	Conditional inference tree analysis of the interaction of linguistic and language-external factors in BSAfE.....	228
5.4	Summary of the findings	230
5.5	Conclusion.....	235
Chapter 6 Synthesis of findings, recommendations for future research and conclusion.....		237
6.1	Introduction	237
6.2	Summary of findings	237
6.2.1	Findings for research question 1	238
6.2.2	Findings for research question 2.....	240
6.2.2.1	Findings for research questions 2.1 and 2.2.....	241
6.2.2.2	Findings for research question 2.3	244

6.3 Findings for research question 3: Theoretical implications for studies of language change, World Englishes and across written registers	247
6.4 Limitations of the study and recommendations for further research	248
6.5 Concluding remarks	249
Bibliography.....	250
Appendix A: Draft questionnaire for pilot.....	266
Appendix B: Final questionnaire	283
Appendix C: Source type categorisation	299
Appendix D: Calculation of total estimated proportions and normalised frequencies	308
Appendix E: Participant Information and Consent Forms	313
Appendix F: Ethics approval letter.....	317

List of figures

Figure 5.1: Respondents' age by country.....	129
Figure 5.2: Number of years' editing experience.....	131
Figure 5.3: Freelance versus in-house employment.....	134
Figure 5.4: Respondents' conceptualisation of their editorial role.....	137
Figure 5.5: Respondents' first or home language.....	140
Figure 5.6: Mono-, bi- and multilingualism among the respondents, per sample.....	143
Figure 5.7: Respondents' strongest editing language.....	146
Figure 5.8: Proportional frequencies of <i>s</i> - and <i>of</i> -genitives in published academic writing across sub-varieties, in edited texts	174
Figure 5.9: Proportional frequencies of <i>s</i> - and <i>of</i> -genitives in published creative writing in WSAfE and AusE, in edited texts	175
Figure 5.10: Proportional frequencies of <i>s</i> - and <i>of</i> -genitives in published instructional writing across sub-varieties, in edited texts	176
Figure 5.11: Proportional frequencies of <i>s</i> - and <i>of</i> -genitives in published popular writing across sub-varieties, in edited texts	176
Figure 5.12: Proportional frequencies of <i>s</i> - and <i>of</i> -genitives in reportage across sub-varieties, in edited texts.....	177
Figure 5.13: Conditional inference tree for the proportion of <i>s</i> -genitives, with register and sub-variety as predictors	180
Figure 5.14: Conditional inference tree for the proportion of <i>s</i> -genitives, with register, sub-variety and edited status as predictors.	184
Figure 5.15: Variable importance plot for unedited writing, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables ...	204
Figure 5.16: Variable importance plot for edited writing, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER, AUTHOR LANGUAGE and EDITOR LANGUAGE as predictor variables.....	205
Figure 5.17: Variable importance plot for unedited AusE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables ...	206
Figure 5.18: Variable importance plot for edited AusE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITOR LANGUAGE as predictor variables	207
Figure 5.19: Variable importance plot for unedited WSAfE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables.	208
Figure 5.20: Variable importance plot for edited WSAfE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS,	

SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITOR LANGUAGE as predictor variables.....	209
Figure 5.21: Variable importance plot for unedited AfrE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables...	211
Figure 5.22: Variable importance plot for edited AfrE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER, and EDITOR LANGUAGE as predictor variables.....	211
Figure 5.23: Variable importance plot for unedited BSafE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables...	213
Figure 5.24: Variable importance plot for edited BSafE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITOR LANGUAGE as predictor variables.....	213
Figure 5.25: Conditional inference tree for AusE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER and EDITED STATUS in conditioning genitive alternation	221
Figure 5.26: Conditional inference tree for WSafE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER and EDITED STATUS in conditioning genitive alternation	224
Figure 5.27: Conditional inference tree for AfrE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITED STATUS in conditioning genitive alternation.....	227
Figure 5.28: Results of the conditional inference tree analysis for BSafE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, SEMANTIC RELATION, EDITOR LANGUAGE and EDITED STATUS with the response variable (with two levels <i>s</i> -genitive and <i>of</i> -genitive)	229

List of tables

Table 3.1: Animacy scale for plotting possessor noun type	81
Table 3.2: Distribution of genitives for AusE and AmE (calculated from Ford and Bresnan (2015)).	88
Table 3.3: Distribution of genitive constructions across modalities and genres (adapted from Grafmiller (2014, p. 476)).....	91
Table 3.4: Genitive constructions available in Afrikaans, English, Southern Sotho and Zulu.	96
Table 4.1: Categorisation of norm orientations.....	107
Table 4.2: Text and approximate unique author and editor count per register, per sub-variety.....	115
Table 4.3: Word count per register, per sub-variety.....	116
Table 4.4: Representation of editors across registers	116
Table 4.5: Language codes used for authors' and editors' languages.....	119
Table 4.6: Editorial intervention labels used in the data annotation	120
Table 5.1: Respondents' gender by country	130
Table 5.2: Respondents' highest qualifications.....	131
Table 5.3: Categorisation of qualification areas.....	132
Table 5.4: In-house employment details per country	134
Table 5.5: Categorisation of text types.....	135
Table 5.6: Respondents' experience editing different text types	136
Table 5.7: Comparison of first or home language and strongest language: South Africa.....	142
Table 5.8: Percentage of respondents with second, third, fourth and fifth languages, per sample	142
Table 5.9: Respondents' second strongest languages	144
Table 5.10: Respondents' third, fourth and fifth strongest languages.....	145
Table 5.11: Combination of respondents' first or home, strongest and strongest editing language(s): South Africa.....	147
Table 5.12: Respondents who edit in second, third, fourth and fifth languages	147
Table 5.13: Number of responses and number of distinct sources per source type category.....	149
Table 5.14: Number and percentage of respondents per source type category	151
Table 5.15: Number of responses per norm orientation for the top five source types: Australia.....	154
Table 5.16: Most frequently selected dictionaries by the Australian respondents	155
Table 5.17: Most frequently selected style guides by the Australian respondents.....	156
Table 5.18: Most frequently selected usage guides by the Australian respondents	157
Table 5.19: Most frequently selected editing and writing sources by the Australian respondents	158
Table 5.20: Most frequently selected language reference works by the Australian respondents.....	159
Table 5.21: Number of responses per norm orientation for the top five source types: South Africa.....	161
Table 5.22: Most frequently selected dictionaries by the South African respondents	162
Table 5.23: Most frequently selected style guides by the South African respondents.....	164

Table 5.24: Most frequently selected language reference works by the South African respondents	165
Table 5.25: Most frequently selected editing and writing sources by the South African respondents	166
Table 5.26: Most frequently selected usage guides by the South African respondents.....	167
Table 5.27: Responses for glossaries, thesauruses, people and various online and print sources	168
Table 5.28: Overall proportional frequencies of the <i>s-</i> and <i>of-</i> genitives in the sub-varieties, in edited texts.	173
Table 5.29: Normalised frequencies (per 1 000 words) of <i>s-</i> and <i>of-</i> genitives per sub-variety and register, in edited texts	179
Table 5.30: Proportional frequencies of <i>s-</i> and <i>of-</i> genitives across variety, register, and unedited and edited texts	182
Table 5.31: Normalised frequencies (per 1 000 words) of <i>s-</i> and <i>of-</i> genitives per sub-variety and register, in unedited and edited texts.....	187
Table 5.32: Percentages of the types of editorial changes influencing the frequencies of the <i>s-</i> and <i>of-</i> genitive constructions across registers in AusE	190
Table 5.33: Percentages of the types of editorial changes influencing the frequencies of the <i>s-</i> and <i>of-</i> genitive constructions across registers in AfrE.....	192
Table 5.34: Percentages of the types of editorial changes influencing the frequencies of the <i>s-</i> and <i>of-</i> genitive constructions across registers in BSAfE	194
Table 5.35: Percentages of the types of editorial changes influencing the frequencies of the <i>s-</i> and <i>of-</i> genitive constructions across registers in WSAfE	198
Table 5.36: Summary of the variable importance scores.....	215

Chapter 1

Introduction, contextualisation and problem statement

1.1 Introduction and rationale

The global spread of English and its diverse roles in societies across the globe have given rise to different varieties of English, which have been widely studied within the World Englishes paradigm since the 1980s. Various influential frameworks and models have been developed to categorise these varieties of English and to describe their functions and development, including Kachru's (1985) Three Circles Model and Schneider's (2003a, 2007) Dynamic Model of Postcolonial Englishes (henceforth PCEs).

In both models, an essential cause of the differences between these varieties is the processes of language change set in motion in the different contexts in which these transplanted varieties of English have developed (Schneider, 2007). Language change comprises two distinct, but equally important processes: innovation, which is the creation of new forms or usage patterns, and propagation, which is the conventionalisation of these forms (Croft, 2000). The notion of norms is crucial for understanding these two processes of change, because "innovation is essentially language use beyond convention, and propagation is essentially the establishment of a new convention" (Croft, 2000, p. 4).

Normativity is also a central concept in the models of the varieties of English proposed by Kachru (1985) and Schneider (2003a, 2007). Kachru (1985) clusters different varieties of English into three circles on the basis of national borders and whether English as used within those borders can be seen as norm-providing, norm-developing or norm-dependent. In Inner Circle countries English is used by native users, and English usage in these countries serves as norms for other countries. Inner Circle countries include the United States, Britain, Australia and New Zealand. Outer Circle countries are those countries in which English is used by native users and non-native, second-language users. Such countries are considered to be in the process of developing their own norms, and include Singapore and India. In Expanding Circle countries, English is used as a foreign language and is acquired for the purposes of international communication, such as in China or Brazil. Expanding Circle countries are norm-dependent and rely on the norms of Inner Circle countries to model English usage (Kachru, 1985, p.17). Normativity is just as crucial in Schneider's (2003a, 2007) Dynamic Model, which, rather than clustering varieties on the basis of speaker status, proposes an underlying developmental process of postcolonial varieties of English in which at least two parties in the colonial contact setting (the settler and the indigenous populations) move through five stages of identity rewriting and linguistic transformation processes. Varieties progress from an exonormative orientation in the early stages of the Dynamic Model to an endonormative orientation in the later stages of the model, ultimately culminating in the emergence of a unified national identity and a new variety of English with new linguistic norms (Schneider, 2003a, 2007). The five stages of the model are Foundation, Exonormative Stabilisation, Nativisation, Endonormative Stabilisation, and Differentiation. Two stages in the model,

namely Exonormative Stabilisation and Endonormative Stabilisation, explicitly address the normative consolidation of PCEs, reflecting a shift in orientation from the external norms of the imperial centre, to the internal norms of the newly forged variety.

Despite the importance of norms in models of varieties of English, very little attention has been given to the processes by which local norms develop within the context of different varieties, particularly in PCEs, particularly in written language, and particularly with regard to the development and influence of local norm-giving authorities, such as writers, teachers, media practitioners, examination bodies, publishing houses and influential opinion leaders (Bamgbose, 1998, p. 4). This thesis is interested in the role of one further type of norm-giving authority: editors of written published texts. While editorial practices in World Englishes have not been widely studied, some researchers have proposed that editorial practices may be viewed as mechanisms that aid the conventionalisation of innovations and changing usage practices in written registers, while simultaneously acting as forces for stability. For example, Kruger and Van Rooy (2017) argue that the editorial processes that all published texts undergo may weed out innovative features and usages, thereby blocking their acceptance, but they may also allow such features into written texts and thus encourage their dissemination. In so doing, editing acts as a mechanism that helps conventionalisation along. Thus editors' acceptance of innovative features and changing usage practices in a particular variety may be a significant marker of and contributing force to the endonormative stabilisation of the variety, firmly situating editorial practice as an important force in the development of different varieties of English towards endonormativity.

Apart from the work of Kruger (forthcoming), Kruger and Van Rooy (2017), and Van Rooy and Kruger (2016), there is hardly any work on the role of editorial intervention in the processes of ongoing change in different varieties of English. The dearth of research in this regard is the consequence of various theoretical, conceptual and methodological difficulties.

The first difficulty relates to dominant linguistic-theoretical paradigms of the latter half of the twentieth century that abstracted the study of language away from language use, and consequently from the variation that arises in language through use and which drives language change. However, the recent emergence of usage-based theoretical approaches that integrate cognitive and social factors in understanding language, have allowed linguists to re-examine and consolidate the realities of language, making available meaningful systems to explore, describe and explain the variation that arises in the different contexts in which language is used (Van Rooy, 2010).

The second factor relates to the centrality of spoken language in studies of change and the tendency to dismiss written language as its derivative. Consequently, less attention has therefore been afforded to the important role that written language plays in shaping language use and in driving language change. Furthermore, when written language, and in particular published written language, is investigated, there is

very little comprehensive reflection on the unique production circumstances for this form of language. For example, although the potential influence of editorial intervention is often noted by scholars of language variation and change hardly any work has investigated what this influence entails, with most scholars simply framing editorial intervention as a factor that ‘dirties’ language data. As a result, editorial intervention is largely dismissed as a factor that weeds out variation and blocks change, without any real empirical evidence to support this.

The conceptualisation of the editorial role is further complicated by the different contexts in which the varieties of English develop as well as their differential progression through the stages of the Dynamic Model (in which different varieties, at different stages, have varying and changing norm-orientations). Since norm-orientation and the acceptability of local norm-giving authorities are important considerations in the conventionalisation of innovative usage patterns, the question that arises is how editors working with different varieties of English in different contexts respond to the presence of innovative forms or patterns of use. Two additional complicating factors are that different written published registers¹ have been shown to be variably receptive to language change (Hundt & Mair, 1999), while editorial intervention has also been shown to vary across registers (Law, 2011, 2014). This raises further questions about the variability of the editorial role across different registers.

The last difficulty relates to the methodological challenges associated with investigating the editorial role in process of language variation and change. The foremost practical challenge is that it is difficult to obtain texts in both their original and edited forms since original forms are not publically available (and in many cases no longer exist). Furthermore, corpora that seek to investigate editorial changes need to be substantive in size: because editors do not change every word in the texts that they edit, large numbers of texts are needed to ensure that a meaningful number of interventions are included.

This study attempts to address the scarcity of research on the role of editorial intervention in processes of variation and change in different varieties of English. Specifically, it engages with usage-based theories of language use and language change, models of World Englishes and register variation in order to develop a theoretical framework within which to explore and explain the role of editorial intervention in the progression of the varieties towards endonormativity. The empirical investigation takes the form of a case study focusing on two varieties of English that have progressed along the stages of the Dynamic Model

¹ In this study, a register is regarded as a variety of language that is defined in terms of its situational characteristics and communicative purpose (Biber & Conrad, 2009, p. 6). Furthermore, there are usually important linguistic differences across different registers that correspond to the differences in the situational characteristics and communicative purposes of registers (Biber & Conrad, 2013, p. 3). Therefore, the linguistic choices that give rise to the linguistic differences across registers, and consequent variation among them, are highly systematic and functional because they are linked to particular non-linguistic factors such as the addressor’s purpose in communication, the topic, the relationship between the addressor and addressee, the type of readership and the production circumstances (Biber & Conrad, 2009, p. 3). The concept of register is dealt with in more detail in Section 2.5.

differentially, namely Australian English (henceforth AusE) and South African English (henceforth SAfE). The empirical study has two primary components. First the language profiles of and norm-providing sources used by editors in the two countries are investigated in order to analyse the interplay between overt and covert norm orientations. Subsequently, a corpus-based analysis is carried out of the influence of editorial choices on one linguistic feature that is known to currently be undergoing change, namely genitive alternation (that is, the choice between constructions such as *the CEO of the company* (referred to as the *of*-genitive) and *the company's CEO* (referred to as the *s*-genitive)), across five published written registers in the two varieties. This analysis draws on an innovative corpus of unedited texts and their edited counterparts, in the two varieties.

1.2 Contextualisation

As already discussed in Section 1.1, norms are central to the two processes of language change because in the first a variant feature or usage pattern (that is, one that deviates from the norm) is introduced into the linguistic system (giving rise to variation), and in the second this feature or usage pattern spreads throughout the community, eventually becoming conventionalised and being established as a 'new' norm. Labov (1966) and Weinreich, Labov, and Herzog (1968) were the first to show that variation is a normal part of language and language use, but they were also the first to show that it is variation that drives language change and that there is structure to variation (and by implication, to the normative pressures that act on this variation).

Advances in the theoretical understanding of language have given rise to the family of usage-based approaches to investigating language structure, variation and change. These theoretical approaches have highlighted the important role of both cognitive and social factors in language use, which is seen as the central force that shapes language structure, variation and change. For example, the two processes of language change are enabled by distinct mechanisms: the mechanisms for innovation are functional and involve individual psycholinguistic re-mappings of the link between form and function, while the mechanisms that provide for propagation are the cognitively and socially conditioned selection and conventionalisation of innovations by a society of language users (Croft, 2000, p. 166). Thus, norms have many dimensions. In cognitive usage-based linguistics, norms are understood as the cognitive entrenchment of a feature as a result of exposure to it in actual instances of language use. In sociolinguistics, norms operate at the level of the language community, and are regarded as the conventionalised usage patterns within a community of language users. Given the importance of conventionalisation for the establishment of norms and given the centrality of norms for World Englishes, it is important to investigate how cognitive conventionalisation and social conventionalisation interact in the processes of ongoing language change.

Studies of recent language change in English have placed significant emphasis on the development of literacy and the changing, bidirectional ways in which spoken and written language interact (Leech, Hundt,

Mair, & Smith, 2009; Mair, 2006). This interaction is evident in two broad trends of change in English: colloquialisation and densification. Colloquialisation is the narrowing of the stylistic gap between speech and writing, or the shift to a more speech-like style in writing, as a result of a trend towards informality (Leech et al., 2009; Mair, 2006).² This is manifested through, for example, an increase in the use of contracted verbs and negatives, and first person pronouns in writing (Leech et al., 2009). Densification is the compacting of more information into fewer words in response to the information explosion, and is associated with the development of more dense and complex phrasal (rather than more expanded clausal) discourse styles (Biber & Gray, 2012; Leech et al., 2009; Mair, 2006). Densification can be observed in, among others, increased lexical density and the increasing use of nominalisations in written texts (Leech et al., 2009).

An important point to have emerged in studies of these trends of change in English is that the trends cannot be generalised to the language as a whole (Biber, 1999), precisely because English is used in many different contexts and under vastly different conditions. Register is one crucial context of use, and Biber (2012) argues that patterns of linguistic variation and use are different in spoken and written registers, and that studies of language variation and change should therefore separate the two registers. Researchers have also argued that different written registers display divergent behaviour in terms of their receptivity to the features associated with ongoing change, including innovative usages associated with language contact (Biber & Gray, 2012, 2013; Hundt & Mair, 1999). Biber and Gray (2013), for example, have found differences in the dissemination of features associated with ongoing language change in newswriting, academic prose and popular writing. Hundt and Mair (1999) argue that as a result of this varying openness to broader trends of change, registers can be described as ‘agile’ or ‘uptight’ in relation to their acceptance of innovative forms and usage patterns.

In addition to distinguishing between spoken and written registers, as well as different written registers, researchers have also pointed to differences in trends of change in different varieties of English. Hundt (2007), for example, explores the frequency of the progressive passive in different varieties of English, and states that an analysis and comparison of this feature in different varieties demonstrates that globally used grammatical patterns display different behaviour in local usage settings. Situating this within the World Englishes paradigm, Hundt (2007, p. 288) explicitly concludes that different varieties of English do not necessarily cluster together in their usage of global grammatical structures, and adds that:

² Various labels, such as ‘popularisation’ and ‘informalisation’ are often associated with colloquialisation. Leech et al. (2009, p. 239) explain that in principle there is a difference between informalisation and colloquialisation. Informalisation has to do with informality of style and is related to the absence of distance between a text’s creator and its recipient, something that is strongly associated with speech; while colloquialisation is related to the adoption of speech-like patterns in writing (Leech et al., 2009, p. 239). This study notes the difference between the two terms, but does not draw a sharp distinction between them and treats informalisation and colloquialisation as aspects of the same trend.

different local and stylistic preferences or pressures, substrate transfer influence and learner phenomena in contact varieties, ongoing diachronic change as well as the possibility of parallel (but largely independent) internal developments are some of the factors that are likely to account for the complex patterns of variation we find.

Genitive alternation, is one grammatical feature that is currently known to be undergoing change (and for which there is evidence of variation across registers and varieties of English (see for example, Rosenbach (2014)). This change is demonstrated in the ongoing increase in the use of the *s*-genitive at the expense of the *of*-genitive. The increasing use of the *s*-genitive is ascribed to its extension to non-human possessors in present-day English, but there is less agreement on whether this extension is motivated by linguistic factors or stylistic factors. Explanations related to linguistic factors argue that the extension of the *s*-genitive could be due to either a loosening of the animacy constraint for particular subsets of nouns lower down on the animacy scale, or to a general spread of the form to inanimate noun possessors (Hinrichs & Szmrecsanyi, 2007). Explanations that invoke stylistic factors argue that the increasing use of the *s*-genitive with inanimate nouns is due to changing discourse practices (such as the trend towards colloquialisation or densification) (Hinrichs & Szmrecsanyi, 2007).

Cross-varietal investigations reveal that the patterns of genitive alternation are different across varieties of English. For example, Heller, Szmrecsanyi, and Grafmiller (2017b) show that the *s*-genitive occurs in higher proportions in first-language varieties of English than in second-language varieties and attribute this to a preference for the more transparent and analytic construction in second-language varieties. However, Rosenbach (2017) argues that in second-language varieties, substrate influence might also play a (potentially stronger) role and that this could either favour or disfavour the use of the *s*-genitive. As already discussed in Section 1.1, the differences between varieties of English result from processes of language change that are uniquely shaped by the different contexts in which varieties have developed, in which the complexity of factors identified by Hundt (2007) play uniquely configured roles. Given the above, the question that arises is: What is the interaction between broader trends of language change in written language, register effects, and specific language change in different varieties of English that are subject to different contextual forces, and consequently different opportunities for and constraints on innovation and shifting usage patterns and dissemination?

Attempting to generalise from this variability, Schneider (2007) proposes the Dynamic Model to account for the development of different varieties of English that share their origins in British colonisation activities. In the model, PCEs emerge as a result of the contact between the settler and indigenous populations involved in the colonisation process. Variation across PCEs is due to both the development of innovative features and usage patterns, and their selection and conventionalisation within the different contexts in which the varieties develop – but also to the different varieties' responses to broader trends of change.

The Dynamic Model conceptualises the development of PCEs across five consecutive stages, ultimately culminating in a new variety of English with hybrid roots and new linguistic norms (Schneider, 2007). At each of these five stages, four parameters can be observed:

(1) Extralinguistic factors, like historical events and the political situation, result in (2) characteristic identity constructions on the sides of the parties involved. These, in turn, manifest themselves in (3) sociolinguistic determinants of the contact setting (conditions of language contact, language use, and language attitudes), which, consequently, cause specific (4) structural effects to emerge in the form(s) of the language variety/-ies involved. (Schneider, 2007, p. 31)

In the first phase, Foundation, English is brought to a new territory through of group of English settlers who normally come from different regional backgrounds. The contact between these settlers in the new territory results in a levelling off of the different English dialects present through the process of koinéisation, and a relatively homogenous new variety, the settler (STL) strand, emerges (Schneider, 2007). Some toponymic borrowing does occur through contact between the settlers and the indigenous population, which is largely restricted at this stage (Schneider, 2007). While the newly emerged variety stabilises during Phase 2, Exonormative Stabilisation, the settler population views itself as an outpost of Britain and the linguistic orientation at this stage is towards the colonial centre. However, increased contact among the settler and indigenous populations results in the spoken form of the STL stand moving towards a local form of English, with settlers also starting to adopt indigenous words (Schneider, 2007). Within the indigenous population, increased contact with the settler community means that bilingualism starts to spread, marking the emergence of the indigenous strand (IDG strand). This represents the start of structural nativisation, in which one population group shifts to the language of the other, invariably resulting in some transfer of phonology and structure (Schneider, 2007). This process acts as a catalyst for the movement of the variety into Phase 3 of the model, Nativisation.

In Phase 3, the settler and indigenous populations realise that they are both permanent residents, and the cultural and linguistic gap between the two populations starts to narrow (Schneider, 2007). The use of English in important and formal functions places pressure on the indigenous group to accommodate to English, leading to widespread second-language acquisition. The heaviest restructuring of the English language occurs during this stage, mostly in innovations and structural nativisations that spread from the IDG to the STL strand. A new variety that starts to form its own linguistic norms beings to emerge; however, the acceptance of this norm is disputed by some members of the settler community while other STL-strand users accommodate to the features of the local variety. During Phase 4, Endonormative Stabilisation, following several important events the settler population severs ties with Britain. This results in the emergence of a form of unity among the settler and indigenous populations, in which the cultural and linguistic differences between them disappear (Schneider, 2007). A local identity therefore develops, alongside the acceptance of a local linguistic norm in both spoken and formal written usage (which is

codified in dictionaries, style guides and grammars). However, in the early parts of Phase 4 different norm orientations between conservative and innovative settler users persist, giving rise to tensions between the written standard norms of the imperial centre and the new locally codified norms. This tension is usually felt more strongly by educated users, who find themselves having to choose between the new local norm and the external norm. Over time this tension dissipates and the local norm becomes accepted as the standard to aspire to, leading to the endonormative stabilisation of the new variety (Schneider, 2007, p. 50). Having consolidated its national unity and own linguistic norms, the new variety continues to change through the normal processes of language change, and the variety moves into Phase 5 of the model, Differentiation, in which regional and social dialects of the new variety emerge.

Normativity is thus central to the Dynamic Model, because as varieties progress through the various stages, they gradually move from an exonormative orientation to an endonormative orientation in which a new linguistic norm reflective of the local community gains currency within the community and eventually comes to be accepted. In especially the third and fourth stages of the model, competition arises between different norms, including the competition between written standard and spoken vernacular norms, and overt prestige and covert solidarity norms (Schneider, 2007). Competition also arises between two kinds of overt norms in Phase 4: those that reflect the external, metropolitan standard, and those that reflect the local codified norm. These different competing forces raise challenges in deciding when an observed feature of language use is an innovation and when it is simply an error (Bamgbose, 1998), which is particularly salient in the development of the native and non-native varieties of World Englishes, and especially in their movement from Phase 3 of the model into Phase 4.

Van Rooy (2011) argues that errors can be distinguished from conventionalised innovations on the basis of two criteria: grammatical systematicity and acceptability. With regard to the latter, the acceptance of a new feature as a conventionalised innovation is rooted both in its widespread use within the community (or increased communal patterns of usage) and its acceptability by norm-giving authorities in the local community, such as teachers and editors, which implies convergence between overt and covert norm orientations.

Studies of language variation and change in World Englishes tend to downplay the influence of written language in these processes, with most researchers implicitly or explicitly departing from the principle that linguistic change has its origin in spoken language. The work of editors is seen as ‘interfering’ with these natural processes of change (Kruger, forthcoming; Kruger & Van Rooy, 2017). Because editorial work within the media and publishing houses serves a gatekeeper role in enforcing selected usage practices it is viewed as having a conservative influence on the language of written texts (Peters, 2006). Editorial intervention is thus framed as a factor that gets in the way of language change, rather than as a factor that forms part of the process.

However, researchers are increasingly noting the important role of written language in processes of language change more generally, and in varieties of English specifically. For example, Biber and Gray (2011) argue that innovative linguistic usage can emerge out of the communicative demands of written language, suggesting that written language forms an intrinsic part of contemporary processes of language change, and may be both a source of innovations and new usage patterns and an important mechanism in the dissemination and ultimate conventionalisation of these usages. In the context of World Englishes, the production of written texts plays an important role in affirming and propagating the conventionalised status of new features in a variety, which opens up the opportunity for editorial intervention to form part of these processes (Kruger & Van Rooy, 2017).

Cameron (2012) argues that the normative practices of editors should be included in studies of language change, because this forms part of the answer to important questions about the role of prescriptivism in language change. This is a particularly important point for this study, especially in light of Croft's (2000) assertion that studies of language change must account for the mechanisms that act as forces for change, but also for those mechanisms that act as forces for stability. In this regard, editorial practices that block the acceptance of or remove innovative usages in written texts may be viewed as mechanisms that act as forces for stability. However, as already discussed in Section 1.1, they may also act as forces that allow innovative features into written texts, legitimising their use and facilitating their further dissemination (Kruger & Van Rooy, 2017).

It seems clear that editing may provide important insights into patterns of conventionalisation and normativity in varieties of English. However, most existing studies on editing have focused on delimiting the types of editorial tasks specific to different contexts and have not focused on the actual changes that editors make to texts, nor what the consequences of these changes are. There is therefore very little evidence, in both the fields of editing and World Englishes, for what editors actually do when amending texts and how these changes influence the language of texts. Interpreted within the context of World Englishes and the development of different varieties of English, it is therefore not clear to what extent editorial practices influence the prevalence and dissemination of innovative usages in different registers in the different varieties of English, and whether this intervention acts as a mechanism that inhibits or aids conventionalisation and legitimisation.

It is known from existing research that the nature of editorial work is highly variable, and while some degree of editorial intervention is a given for all published texts, this intervention varies in scope (ranging on a continuum from light copyediting to substantive editing of content, style and structure that sometimes includes rewriting), and is largely determined by the type of text being published and the context of its production (Kruger, 2017; Law, 2011, 2014). Editorial tasks and the dimensions of editorial work are typically classified into two broad categories: copyediting and substantive editing. Copyediting comprises tasks that are done to bring a document into conformance with pre-set rules, including rules related to

spelling, punctuation, syntax and idiom, and good usage, as sanctioned in authoritative sources such as house styles, dictionaries, style manuals and usage guides (Mossop, 2014). Copyediting also ensures consistency in the application of these rules as well as in the treatment of matters such as terminology and the position, numbering and appearance of headings (Mossop, 2014). As such, copyediting is line-by-line, micro-level work that is normative in orientation, with editors enforcing selected usage practices (Peters, 2006).

Substantive editing is much more interventionist than copyediting because it aims to improve the overall scope, level, length, presentation and organisation of a text and is therefore concerned with the communicative success of the text (Butcher, Drake, & Leach, 2006). Although substantive editing does not invoke the kinds of prescriptive rules that form the basis of copyediting, it still has a normative dimension because of the normative expectations associated with register and audience (Kruger & Van Rooy, 2017). For example, stylistic editing, one sub-type of substantive editing, is concerned with ensuring the readability and clarity of a text by tailoring the language to the audience and their use of the text, and smoothing the language to ensure it is suitable to the text type (Mossop, 2014). This may be a point at which editorial intervention invokes covert norms, since it not only forces editors to think about the ways in which people use language but may also be an environment in which editors are more likely to impose their own stylistic and usage preferences. Content editing, another sub-type of substantive editing, is concerned with the coverage of the topic, and may sometimes necessitate additions to the text (Mossop, 2014). In this case the editor is required to write or rewrite sections of text. Thus content editing is also a kind of intervention in which editors' own (covert) normative preferences and stylistic and usage idiosyncrasies may find their way into texts.

Given the variability of editorial intervention in different contexts and the potential role of editorial intervention in the processes of ongoing change in published written registers and across different varieties of English, two questions arise, which form the impetus for this study:

- First, what kinds of norms do editors invoke and apply in their editorial practice? Here, norms are understood to include both overt norms (the norm-giving authorities that editors use to guide their editorial choices) and covert norms (the communal norms acquired through editors' personal experiences and interactions, as shaped by their sociolinguistic background and evident in the changes that they actually make to texts).
- Second, what role does editorial intervention play in accelerating or inhibiting language change in the different written registers in different varieties of English, both in terms of features specific to a variety and the more general processes of language change?

As already mentioned in Section 1.1, this study focuses on two varieties of English that offer a meaningful comparison to investigate these questions: AusE and SAfE. AusE represents a variety of English that is

entering Phase 5 of the Dynamic Model and is widely considered to be a classic Inner Circle variety in which the STL strand has prevailed (Schneider, 2007). Furthermore, AusE has achieved endonormative stabilisation, and is accepted as one of the major global reference varieties of English. The variety has been codified in several national dictionaries and its endonormative orientation has been promoted in other sources with the publication of various AusE style guides and usage guides (Delbridge, 1999; Leitner, 2004). The language profiles of AusE editors probably do not differ all that much from the AusE authors that they edit, which suggests that the covert norms of the two groups of language users are in all likelihood closely aligned, and that these norms are also closely aligned to the overt, codified norms of AusE. Furthermore, AusE has been shown to be particularly receptive to the features associated with colloquialisation, and is described as well known for its pervasive and persistent informality (Collins, 2014). In terms of the corpus analysis presented in this study, the *s*-genitive is the less formal of the two genitive constructions (Grafmiller, 2014), and the inclusion of AusE offers a good opportunity to investigate the extent to which AusE participates in this change and the extent to which the use of this less formal variant enjoys acceptance among editors in a range of written registers known to vary in terms of their formality.

South Africa, on the other hand, presents “a complex sociolinguistic constellation” in which a comparatively high number of distinct, compartmentalised language communities have interacted with each other at different points in time and under varying social circumstances (Schneider, 2007, p. 173), giving rise to a complex internal variation structure with a number of sub-varieties developing in contact with each other – such as the native sub-variety White South African English (henceforth WSAfE) which represents the STL strand, and the indigenised second-language sub-varieties Afrikaans English (henceforth AfrE) and Black South African English (henceforth BSAfE). As a consequence, English as used in South Africa is difficult to classify as an Inner or Outer Circle variety because it qualifies for both categories. In addition, while there is some evidence to suggest that SAfE is slowly moving towards endonormativity, the variety has not enjoyed the same level of codification as AusE and in the absence of clear codified norms for South African usage, the set of overt norms governing the use of English appears to be an exonormative one, with the orientation mostly towards British norms.

The complex sociolinguistic constellation of the different sub-varieties in South Africa also gives rise to complex sociolinguistic profiles for editors of English in the country. Given the importance of English in high functions and the fact that functional bilingualism is a key sociolinguistic feature in South Africa (Coetzee-Van Rooy, 2013, 2014), editors of English in the country are usually users of WSAfE or AfrE. Evidence for this can be found in a review of the South African Translators’ Institute’s (SATI) database of English editors, which shows that native users of English constitute approximately a quarter and Afrikaans–English bilinguals just less than half of this group (SATI, 2015). Furthermore, only a handful of this group of editors reports one of the nine official African languages as their first language, while the remaining

editors report a European language as their first language. This means that editors in South Africa do not necessarily have the same linguistic backgrounds (and therefore covert norms) as the authors that they edit, which adds a layer to the already complex notion of normativity. In terms of the feature selected in this study, South Africa offers an interesting case to investigate how the broader trends of change play out in this variety. Since indigenised second-language varieties show a tendency towards formality (Van Rooy, Terblanche, Haase, & Schmied, 2010), and since BSAfE (an indigenised second-language variety) is the most widely-used sub-variety in South Africa, it seems likely that the choice of genitive construction in this country would be towards to more formal of the two constructions.

The selection of the genitive alternation as the grammatical feature investigated in this study is rooted in the fact that there is evidence of ongoing change in this feature, and that it is an example of a grammatical feature whose use is linked to the broader trends of language change, colloquialisation and densification. Furthermore, it is known that the use of the feature is influenced by register, and variability across varieties of English has also been documented. In the two varieties selected (for which genitive alternation has not been widely investigated), the particular sociolinguistic settings are likely to condition the ways in which this feature is used (and treated by editors) in distinct ways. The two varieties therefore each present a different case for the potential interaction between overt and covert norm orientations, which allows for a meaningful comparative variationist investigation of the role of editorial intervention in the processes of ongoing language variation and change in World Englishes.

1.3 Research questions

The above discussion gives rise to the following research questions:

- 1) How do broad trends of language change, the different contextual forces of the varieties of English, register effects and editorial practice interact to influence language variation and change?
- 2) How have AusE and SAfE progressed along the stages of the Dynamic Model, and what are the implications of this for editorial work in AusE and SAfE?
 - 2.1) How do the language profiles of AusE and SAfE editors differ and how are these profiles similar or different to the language profiles of the authors that they edit, reflecting similarities and differences in covert norm orientation?
 - 2.2) What overt norm-providing instruments do editors of AusE and SAfE use to guide their editorial choices, and how and why does this differ in each of these two varieties?
 - 2.3) How do editors of AusE and SAfE respond to a feature known to be undergoing language change, namely the genitive alternation, and how are these responses shaped by the interplay between overt and covert norm orientation?

- 3) Based on the integration of theoretical and empirical work, how does editorial intervention interact with broader processes of language change, specific language change in the varieties of English and register effects to influence different opportunities for and constraints on the processes of dissemination and conventionalisation in a variety's progression towards endonormativity?

1.4 Research objectives

The objectives of this study are to:

1. Synthesise usage-based theories of language in a way that integrates cognitive and social perspectives for understanding language structure and variation with models of World Englishes and editorial practice in order to construct an innovative theoretical framework within which to investigate the interaction between the processes of ongoing language change and editorial work in different registers and different varieties of English.
2. Explore and compare AusE and SAfE's progression along the stages of the Dynamic Model by investigating and comparing:
 - 2.1. the sociolinguistic profiles of editors of AusE and SAfE,
 - 2.2. the norm-providing instruments used by editors of AusE and SAfE and,
 - 2.3. editors of AusE and SAfE's treatment of a feature for which there is known evidence of ongoing change and which is known to be linked to the broader processes of language change, namely the genitive alternation.
3. Synthesise the theoretical and empirical findings of the study to determine how editorial intervention in AusE and SAfE interacts with broader processes of language change, specific language change in the two varieties and register effects to influence different opportunities for and constraints on the processes of innovation and dissemination; and establish if and how this influences the progression of the varieties towards endonormativity.

1.5 Methodology

The methodology used to answer the research questions outlined above comprises three components: the construction of an innovative theoretical framework, a survey of the norm-providing instruments used by and sociolinguistic profiles of editors of English texts in Australia and South Africa, and a corpus-based investigation into how editors in the two varieties and across different published registers respond to the presence of a feature known to be undergoing change, namely the genitive alternation.

In order to answer research question 1, an innovative theoretical framework was developed that synthesises existing literature on theories of language change, research into World Englishes, register effects and editorial intervention in order to understand the interaction between processes of language variation and change and editorial work in different registers in different World Englishes contexts. The framework pulls

together various aspects of the theories and models in order to position the role of editors within a usage-based approach that integrates cognitive and social perspectives for understanding linguistic variation and change within the context of World Englishes.

In order to answer research questions 2.1 and 2.2, the first part of the empirical component of the study involved the design and distribution of a questionnaire among editors of English texts in Australia and South Africa. The purpose of the questionnaire was twofold: first, it gathered information on the demographic and language profiles of editors of English texts in Australia and South Africa; and second, it gathered information on the kinds of sources that these editors draw on to guide their editorial choices as well as the norm-orientation of these sources. These data are used, firstly, to investigate if and how the sociolinguistic profiles of editors of English in Australia and South Africa differ, providing insights into the covert norms (based on editors' own linguistic experiences) shaping editorial decisions. Secondly, the data from the questionnaire are used to determine the kinds of sources editors in the two varieties draw on to guide their editorial choices, and, more importantly, to determine the norm-orientation of these overt norm-providing sources in order to establish if these reflect norms internal to the variety in which they are used, or if they reflect a norm external to the variety.

The second empirical component of the study involved an investigation into how editors of English texts in the two varieties respond to the presence of a particular linguistic feature associated with ongoing change in the texts that they edit, and how this intervention differs across different registers in the two varieties. As discussed in Section 1.2, the selected feature is the genitive alternation. The second part of the study therefore seeks to answer research question 2.3. Because this study is interested in actual patterns of language use and the choices that editors make when editing texts (rather than relying on anecdotal or intuitive observations of what these choices might be), a corpus-based approach is used. To this end, a parallel, register-differentiated corpus of aligned unedited texts and their edited counterparts, representing each of the two varieties, was constructed. The main corpus is divided into two primary sub-corpora representing the two main varieties under investigation, namely AusE and SAfE, with the latter corpus differentiated into AfrE, BSAfE and WSAfE components. For each variety, five printed written registers are represented: academic, creative, instructional, popular and reportage. Each register is further subdivided to include unedited texts and their edited counterparts, which are aligned. This design allows for a parallel investigation into the occurrence of genitive constructions across the unedited texts and their edited counterparts for each register within each variety, while also allowing for a comparative investigation of the editorial role across the two varieties.

The corpus analysis of the genitive alternation is carried out in three stages. First, an overall analysis of the proportions of the two genitive constructions is presented for the edited, published sub-corpus across the sub-varieties and registers. This analysis provides information on the overall distributional patterns across the different registers and sub-varieties and helps to position these two varieties alongside other varieties

of English in terms of their openness to this feature associated with the two broad trends of change. Subsequently, the effects of editorial intervention on the overall proportional frequencies of the two constructions across registers and sub-varieties are investigated. Drawing on a combination of distributional analyses, statistical evaluation tools, and qualitative investigation, this analyses explores how and why editorial intervention shifts the frequencies of the two constructions across different registers and varieties. Lastly, following the use of variationist approaches to investigate variation in (genitive alternation in) World Englishes, a variationist method is used to explore the interaction between a number of language-external factors (including editorial intervention) and linguistic factors in conditioning the choice of construction. The analysis draws primarily on quantitative analyses, including descriptive statistics and state-of-the-art statistical evaluation techniques (such as random forests and conditional inference trees).

The data from the survey and corpus analysis are synthesised with the theoretical framework developed in order to answer research question 3: to determine how editorial intervention interacts with broader processes of language change, specific language change in the two varieties and register effects to influence different opportunities for and constraints on the processes of innovation and dissemination, and to establish if and how this influences different varieties' progress along the stages of the Dynamic Model.

1.6 Chapter outline

Chapter 2: Theoretical framework: Editorial intervention, language variation, and language change in written varieties of English

Chapter 2 discusses the literature relevant to usage-based theories of language use and language change, and research into World Englishes, register variation and editing. Chapter 2 synthesises these theories and models in order to develop a theoretical framework within which to investigate the interaction between the processes of language variation and change and editorial work in different registers in different World Englishes contexts.

Chapter 3: Varieties, registers and grammatical feature investigated

Chapter 3 sets out a discussion of the two varieties, five registers and grammatical feature selected for the empirical investigation of the role of editorial intervention in influencing different opportunities for and constraints on the processes of conventionalisation and dissemination in different varieties' progression towards endonormativity. In particular, Chapter 3 focuses on the historical development and current status of AusE and SAfE in terms of Schneider's (2007) Dynamic Model. This chapter also unpacks the five registers investigated in this study and explores the differences and similarities among them in terms of their specific situational and communicative characteristics and the varying nature and scope of editorial intervention across the registers. Thereafter, the grammatical feature used as a case study for this thesis, namely the genitive alternation, is discussed. This includes a discussion of the two variant constructions as

Chapter 1: Introduction, contextualisation and problem statement

well as an exposition of the existing body of research on the ongoing change in genitive alternation. Chapter 3 concludes by drawing the different sections together in order to provide a motivation for the selection of the varieties, registers and grammatical feature investigated in this study.

Chapter 4: Methodology: Combining survey and corpus research to investigate the effects of editing in written varieties of English

Chapter 4 outlines the methodology used for the empirical component of the study and discusses the processes involved in the design of the questionnaire and corpus, the sampling procedure followed, the data-collection procedures as well as the various statistical methods used in the analyses of the data collected from the questionnaire and corpus.

Chapter 5: Presentation and analysis of the empirical data: Evidence for the role of editorial intervention in ongoing language change in AusE and SAfE and across written published registers

This chapter presents an analysis of the data from the questionnaire and corpus-based investigation. The results of the analysis of the questionnaire data are presented and discussed first, followed by the presentation of the results from the corpus analysis, and a critical discussion of these results. The chapter concludes by discussing the overall findings of the empirical investigation.

Chapter 6: Synthesis of findings, recommendations for future research and conclusion

Chapter 6 synthesises the findings of the empirical component of the study within the theoretical framework developed in Chapter 2. Recommendations for future research are made before concluding remarks are presented.

Chapter 2

Theoretical framework: Editorial intervention, language variation and language change in written varieties of English

2.1 Introduction

The global spread of English, initially as a consequence of British colonisation activities and more recently associated with globalisation, means that English is now used in many different societies and contexts and for different purposes. The dynamic and continuous expansion of English has given rise to different varieties of English, typically studied within the World Englishes paradigm. Most research on World Englishes has been oriented towards documenting the properties and functions of the different varieties within their contexts of use, pointing out differences as well as similarities among them. This is demonstrated in the various frameworks and models that have been developed to categorise the different varieties and to describe their functions and development, such as Kachru's (1985) Three Circles Model and Schneider's (2007) Dynamic Model of PCEs. The ongoing diversification of Englishes across the globe has also given rise to comparative analyses of and deepened theoretical reflection on the processes that give rise to and drive the differentiation among the varieties.

One such process, namely the development of local norms and the acceptance of such norms by local norm-providing authorities (such as editors, the media and publishing institutions) has not been the topic of much empirical (or theoretical) work, despite the importance of norms for the development of the different varieties. Editors of published written texts may be regarded as one group of language users who play an important role in reinforcing and establishing local norms within the contexts of different varieties of English (Kruger, forthcoming; Kruger & Van Rooy, 2017). Yet, despite their (potentially) important role in the development of local norms very little empirical work has been done on the influence of editors in the development of local norms within varieties of English, and only three studies, Kruger (forthcoming), Kruger and Van Rooy (2017) and Van Rooy and Kruger (2016), have explicitly investigated this role. As a result, not much is known about how editors form part of the processes that drive language variation and change in written registers in different varieties of English, and if they do, how their influence unfolds within the development of local norms within the different contexts in which varieties develop.

As discussed in Section 1.1, several factors have contributed to the dearth of research in this area. The first factor relates to the availability of linguistic-theoretical paradigms to offer meaningful systems for the systematic analysis of variable data in World Englishes (Van Rooy, 2010). Modern theoretical approaches to language, at least since the 1950s, tend to fall into two broad categories: approaches framed within the generative tradition, and approaches that fall into the family of usage-based views that reject this tradition

and which reflect converging ideas from different linguistic approaches (Geeraerts & Kristiansen, 2015; Van Rooy, 2010; Van Rooy & Kruger, 2015; Von Mengden & Croussé, 2014). Within the generative tradition, language is viewed as an idealised entity that is abstracted away from language use. This means that the generative approach to language does not offer the theoretical tools to investigate the variation that arises in language through use and which drives change, limiting the opportunity to investigate editing as a factor in language variation (see further discussion in Section 2.2). Recent new insights into the connection between language structure and language use, most notably within usage-based approaches and the integration of social and cognitive perspectives of language structure and use, have allowed linguists to re-examine the realities of language and have provided scholars of World Englishes with theoretical paradigms that help to explore, describe and explain the variation among and evolution of the world's Englishes (Van Rooy, 2010). Van Rooy (2010) has called for the integration of cognitive, usage-based approaches with variationist approaches in order to develop a framework that allows for the integration (and interaction) of social, cognitive and linguistic factors in models of language variation and change in World Englishes. This has opened up the opportunity to investigate the role of editorial intervention in the processes of ongoing change and in the development of local norms within the context of World Englishes.

A second factor that has contributed to the dearth of research on the role of editors in the development of (local) norms is the centrality of spoken language in studies of variation and change, both in English more broadly, and in World Englishes. Following Saussure, linguists have tended to dismiss written language as a derivative of spoken language. This privileges spoken language as the primary area of investigation and leads to the assumption (and widespread belief) that the source of linguistic change is spoken language. As a consequence, most studies of language variation and change in English invariably turn their attention to spoken language, ignoring the potential influence of written language in the processes of variation and change, and in the development of local norms. The implicit focus on speech is demonstrated in the consistent use of phrases such as 'speaker and hearer' to refer to interlocutors, or 'speech community' to refer to the community of language users, or 'utterances' to refer to the linguistic unit of analysis. More unequivocal examples of the primacy of speech (in especially World Englishes) can be found in various quotes by linguists, such as the following: "For contemporary studies, though, spoken data are *sine qua non*..." (Mesthrie & Bhatt, 2008, p. 2). Recently Biber and Gray (2011, 2012, 2013) have demonstrated the important role of written language in processes of ongoing change by showing that features associated with linguistic change may actually emerge in written language, and can be linked to the specific demands associated with language production in different written registers. Similarly, scholars such as Iwasaki (2015) and Kruger and Van Rooy (2017) have argued (and found evidence to show) that written language actually plays an important role in propagating the conventionalised status of features associated with ongoing change, implicating written language (and by implication, editors) in processes of language variation and change, and the development of local norms in different varieties of English.

The third factor is related to the second: because of the primacy of spoken language, where written language is investigated, there is a tendency to mention the potential influence of editors on the language of published

written texts in passing, without any real attempt to investigate what this influence actually entails. Furthermore, because editorial intervention is widely viewed as a primarily normative process, when its influence is mentioned, it is almost always described as an activity that blocks or removes the linguistic features under investigation, and is therefore framed as an obstacle to change rather than as an important part of the processes of change. Kruger and Van Rooy (2017), however, propose that editorial intervention actually plays a far more important role than previously assumed, and suggest that the acceptance of features associated with ongoing change by editors legitimises their use and provides important information about the conventionalised status of these features, while at the same time forming part of the process in which these features spread.

This thesis aligns itself with the views of Kruger and Van Rooy (2017) and aims to investigate the role of editorial intervention in the dissemination and conventionalisation of the features associated with ongoing change in different written published registers in different varieties of English. It is rooted in the intersection of broader contemporary theoretical linguistic approaches, drawing on and integrating usage-based approaches, theories of language change and research on World Englishes. Section 2.2 discusses the overarching linguistic theories which inform the study. First, in Section 2.2.1, an explanation of the role of cognitive factors in language use and variation is set out. This lays the foundation for Section 2.2.2, which discusses the role of social factors in language use, and focuses on how social factors contribute to language variation. The discussion of the interaction between cognitive and social factors in language structure and language use provides the scaffolding for understanding the processes of language change, which is crucial for situating the role of editors in these processes. Section 2.3 shifts the focus to theories of language change, and in particular to evolutionary theories of language change, in order to identify and explore the processes and mechanisms by which variation arises and change proceeds. Thereafter, in Section 2.4, models for the development and classification of World Englishes are discussed with reference to theories of language structure and language change. Section 2.5 turns its attention to another crucial factor that conditions both language variation and language change: register. Register studies, and in particular the variation in published written texts, forms the focus of this section, which lays the foundation for the discussion of editorial intervention that follows. The complexities associated with the many dimensions of norms and normativity will emerge from the discussions in these sections and therefore Section 2.6 pulls together the main threads of the previous sections by addressing the complexities of norms and normativity within a usage-based approach that integrates cognitive and social perspectives for understanding the role of editors in language structure and variation within the context of World Englishes.

The discussion presented in this chapter therefore seeks to answer the first research question of this thesis by synthesising theories of language change, research into World Englishes, register effects and editorial intervention in order to understand the interaction between the processes of language variation and change and editorial work in different registers in different World Englishes contexts. This provides the basis for the discussion that follows in Chapter 3, which sets out a discussion of the two varieties, different registers and linguistic feature selected for the empirical investigation of the role of editorial intervention in

influencing different opportunities for and constraints on the processes of conventionalisation and dissemination in different varieties' progression towards endonormativity.

2.2 Theoretical approaches to language structure and use

The complexity and systematicity of language have given rise to theories about whether it is used for thinking or communicating, whether it has evolved gradually or abruptly, whether its structure is innate or emerges through use, whether the cognitive processes that underlie its structure are specific to language or not, and how and why it changes (Bybee, 2010). This thesis adopts a usage-based approach to the investigation of language and integrates cognitive and social perspectives for understanding language structure and ongoing change within the World Englishes paradigm. In order to outline the approach used in this study it is first necessary to briefly unpack the foundational assumptions of the generative tradition³ in order to provide the necessary contextualisation for the selection of the usage-based approach.

Linguistic theories in the generative tradition argue that language should be viewed as an idealised entity that is abstracted from its actual use in particular contexts. Generativists frame this idealisation in terms of an ideal speaker-hearer in a completely homogenous speech community, who has perfect knowledge of her language, who is not affected by memory limitations or shifts of attention, and who does not make errors in applying this perfect knowledge of language in instances of use (Chomsky, 1965). This abstracts language away from the cognitive and social forces that give rise to the patterns of variation that occur in everyday language use; a view that is captured in the distinction and strict separation of knowledge of language (competence) and actual language use (performance) and which has its roots in the Saussurean distinction between *langue* (language structure) and *parole* (language use). In building a case for this distinction, generativists argue that investigations of language are complicated by what Chomsky (1965) calls 'Plato's problem' and the belief that the structure of language is far too complex for children to acquire from the imperfect and limited data they receive by exposure to language use. Generativists thus argue that there is an innate language acquisition device specific to humans and to language that guides the acquisition of language. Thus, an individual's knowledge of language, their competence, is innately given, is a property of the mind that grows and matures along an intrinsically determined course, and is triggered and shaped by the effects of the individual's environment (Chomsky, 1986). Within this framework, it is argued that linguistic theory should not focus on linguistic behaviour or the products of this behaviour (performance), but should rather investigate the state of mind, or innate language faculty, that gives rise to the behaviour (Chomsky, 1965, 1986).

The distinction between competence and performance is further captured in the Chomskyan notions of E-language and I-Language (Chomsky, 1986). E-language refers to an externalised form of language that is

³ The overview of the generative tradition that is given here is by no means comprehensive. Instead, the discussion focuses on the fundamental assumptions of generativism that are rejected by the usage-based approach.

characterised by its use by populations of users who share certain regularities in action or belief. This externalised language is governed by factors unrelated (and therefore external) to grammar, and is therefore understood independently of the properties of the mind. I-language, or internalised language, on the other hand, refers to the notion of structure as the most important component of the knowledge of language and frames this as an element of the language user's mind. The study of the knowledge of language pursued by generativists is therefore a study of this innate grammar, which is the theory of the I-language (Chomsky, 1986).

Although generativists, like cognitive linguists in the usage-based approach, view language as seated in the mind, they argue for the modularity of language. Pinker (1995) captures this modularity in the following statement:

Language is a complex, specialized skill, which develops in the child spontaneously, without conscious effort or formal instruction, is deployed without awareness of its underlying logic, is qualitatively the same in every individual, and is distinct from more general abilities to process information or behave intelligently. For these reasons some cognitive scientists have described language as a psychological faculty, a mental organ, neural system, and a computational model. (p. 18)

The argument for this modularity, together with the idealisation away from variation, means that generativists focus on language structure or competence and view this inherent knowledge as universal – giving rise to the notion of a Universal Grammar. It is this Universal Grammar that is given, shared by all humans, and which is seated in the language-specific component of the mind. Thus, the cognitive organisation of language, or knowledge of language, is separate from the rest of cognition and language is processed by domain-specific cognitive processes. The object of study in the generativist tradition is therefore language structure and the grammar that gives rise to this structure, which is taken to be divorced from performance, lexis, social usage, and the rest of cognition, and is framed as an autonomous, static system of grammatical principles that is rule governed. The consequence of this is the claim not only for the autonomy of the language faculty, but also for the autonomy of certain areas of linguistic description such as syntax and morphology, giving rise to the form-focused linguistic orientation that is characteristic of the approach (Geeraerts & Kristiansen, 2015, p. 368).

As a result of these views, it has been argued that the generative approach decontextualises the study of language, and abstracts important aspects of language, like meaning and function (and the lexicon as a major repository of meaning), context of use, and variation away from linguistic theorisation (Geeraerts & Kristiansen, 2015). The generative approach also cannot account for the fact that variation is an inherent feature of language, with linguists working in alternative paradigms calling the approach into question and casting doubt on central aspects of the theory such as autonomy, modularity, innateness, and the form-focused hierarchical account of syntactic structure (Schmid, 2012). Instead arguments have been made for the importance of such aspects as communication, cognition, social forces and processing for the development and organisation of grammar (Diessel, 2011). In this regard, cognitive usage-based linguists

such as Bybee (2010) and sociolinguists such as (Labov, 1994, 2001, 2010) state that because languages exhibit regularity in their shape and structure, while also exhibiting considerable variation, and because languages change, studies that seek to understand language should focus on the dynamic social, pragmatic and cognitive forces that give rise to its structure and variability, and which drive its change.

The usage-based approach thus seeks to reverse the decontextualising tendency of generativism, and argues for a dialectic relationship between language structure and language use (Geeraerts & Kristiansen, 2015). The term ‘usage-based’ was first coined in the late 1980s in Langacker's *Foundations of Cognitive Linguistics* (1987), in which he proposed the term to describe an approach to the study of language in which the mental representations of language are rooted in language use, and not in some innate language faculty. In doing so, Langacker unified the study of language use (performance) with that of language structure (competence). The term was thus initially used to highlight a theoretical and methodological distinction between cognitive linguistics and generative linguistics; however, Bybee (2013) states that although Langacker may be credited with coining the term, the approach actually has its roots in the earlier work of functional-typological linguists who proposed that grammar is created through the conventionalisation of commonly used discourse patterns – a view that was later enriched by the recognition of the cognitive processes that feed into linguistic structure. Thus contemporary conceptions of the approach, which represent a cluster of approaches including cognitive grammar, construction grammar, and functional grammar (Van Rooy, 2010), are characterised by the view that grammar is based on usage, and therefore usage data are important in the construction of linguistic theories and the description of language (Von Mengden & Coussé, 2014).

Usage is the communicative interaction between interlocutors when language is produced and comprehended, and the data gleaned from these usage experiences creates and shapes the cognitive representation of language (Bybee, 2013). Thus, language structure is informed by both cognitive processes and social factors. Usage events play a dual role in language: they result from and shape the linguistic system, forming a kind of feedback loop between structure and use (Kemmer & Barlow, 2000). Usage data are therefore central to shaping grammar, serving as input for language users, but also providing a window onto how grammar is structured. This informs several further points on which the usage-based approach differs from the generative approach.

In the first instance, the usage-based approach does away with the competence/performance dichotomy because it argues that linguistic structure emanates from usage events and that there is therefore a dialectic of usage and structure, competence and performance. In other words, an individual's knowledge of language is a result of their exposure to language. In this way, grammar is not only the knowledge repository employed in language use, but it is also the product of language use (Tummers, Heylen, & Geeraerts, 2005).

The recognition that experiences with language constantly create and affect knowledge of language means that linguistic structure is dynamic and constantly changing, or emergent (Bybee, 2010, 2013; Diessel,

2011, 2015). This view places the notion of variation as a central aspect of both language structure but also of language use. Since no two linguistic experiences or usage events are exactly the same and because no two individuals' previous experiences are the same, the cognitive and contextual aspects of the usage event interact to define and continuously redefine the language system (Tummers et al., 2005). Grammar therefore emerges through and is shaped by language use and is not static, given or “pre-ordained by God, by gene, by school curriculum, or by other human policy” (Ellis, 2011, p. 655).

Lastly, usage-based linguists who work from within a cognitive approach argue that the mental representation and organisation of language are grounded in domain-general rather than domain-specific cognitive processes that are applied during usage events (Bybee & Beckner, 2015; Diessel, 2011). Thus they take as their null hypothesis that language is an extension of other cognitive domains and argue that language structure is shaped by the same general cognitive processes involved in non-linguistic cognitive activities, such as visual and auditory perception or reasoning (Bybee & Beckner, 2015; Diessel, 2011).

However, individual linguistic encounters take place within the context of social interaction, and are therefore influenced by the various social, cultural and pragmatic factors of those interactions. This may seem obvious, but it is only very recently that various linguists working within the broader usage-based paradigm have called for the integration of cognitive, social, cultural and pragmatic factors in order to allow for the study of the total linguistic fact (Dąbrowska, 2015, 2016; Geeraerts & Kristiansen, 2015; Hans-Jörg, 2016; Harder, 2010, 2015). Backus and Spotti (2012), Clark and Trousdale (2010), and Von Mengden and Croussé (2014), for example, point out that despite the fact that modern sociolinguistic and cognitive linguistic approaches share an interest in usage data, linguists in each of the two approaches have not really been interested in the work of the other. Dąbrowska (2016) summarises this as follows:

At the risk of stating the obvious, human beings are social creatures through and through, and social influences shape virtually every aspect of our lives from tool use and morality to parenting behaviour and communication. Language is not only an instrument for social interaction; it is also a system that emerges through interaction, and we cannot hope to understand its structure without considering both cognitive and social factors and their interactions. (pp. 486–486)

Cognitive approaches and sociolinguistic approaches (especially variationist sociolinguistics) share certain assumptions about the importance of language use, and therefore offer complimentary approaches (and methods) for investigating the variation that arises in language use and which drives language change: both view synchronic and diachronic perspectives as being related and argue that a proper theory of language needs to relate the two perspectives; both seek to investigate the impact of language use on knowledge of language and therefore both focus on usage data as the primary research data; and lastly, both view meaning as an important component of language (Backus, 2015).

The theoretical approach in this study is thus rooted in the intersection of cognitive and social perspectives for understanding the cognitively and socially motivated variation that arises in language use and which

drives the ongoing diversification of and change in World Englishes. The following sections set out the primary viewpoints within each of these approaches that are relevant to the questions that inform this study. Section 2.2.1 outlines important aspects associated with the cognitive usage-based approach that can be used to explain variability in language use. Following this, a discussion of the importance of social factors in influencing language use is presented in Section 2.2.2, with specific focus on variationist sociolinguistics.

2.2.1 The cognitive organisation of language

To account for the systematic conventionality of language and the fact that it appears to follow systematic paths of change, cognitive usage-based linguistics proposes a framework for the emergence of language structure through language use that incorporates exemplar models, the interaction of domain-general cognitive processes, networks of rich memory representations, and the effects of frequency (Bybee & Beckner, 2015).

Any approach that seeks to study language must define what it considers to be the basic linguistic unit. Because usage-based approaches claim that linguistic experiences create and impact the cognitive representations of language, it should come as no surprise that explanations of what these linguistic experiences entail are central in definitions of the basic linguistic unit:

The constructions that speakers know are directly associated with phonological, morphological, and syntactic properties, along with conventionalised meanings, possible variants, and the social contexts in which you are likely to use and hear them. In simplest terms, your knowledge of a construction is the sum total of your experience with [it]. (Hilpert, 2014, p. 2).

Since linguistic structure is grounded in the semantics and pragmatics it encodes in concrete instances of language use, studies of linguistic structure within the usage-based approach cannot, and do not, exclude meaning and context of use (Behrens, 2009; Evans, 2017; Geeraerts & Kristiansen, 2015). In this regard, usage-based linguists draw on the notion of a construction (originally described in construction grammar) to foreground this relationship and state that constructions are conventionalised form-meaning pairings (Bybee, 2010). In practical terms, these pairings are manifest or observable in patterns of morphemes or words that are used so often that they are accessed together as processing units or chunks and therefore become conventionalised (Bybee, 2013). These conventionalised units may, over time and in a variety of ways, develop particular pragmatic uses that give rise to special meanings or which lead to idiosyncrasies in their form.

The central status of meaning in this pairing means that constructions cover the full range of linguistic levels and include the integration of lexical items with grammatical structure (Bybee & Beckner, 2014; Evans, 2017; Geeraerts & Kristiansen, 2015). These units are framed at different levels of representation, and are represented in a lexicon–grammar continuum. At one end of the continuum these constructions may be highly specified and concrete in terms of their form and function (such as idiomatic expressions or lexical

items), while at the other end, they may be highly schematic and abstract (such as sentence-level constructions) (Bybee, 2013, p. 57).

To account for how experiences with concrete linguistic tokens constantly create and impact on cognitive representations of language, it is necessary to explore the processes through which these linguistic units are structured in the mind. The most important domain-general cognitive process invoked within the usage-based paradigm to explain the cognitive organisation of language is categorisation, which is the process of classifying or grouping experienced entities that are perceived as instances of the same kind, into categories (Bybee & Beckner, 2015). The storing and representation of these categories as cognitive structures has been described in terms of two models: prototype models and exemplar models. Within the prototype approach, categories are constructed on the basis of highly abstract representations of linguistic instances in which the aggregate properties that are characteristic of the concept are represented in a single prototype (Divjak & Arppe, 2013). Thus, the process of categorisation is taken to be one where the encountered linguistic instance is compared to the prototypes (abstractions) that are available in memory and are categorised under the most similar prototype representation (Divjak & Arppe, 2013). Exemplar theory, on the other hand, argues that categories are constructed on the basis of detailed memory traces of all individually encountered linguistic instances (Bybee & Beckner, 2015). Because every piece of information encountered in a linguistic experience leaves a memory trace, tokens that are similar or identical reinforce each other, and over time clusters of overlapping tokens, or exemplars, are formed (Diessel, 2016).

There is considerable literature on the use of exemplar models versus prototype models in studies of language structure (Divjak & Arppe, 2013). It is not within the scope of this thesis to discuss this literature, and therefore this thesis adopts the model deemed most relevant to this study, namely exemplar models. Many usage-based linguists, such as Bybee (2007a, 2007b, 2010, 2013), Bybee and Beckner (2014, 2015), and Diessel (2011, 2015, 2016) draw on exemplar models; however, it may also be possible that exemplar categorisation may display prototype effects. In other words, the mental representations of linguistic categories are constructed on the basis of exemplar models, but these categories are subject to prototype effects, most notably through frequency and category membership centrality (Bybee & Beckner, 2014, 2015).

Exemplar models provide a way in which to explain how all instances of language use continuously shape language structure because exemplars are formed on the basis of the similarity matching that occurs between currently-experienced and already-experienced linguistic tokens (Diessel & Hilpert, 2016). Exemplars are thus used to interpret and produce utterances and form the basis on which novel tokens are licensed and conventionalised tokens are recognised (Hruschka et al., 2009). Furthermore, exemplars contain all the information that a language user can potentially perceive during a linguistic event, including procedural knowledge about how to use a particular token, as well as information related to its frequency and statistical patterning, and may thus display prototype effects where certain exemplars may be described

as better members or more central members of a category than others (Bybee, 2010; Bybee & Beckner, 2015).

The ability to remember detailed information about particular tokens is made possible by the fact that the human mind is capable of storing rich memory representations, but these representations are always generalisations over experience and are therefore partially abstract (Goldberg, 2013). As a result, the categories that emerge in memory are not as detailed or comprehensive as the linguistic event in which they were interpreted or produced (Langacker, 2008). The generalisation and abstraction of concrete tokens leads to pattern forming and recognition, giving rise to schemas (a kind of grammatical template, or abstract pattern of concrete tokens) (Diessel, 2011). Schemas can be described on a cline of complexity and vary across a continuum of generality and abstractness (Diessel, 2015). Abstractness is related to the degree to which a construction's slots are schematic, or not specified. Thus, highly abstract constructions are those with a high degree of unspecified slots, while more concrete constructions are less abstract because they contain highly specified slots.

If during a usage event a linguistic token is perceived to be identical to an existing exemplar then it is mapped on to it, reinforcing its mental representation and thereby strengthening it, leading to its cognitive entrenchment (or its cognitive conventionalisation) and making it a stronger or 'better' member of a category (Bybee, 2010, p. 19). This increases the likelihood that it will be selected the next time an exemplar from the category is used. Alternatively, if some aspect of a token's form or function is perceived as similar to an existing exemplar, but not identical to it (for example, if it offers a different way of saying the same thing), then it is stored near to it (as well as to other exemplars with which it shares similarities), effectively establishing a new exemplar and extending the category(-ies) it occupies (Bybee, 2010). This process of similarity matching, or categorisation, together with the human capacity to store rich representations of linguistic experiences and the abstraction of linguistic patterns from these experiences, gives rise to a network of gradient categorical representations, or schemas, of language in the mind.

However, categorisation (together with entrenchment and schematisation) is only possible if linguistic units are repeated in usage events, and if humans are able to keep track of that repetition. The importance of repetition is highlighted in the widely accepted notion within usage-based approaches that conventionalisation through repetition creates grammar (Bybee, 2007b). Thus, if linguists want to investigate the conventionality of language, but also its variability, then they need to look to the effects of frequency in shaping language. Thus, frequency plays a crucial role in language structure because if there are quantitative changes in a user's language, then the language system itself changes (Behrens, 2009; Pfänder & Behrens, 2016).

Research into frequency, typically carried out using corpus-based approaches, demonstrates that its effects play out at different levels of linguistic representation, but also that there are different kinds of frequency (Ambridge, Kidd, Rowland, & Theakston, 2015). The effects of frequency may be observed at the level of

concrete items as well as at the level of abstract categories, and therefore frequency affects the full range of linguistic units (Ambridge et al., 2015; Behrens & Pfänder, 2016; Bybee, 2010; Diessel, 2016; Pfänder & Behrens, 2016). Different kinds of frequency effects may also be distinguished, including token frequency, type frequency, absolute frequency and relative frequency. Type and token frequency are the most relevant to this study and are discussed in further detail.

Token frequency refers to the number of times a particular linguistic unit occurs in a given text (such as, for example, the number of times the word *lie* appears in running text) (Behrens, 2009). Generally speaking, token frequency is taken as a measure of the degree to which a particular item is entrenched in a user's mental representation (Ambridge et al., 2015; Evans, 2017; Langacker, 2008). The higher the token frequency, the more entrenched the token is, the more conventionalised it is considered to be and the more likely it is to be selected in language use. The entrenchment of a unit due to its higher token frequency thus contributes to its stability (Pfänder & Behrens, 2016). Furthermore, because repetition leads to entrenchment, highly frequent tokens are also considered to be more easily accessible than lower frequency items and they therefore tend to resist the kinds of reformations associated with language change, such as analogy (Pfänder & Behrens, 2016). (The kinds of reformations associated with language change are discussed in Section 2.3). High token frequency thus also has a conserving effect (Bybee, 2007b; Pfänder & Behrens, 2016).

A further consequence of high token frequency is that of chunking. Chunking occurs when sequences of linguistic units are combined into more complex chunks as a result of frequently occurring together (Bybee, 2010). In other words, chunks form the sequential relations between units that have co-occurred so often, they are stored and accessed as a single unit. One consequence of this is that it may lead to the opacity of the chunked unit's internal structure (Bybee, 2010; Pfänder & Behrens, 2016). Because these units may be of varying levels of complexity, they provide "the cognitive basis for the morphosyntax and hierarchical organisation of language" (Bybee, 2010, p. 34). Frequency, or repetition, is therefore a particularly important mechanism that enables chunking, which, together with categorisation, results in a powerful mechanism by which conventional sequences that are used often become established as the appropriate or conventional way to say something.

Type frequency is related to patterns and constructions, such as the English Past Tense Construction VERB-*ed*, or the availability of two constructions to express genitive relations in English, the *s*- [NP₁'s NP₂] and *of*-genitive construction [*the* NP₂ *of* NP₁] (Pfänder & Behrens, 2016). Type frequency refers to the number of items that exemplify the pattern or the number of distinct items that can occur in the open slots of the construction (Bybee, 2007b). For example, more verbs occur with the VERB-*ed* construction than with the vowel-change XX-*ew* construction (*flew*, *blew*), and therefore the VERB-*ed* construction has a higher type frequency and is the more productive of the two. Since constructions with a higher type frequency tend to attract new items more easily than constructions with a lower type frequency, type frequency may act as a trigger for change (Pfänder & Behrens, 2016).

Frequency effects therefore appear to be contradictory, but these contradictory effects can be reconciled if a distinction is made between token and type frequency (in terms of what is counted, but also what the effects of each are), and if it is assumed that frequency interacts with other processing factors like recency or priming effects, and salience (Pfänder & Behrens, 2016).

Usage-based approaches therefore frame language as a phenomenon that is both structured and variable, and grammar as emergent and structured by the interaction of an individual's exposure to actual language use in context and domain-general cognitive processes (Bybee, 2010). Thus, the full circle view of language within the cognitive usage-based approach is that there is an intimate relationship between linguistic structure and language use, which includes comprehension and production in context (Von Mengden & Coussé, 2014). Of critical importance to this thesis are the effects of frequency and the processes by which the differential exposure of an individual to (different) constructions leads, on the one hand, to the cognitive conventionalisation of a construction, and on the other, to a form of cognitively motivated variation. In this way, language users' knowledge of (the conventions of) language (including editors) is shaped by their differential exposure to language in the language communities in which they interact. This gives rise to variation within individual language users' cognitive representations of language, subject to the different contexts in which they experience language, as well as variation across individual language users' cognitive representations of language.

Because the usage-based framework outlined above is primarily concerned with explaining how the usage events that language users encounter are processed by domain-general cognitive processes and how repetition and these cognitive processes interact to give rise to the mental representation of language structure in the individual user's mind, the discussion so far has focused on the cognitive organisation of language from the perspective of the individual. But, language is used by people in varying social and pragmatic contexts to communicate, all of which influence the particular selections that users make. Dąbrowska (2015) thus argues that to really explain language structure (and the variation that arises in this structure), it is important to distinguish between individual grammars, the patterns found in the language as a whole and the complex interactions between language in the mind and language in the community. Therefore, while frequency may play an important role in shaping the cognitive organisation of grammar and while that may be represented at the individual, cognitive level, an important question in this respect is how this structure is brought about and what external factors can explain why some linguistic variants occur in language use settings more frequently than others. The answer to this question is to be found in the integration of the cognitive factors discussed in this section with the social and pragmatic (language-external) factors within which usage events occur, and to anchor this interaction within a theory of language use and change in which variation is central.

In the following section, Section 2.2.2, the social structure of language and language variation, as set out in the sociolinguistic paradigm, are discussed. Specifically, this section focuses on the variationist sociolinguistic paradigm because it holds the potential to provide a framework within which to integrate

the investigation of language variation and change with the mental processes that are at play when users communicate and the social contexts and processes in which this communication unfolds (Bybee, 2010, 2015; Labov, 1994, 2001, 2010).

2.2.2 The social organisation of language

The interrelated patterns of individual user experience, social interaction and cognitive processes give rise to the structure of language (Labov, 1994, 2001, 2010), and their integration allows for a unified account of variation at all levels of linguistic organisation and continuous change within language users and across language communities. There is therefore variation not only within individuals' mental representations of language (as discussed in Section 2.2.1) but also across the individual grammars of the individuals that make up the community, all of which are embedded in and connected by the social, external context in which language use takes place. Therefore, a proper theory of language is one that also takes this variability and the social aspects of language use into account (Labov, 1994; Van Rooy, 2010).

The connection between language use and society is investigated within sociolinguistics, an area of linguistic enquiry also connected to dialectology, historical linguistics and language contact. Sociolinguistic investigations of the connection between language and society endorse the view that language can only be properly understood if the many layers of the social context in which it is used are taken into account (Wodak, Johnstone, & Kerswill, 2011). While sociolinguists are united in their focus on language in use, sociolinguistic investigations are expansive in scope and may broadly be divided into two distinct lines of enquiry: macro-sociolinguistics, which takes society as its starting point and focuses on language as a crucial factor in organising societies or communities (typically addressing issues such as language planning, bilingualism, and intercultural communication); and micro-sociolinguistics, which takes language as its starting point, treating social factors as influencing the structure of language (typically in investigations of conversations, narratives, language use and linguistic variation) (Romaine, 2000; Tagliamonte, 2012; Wodak et al., 2011). Depending on which aspect of language forms the focus of investigation, sociolinguistic investigations diverge into many (interdisciplinary) sub-disciplines, including sociohistorical linguistics and variationist sociolinguistics (Tagliamonte, 2012).

The term 'sociolinguistics' was first coined in the 1950s, with the approach gaining momentum in the 1960s as a result of the work of Labov (1963, 1966) and Weinreich et al. (1968). Much like the family of approaches clustered under the cognitive usage-based paradigm, sociolinguistic investigations were triggered by disagreement with the central tenets of mainstream theoretical linguistics, especially generative grammar and its dismissal of variation and change in language as irrelevant to the study of language (Wodak et al., 2011).

Although linguists have always been interested in language change, it has been one of the enduring puzzles in the study of language, and only in the last few decades have linguists made strides in understanding language change (Deutscher, 2005, p. 61). Historical linguistics, which investigates the histories of

language and language change, typically offers explanations of what had changed in a language by comparing a language at different points in time. But the problem with such an approach is that it cannot explain the processes through which language changed or the mechanisms that brought the change about. The progress that Deutscher (2005) alludes to is due to the fact that research in this area during the second half of the twentieth century has been characterised by the convergence of ideas and approaches from different areas of linguistic investigation, which has allowed researchers to explore *why* language changes and *how* this change is brought about, rather than simply investigating *what* has changed.

In the mid-twentieth century, Labov (1963, 1966) and Weinreich et al. (1968) argued that the (social) forces that operated to change a language in the past also operate on present states of language, and that synchronic language states represent a stage in long-term change, or ongoing change, and therefore, the present can be used to explain the past (Labov, 1994). Language change is now recognised not only as a phenomenon of the past, but as a phenomenon that is just as evident in ongoing changes, and which is observable in the variation that occurs in synchronic language states (Bybee, 2015, p. 1).

Labov (1963, 1966) investigates the variation that arises in everyday language use and incorporates social factors in his analyses of the linguistic features that are subject to variation. His argument is rooted in the rejection of the “ideological barriers” in the theoretical linguistics of the time, and the fact that language is always situated in context (Labov, 1972, p. xiii). Most crucially he argues that variation is an inherent part of language, and that it is through synchronic variation that language change proceeds. The result of this early work is the recognition of language as a structured system that is also variable, and that this variability is conditioned by linguistic (language internal) and social (language external) factors. Furthermore, Labov (1963, 1966) shows that variation is found and patterned both within individuals and across individuals, and demonstrates that these patterns of variation can be studied quantitatively and can be correlated with social and stylistic factors.

The quantitative patterning of variation in language and its use as a method to investigate language change is further highlighted by Weinreich et al. (1968), who state that “[n]ot all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity” (p. 188). In other words, language users have choices or options in how they use language, and these choices or options form predictable patterns on the basis of the linguistic system and social factors.

Sociohistorical linguistics, or historical sociolinguistics, is an interdisciplinary field of linguistics in which the traditionally separate fields of historical linguistics (language change) and sociolinguistics (language variation) converge. In broad terms, the two fields differ with regard to their foci: while historical linguistics focuses on the histories of languages and language change, sociolinguistics focuses on the role of language in society and the variation that arises in language due to social factors. Sociohistorical linguistics, then, is a sub-discipline of sociolinguistics and historical linguistics, bringing together the two fields and emphasising the role of social, external factors in the study of language change (Nevalainen, 2008).

Weinreich et al. (1968), who first drew attention to the importance of incorporating external, social factors into diachronic analyses of language, are widely considered to have marked the birth of the discipline (Russi, 2016), while Romaine (1982) is regarded as having laid its foundation. Romaine (1982), for example, shows that it is possible to draw on quantitative sociolinguistic models to interpret relativisation in mid-sixteenth-century Scots, and demonstrates that sociolinguistic methods can be used to correlate linguistic variation with external factors in historical data (Nevalainen, 2008, p. 561). In so doing, Romaine (1982) shows that “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” (Wasserman, 2014b, p. 23). Sociohistorical linguistics therefore assumes that language is a product of both historical processes and social factors and views language-internal and language-external factors as systematically interconnected in language change (Russi, 2016).

Sociohistorical linguistics is not the only sub-discipline of sociolinguistics that emphasises the inter-relatedness of synchronic variation and diachronic change, nor that highlights the importance of linguistic and social aspects in understanding language variation and language change. Variationist sociolinguistics, which investigates variation in language and how this relates to language change, also has its roots in the early work of Labov (1963, 1966) and Weinreich et al. (1968). Variationist sociolinguists investigate language use embedded in social contexts, and how this use can be understood not just from a communicative angle but also as language users’ conscious or subconscious acts of identity and social distinction (Auer, Peersman, Pickl, Rutten, & Vosters, 2015). Specifically, variationist sociolinguistics views language as a dynamic, emergent system of usage characterised by social factors: the system comprises multiple agents who *interact* with one another; these agents are the users in the *language community* whose behaviour (which is influenced by several competing factors, such as perceptual mechanics, social motivation, and pragmatic context) is based on past interactions; because past interactions, together with current interactions, feed into future behaviour, the system is adaptive (Auer et al., 2015).

Variationist sociolinguists draw on the notion of a linguistic variable, which, very simply, is two or more ways of saying the same thing, and therefore offers language users a choice among variants which have the same referential value within the same grammatical system (Tagliamonte, 2012). These variants form a linguistically defined set of some type: they may be delineated at any level of grammar, but they have a structurally defined relationship in the grammar. The variants co-vary, and this co-variation may be correlated with linguistic factors (or the structural considerations and relations within the language system), and social factors, such as sex, age, social class, ethnicity, race and community size (Tagliamonte, 2012). This co-variation and its correlation with the independent linguistic and social factors, means that variants are systematically patterned, and therefore their distribution in language use in a community of language users can be studied quantitatively (Tagliamonte, 2012). Genitive alternation, which is the feature selected for investigation in this study (see Sections 1.2 and 1.3), is an example of one such linguistic variable: there are two major adnominal constructions available that offer two ways of saying the same thing; the variants

are known to co-vary and this co-variation is known to be conditioned by very specific linguistic factors (see Section 3.4.3.1), and various social and discourse-pragmatic factors (such as variety of English and register).

Language-internal and language-external factors are therefore crucial in understanding language structure, language use and language change because they are closely interwoven in how variation in a language patterns. Furthermore, linguistic variables typically involve variants that have been assigned social meaning (and are called sociolinguistic variables). These evaluations of language are normally centred on issues of stigma and salience and are due to the external evaluations of language users (Labov, 1994). As a result, these evaluations differ from one community to the next and are therefore embedded in the social fabric of the community. The manner in which linguistic variables pattern in a community and how the community evaluates the variants therefore provides information about how that community is organised (Tagliamonte, 2012, p. 6).

As discussed above, not all variation leads to change, but all change requires variation. Therefore, if the linguistic variable under investigation is known to be undergoing change, then the investigation of that variable and how it patterns in terms of linguistic and social phenomena, can shed light on how and why language change occurs and proceeds through the community.

Situating this within the usage-based framework, Bybee (2007a, p. 945) explains that as language comes to be viewed as a negotiated set of social and cognitive behaviours, the importance of studies of language change increases because “language change provides evidence for the nature of linguistic representations and processing, and thus provides a window on synchronic mental representation and the forces that create grammar”. Thus, the processes and factors that give rise to language structure are also the processes and factors that give rise to synchronic variation and lead to language change. The next section, Section 2.3, investigates in more detail theories of language change, focusing in particular on evolutionary theories of language change, because these theories have been invoked in models of World Englishes, such as Schneider’s Dynamic Model (2007), and because the various mechanisms proposed in these theories can be usefully linked to the cognitive and social factors that drive language structure and language variation.

2.3 Language change: Evolutionary theories

As mentioned in Section 1.1, language change occurs through two equally important, but separate processes. First, variation is introduced into the language system (typically as a result of individual psycholinguistic processes), either through the introduction of a new feature that comes into competition with an existing one, or through a quantitative shift in the frequency of an already-existing variant (which leads to a concomitant decline in the frequency of other competing variants). Second, this variant or usage pattern diffuses through the community or society of language users, eventually becoming wide-spread enough to be considered conventional (Croft, 2000).

Although there is general agreement that change proceeds through this two-part process, there is considerably less agreement on the entities and mechanisms involved in the two processes (Croft, 2000, p. 43). The emergence of linguistics as a field of scientific enquiry and of evolutionary theory coincided in the nineteenth century and subsequently gave rise to debate on the origin of language and the relationship between language change and biological evolution. During that time, a number of theories and conjectures were put forward but they were so unusual that in 1866 the *Société de Linguistique de Paris* banned discussion of the topic (Christiansen & Kirby, 2003, p. 300).

Contemporary historical linguists have started to re-explore the connection between linguistic evolution and biological evolution and have adopted theories of biological evolution to construct evolutionary theories of language change (da Silva, 2010, p. 806). Evolutionary frameworks are well suited to explaining language change because they provide a framework that embeds language structure within the ecology, or larger context, in which language is used, and allow linguists to propose and explore plausible (cognitive, social, pragmatic) mechanisms for “the realization of language structure, the generation of variation in language structure, and the selection of structural variants in an evolving population” (Croft, 2013, p. 102).

Evolutionary theories of language change propose that language change is an example of the same or similar processes found in evolutionary biology. Two evolutionary models of language change are discussed in this section. Although the two approaches differ in how they frame language change in evolutionary terms, they are similar enough to allow for meaningful comparison. Mufwene (2001, 2008) proposes a model of language change that draws on concepts from evolutionary biology, such as *variation*, *competition*, *selection* and *founder principle*, and is rooted in the notion that all language change proceeds through contact, while Croft (2000) draws on David Hull’s Generalised Theory of Selection (1988) to propose the Theory of Utterance Selection to account for language change. The two approaches are selected here because both take usage-based views of language and because both are invoked in explanations of language change in models of World Englishes.

Mufwene (2001, 2008) states that the development of his evolutionary model of language change is inspired by population genetics and macroevolution. He favours the use of *language evolution* rather than *language change* because in addition to covering the structural and pragmatic changes that a language may undergo, *language evolution* covers language speciation, and language birth and death (Mufwene, 2008, p. 16). Although his model analogises language change to biological evolution, he states that the model is analogical only to the extent that it is inspired by scholarship on biological evolution, and therefore does not set out to match all properties of language change to biological evolution. Croft (2000), on the other hand, views analogical approaches to language evolution as opportunistic and therefore unreliable, and argues for a generalised approach to the development of an evolutionary model of language change. The generalised approach is, according to Croft (2000), better because it abstracts the essential properties of evolution from the phenomenon being studied and therefore resolves any issues that are the consequence of analogising processes from one scientific domain to another.

In his competition-and-selection model, Mufwene (2008) argues that language evolution can only properly be accounted for if language is analogised to a viral species, because languages need users to ensure their continued existence. Because language users are constantly changing languages as they use them, they are responsible for the changes that languages undergo and are therefore the unwitting agents of change (Mufwene, 2008, p. 13). This highlights two crucial arguments in contemporary theories of language evolution. First, language change is a result of the acts of individual language users as they communicate. Second, language users do not intend to change their language; rather, they wish to attain some other goal (such as to communicate efficiently) and in the course of doing so, alter the language (Croft, 2008; Deutscher, 2005; Mufwene, 2008).

In the competition-and-selection model, language is viewed as a population of individual language users' idiolects, each of which displays certain idiosyncrasies that give rise to variation in the language, despite the fact that there may be similarities between the idiolects in the population (Mufwene, 2008). In drawing on the population analogy, Mufwene (2008, p. 11) calls attention to a crucial understanding of language for his theory: "languages are communal systems comprising the varying, yet similar, idiolects of individual users, and are constructed on the basis of similarities in the way people speak in order to facilitate mutual intelligibility". Convention, or similarities in the way people speak, is therefore established at the broader, population level of the language, while variation operates at the level of the individual.

Croft (2000) too describes a language in population terms. In line with usage-based approaches, he views language as comprising the real, existing actual occurring population of utterances produced and comprehended in context. He therefore emphasises that the real object of linguistic enquiry is "a spatiotemporally bounded set of actual utterances, and not a set of potential utterances" (Croft, 2000, p. 26) (which stands in contrast to the generative view of linguistic enquiry). Croft (2000) and Mufwene (2001, 2008) therefore both view languages in population terms but they differ in the sense that Mufwene (2001, 2008) views a language as comprising the population of individual users' idiolects, while Croft (2000) views languages as comprising a population of actual occurring utterances. A further dissimilarity between the two, and one that is related to how each views language, is the question of what gets replicated during the process of language change, and consequently what the unit of selection is.

Because Mufwene (2008, p. 116) views a language as comprising a population of idiolects, he states that selection works directly on idiolects. He argues that the two-part process of language evolution proceeds through the competition and selection of linguistic features that individual users in contact situations contribute to a *feature pool* of possible linguistic forms and variants. This feature pool comprises the "total amount of linguistic knowledge language users have, including the variants that compete for the same structural and communicative functions [...] both at the individual and at the population levels" (Mufwene, 2008, p. 17). New features are introduced into the communal language system through the contact between and interaction of individual users' idiosyncratic idiolects in communicative acts and is driven by "the accommodations which speakers make to each other, the adjustments they make to new communicative

needs, and the simple condition of imperfect replication during language transmission” (Mufwene, 2001, p. 18).

Contact plays a fundamental role in Mufwene’s theory of language evolution, because it is through contact that variation is introduced into the language system and how the feature pool is ‘filled’. Every setting of language use is considered a contact setting, and contact is essentially inter-idiolectal, at the level of individuals interacting, even when idiolects of different languages are in contact (Mufwene & Vigouroux, 2012). The most significant contact is that between individuals (and not populations) and the real locus of contact is the minds of individual users, “where structural information is processed, and where features associated with the same or similar functions are brought into competition and can be negotiated during interactions” (Mufwene & Vigouroux, 2012, p. 113). Language users therefore introduce the competition that arises between the variants in the feature pool (Mufwene, 2008).

Selection, which (to some extent) resolves this competition, occurs when certain features are selected from this feature pool and are copied by other language users, resulting in the dominance of one feature compared to its competing variant(s) (Mufwene, 2008). Speakers’ minds are therefore the arenas where competition is resolved. If enough language users in the community make the same selections, community-level selection takes place and it is this that produces the “macroevolutionary developments identified as changes in the language” (Mufwene, 2008, p. 18). In other words, no community-level selection (which *is* language change) can take place without individuals (who introduce variation, add these variants to the feature pool and start and resolve the process of competition through their selections).

Although Mufwene (2008) states that community-level selection cannot take place without individual selections, he points out that the two kinds of selection are not always convergent. This, he argues, demonstrates the competition that occurs in a language and the relevance of both structural and non-structural selections which language users make for determining the evolution of a language. Whatever the source of variation, all variants thus exist and compete within the feature pool until something in the external ecology shifts, causing some features to be selected more often or not, thereby undergoing shifts in their probability, and resulting in some that used to be recessive becoming dominant, and vice versa (Mufwene, 2008, p. 23). The mind filters these external ecological pressures and uses them to resolve the competition among the variants (Mufwene & Vigouroux, 2012). As a result of this, languages do not change in uniform ways, because different idiolects, sociolects and dialects engage in changes differentially relative to their ecologies (Mufwene, 2008).

Ecology, which is a cover term for the range of factors that influences a language’s evolution (including factors that are external and internal to the language), is thus important in Mufwene’s model. While selection plays an important role in language evolution, it is at the mercy of ecology, and therefore the processes of language change are determined by the ecologies in which language users exist (Mufwene,

2001, 2008). The ecology in which languages exist and change is therefore the most influential factor in the process that both gives rise to variation and leads to change.

Although Mufwene acknowledges that change is brought about by both internal and external factors, he argues against the traditional separation of the two, because “no ecological factor alone accounts for everything” (Mufwene, 2001, p. 165). Instead, the external ecology of a language interacts with its internal ecology, and it is this interaction that explains both the causes of change and how change spreads (Mufwene, 2008). To this end, Mufwene and Vigouroux (2012, pp.112–113) argue that studies of language variation and change should describe what the ecology of a language consists of, including both the internal and external ecologies, and the factors that have affected the evolution of the language, such as when the users of the language arrived, the demographic strength of its population of users, the economic and social power of those users, as well as other factors such as users’ age, gender and religion.

Mufwene’s theory therefore makes two important assertions about language evolution. First, the evolution of a language lies in the individual communicative acts of its users which give rise to variation, but it is the cumulative selections of variants from the feature pool by individual language users in these individual communicative acts that result in changes in the language. Second, the communicative acts and activities that individual users engage in are entirely dependent on the ecologies in which they evolve and use language.

Croft (2010) states that there are similarities between his and Mufwene’s approaches to language evolution; however, his primary criticisms of Mufwene’s model relate to the use of analogies that Croft argues do not fit together, what Mufwene conceives of as the replicator, and how he defines language. The Theory of Utterance Selection is based on Hull’s (1988) Generalised Analysis of Selection, in which four key concepts are crucial. Citing Hull (1988, pp. 408–409), he identifies these concepts as follows:

- (1) REPLICATOR – an entity that passes on its structure largely intact in successive replications
- (2) INTERACTOR – an entity that interacts as a cohesive whole with its environment in such a way that this interaction CAUSES replication to be differential
- (3) SELECTION – a process in which the differential extinction and proliferation of interactors CAUSES the differential perpetuation of the relevant replicators
- (4) LINEAGE – an entity that persists indefinitely through time either in the same or an altered state as a result of replication (Croft, 2000, p. 22).

In the Theory of Utterance Selection, a language is viewed as comprising the population of actual utterances produced and comprehended by a speech community. Utterances (the linguistic equivalent of organisms) are comprised of linguemes (the linguistic equivalent of genes). An utterance is the product of the communicative interaction of human behaviour, as it is structured, and semantically and pragmatically interpreted in its context (Croft, 2000). An utterance therefore represents an actually occurring pairing between linguistic form and semantic and pragmatic function. This definition lays the groundwork for

several fundamental points in the theory: utterances convey intended meaning (in context) by the speaker and this meaning is interpreted and understood by the hearer; because they are context dependent, utterances are spatiotemporally bound, and therefore include only actual occurring instances of language use, and not possible instances (Croft, 2000). Thus, in the Theory of Utterance Selection, the lingueme is the replicator, language users are interactors, selection (the locus of language change) is the use of a particular utterance containing a lingueme or set of linguemes, and a lineage relates to the etymology of a linguistic unit (Croft, 2000).

Croft (2000) expands his theory by stating that the two-part process of language change should be viewed as change by replication, and not as inherent change. Innovation is framed as altered replication of the replicator, in which variant linguemes, which do not conform to linguistic conventions, are introduced into the language and added to the lingueme pool (Croft, 2000). Normal replication is the replication of an existing replicator, and does not introduce variation. Differential replication is the propagation of an innovative form in a language through the process of diffusion. This differential replication is a result of “the differential extinction and proliferation of interactors as well as the differential perpetuation of the relevant replicators” (Croft, 2000, p. 31), which is the process of selection and which results in the establishment of a new convention.

Convention is therefore a crucial component of Croft’s (2000) theory because normal replication is essentially conformity to convention; altered replication is a violation of or deviation from convention; and differential replication is the gradual establishment of a new convention. Convention is a regularity of behaviour that is considered to be the common ground of a community and which is used as a coordination device to communicate meaning (Croft, 2000, p. 98). Convention can only be considered a coordination device if (almost) everyone in the community conforms to it and expects others, including additional and new members of the community, to conform to it too. In other words, conventions are norms that regulate language use and they are considered so because they are widely shared and because nearly everyone in the community conforms to, or is expected to conform to, them. Convention is therefore a *social* factor.

The causes of language change may be internal to the language or external to the language. Internal causes of change are linked to the structure of a language, while external causes of change are brought about by social factors, including language contact (Croft, 2000). The distinction between internal and external causes of change is framed in terms of the mechanisms that give rise to change, which may be intentional and nonintentional. The distinction between intentional and nonintentional mechanisms for the two processes of language has its roots in the “popular misconceptions of evolutionary processes [as being teleological, and which can be] explained in terms of the organism aiming towards some goal” (Croft, 2000, p. 64). Croft (2000) argues against this view and states that evolutionary processes are probabilistic, and not deterministic. Language users do not aim to change the language when they use it; rather the language changes as a consequence of language users trying to achieve some other goal, or through the unintended

introduction of an innovation (Croft, 2000). The mechanisms for altered replication, normal replication and selection may therefore be intentional or nonintentional.

Normal replication, that is conformity to convention, is an intentional mechanism, because conventions are the shared common ground in a language community and are therefore based on social interaction (Croft, 2000). This social interaction is driven by language users wanting to be understood by drawing on their mutual knowledge of conventions. Mutual knowledge of conventions also requires individual knowledge of conventions. An individual's knowledge of conventions is called 'competence': it is a psychological phenomenon that varies depending on an individual's knowledge of the community's conventions – and, in line with usage-based approaches, is the result of a language user's experiences with language. 'Convention' is specific to a community and is therefore a social phenomenon that can vary "to the degree to which [it] is established in the community" (Croft, 2000, p. 72). In other words, some conventions may be widely established in a community, while others are less so. 'Competence' may be driven by the purely psychological routinisation of a behaviour, or entrenchment, in which case it may be framed as a nonintentional mechanism in normal replication (Croft, 2000). But, entrenchment is also sensitive to use, is reinforced through use and decays through a lack of use. In other words, the more a form is used (produced and comprehended), the more likely it is that the form will become entrenched. This "variation in entrenchment relative to exposure to use" (Croft, 2000, p. 73) means that when a form is replicated not as a result of a purely automatic psychological habit, but rather because of its reinforcement through exposure to use, then it is actually a mechanism for selection. Therefore, the frequency with which an innovative token is encountered, reinforces its use, leading to cognitive entrenchment in the mind of the language user, and ultimately to its entrenchment in the community of users. Entrenchment may therefore be a nonintentional mechanism for normal replication, but it may also be a nonintentional mechanism of selection.

The mechanisms for altered replication are also either intentional or nonintentional. Innovations in service of expressiveness, or to avoid misunderstanding or for reasons of economy, are all intentional innovations. The unintended result of altered replication arises from a potential discrepancy between an individual's competence, which constantly responds to use, and the conventions of a speech community (Croft, 2000, p. 118). Unintended altered replication is therefore an individual psycholinguistic phenomenon that is brought about by the nonintentional reanalysis of an individual's mapping between a grammatical form and its conventional meaning or function, or form-function reanalysis, and which is based on a simple model of grammatical structure, as presented in usage-based approaches (Croft, 2000, pp. 78, 119).

Selection, the process through which innovative features are propagated, leads to the adoption of a new convention (Croft, 2000, p. 166). The mechanisms for propagation are social because they mostly involve the relationship between a language user and the society to which he/she belongs. The process of selection is therefore related to the structure of a society and the nature of the conventions in that society (Croft, 2000). Selection requires previously created variants to operate on and is driven by language users'

particular social values, which are informed by their social identities and their ties to the society to which they belong. Drawing on sociolinguistic work, Croft (2000) argues that the structure of language user populations and language users' ties to the networks in those populations explain how variants are introduced into different networks and then spread. He invokes the social network theory proposed by Milroy and Milroy (1992), to argue that language users belong to several networks within a population, and their ties to these networks range from weak to strong. Strong-tie individuals are central members of their networks and tend to maintain the conventions of those networks (Croft, 2000). They are therefore less likely to introduce innovations (or propagate variants) into these networks. Weak-tie individuals, on the other hand, are on the periphery of a network(s) because they have multiple ties to several networks, and are therefore more likely to introduce the linguistic innovations of other networks into a network because the innovation (or variant) "is not correlated with strong network ties" (Croft, 2000, p. 179). Although the network tie theory of population structures explains the channels through which variant forms are selected by individuals in a population, it does not explain the mechanisms through which these variants become selected (Croft, 2000). As already mentioned, the nonintentional process of entrenchment (through exposure to use) is one very important selection mechanism that is rooted in individual cognitive representations of language. The following three (intentional) mechanisms are identified to account for the selection of innovative linguistic features in a society: prestige, acts of identity, and accommodation (Croft, 2000, p. 179).

Prestige has to do with the power relationships between different communities, where one community is viewed as being socially dominant or more powerful (Croft, 2000). The implication is that the linguistic variant preferred by this community (which is consequently viewed as the standard) will be propagated into the less powerful community. However, Croft (2000) adds that this may not always be the case and distinguishes between overt and covert prestige. Overt prestige relates to the propagation of the variant of the more powerful community, while covert prestige refers to when the variant of a less powerful community is propagated.

This other factor is typically argued to be based on the selection of a variant as an act of identity with a social group (Croft, 2000). This occurs when a language user selects the variant of particular group because she wishes to identify with, or be identified as belonging to, that group (Croft, 2000). This selection may be directed towards identifying with the more powerful group in the community, but it may also be that a language user wishes to identify with a less powerful group, "particularly speakers in a strong-tie network with a local group speaking a vernacular" (Croft, 2000, p. 181).

The last mechanism for selection, accommodation, also relates to social structure and in particular to a language user's interpretation of another's competence. During accommodation, language users will adjust their language use, and select variants to compensate for the fact that they come from a different speech community or to align their language with the conventions of the speech community to which the person they are communicating with belongs (Croft, 2000). Accommodation thus leads to the convergence of the

language of the interactors in the communication setting. A last factor in social identification is the shifting of stylistic behaviour in different interactional settings. This occurs when the proportion of variants that are selected and used in communicative contexts is directly related to group ranking by class. In other words, “formal styles [which are typically associated with social groups of a higher status] elicit a greater proportion of the higher (standard) variants” (Croft, 2000, p. 182).

The selection mechanisms outlined above lead to the differential perpetuation of variants within a language: language users cause the differential perpetuation of replicators, and the selective advantage of the variants is determined by social factors, such as prestige, acts of identity, and accommodation. However, the propagation of variants is also dependent on the cognitive structures in a grammar that are used by the language user (Croft, 2000). In other words, language users will only continue to (re)select variants as long as they exist in their grammars and are reinforced in use. Croft calls this “the survival of cognitive structures [or entrenchment]”, and states that “[t]he shift in the proportions of the variants of a linguistic variable in usage is brought about by shifts in the degrees of entrenchment of those variants in the grammars of speakers” (Croft, 2000, p. 32). The shifts in the degrees of entrenchment are due to exposure to the variant in language use, and the more often a language user encounters and uses a variant, the more entrenched it becomes in the mind, resulting in a shift in the variant and increasing the likelihood that the variant will be (re)selected. Fundamentally then, the frequency of features as well as the social forces that condition these frequencies, are unified as the most significant factors driving language evolution but are also the factors that explain language structure and language variation.

In incorporating the effects of language contact in language change, Croft (2000) states that the effect of contact on an acquired language may be understood as substratum interference. In this regard, the only way to account for substratum influence is by looking at the size of the shifting population, its social status and the length of time for shift (Croft, 2000). Significant substratum influence is most probably brought about in situations where the shifting population is larger than the native population, where the shifting population is bilingual, and when the shift is relatively rapid. The underlying mechanism enabling substratum influence is the “degree to which native speakers of the acquired language accept shifting speakers’ ‘errors’ as new conventions of the acquired language” (Croft, 2000, p. 204).

As mentioned in the conclusion to Sections 2.2.1 and 2.2.2, language change (and by implication synchronic variation) cannot be separated from language structure, language in the mind or language in society, and all of these features of language therefore play a crucial role in understanding language (Croft, 2000).

A crucial point that emerges from the discussion of language structure, use and change that is essential to this thesis is the centrality of convention in theories of language. But the notion of convention (or norms) has many dimensions. In the cognitive view of language (discussed in Section 2.2.1), a convention is viewed as the cognitive entrenchment of a feature as a result of exposure to it in actual instances of language use. In the social view of language (as discussed in Section 2.2.2), a convention is defined at the level of the

language community, and occurs when a feature is propagated within and diffuses through a community of users. Typically, these two forms of conventionalisation are studied in isolation. But given the centrality of conventionalisation for the establishment of norms, it is important to investigate how cognitive conventionalisation and the different dimensions of social conventionalisation (such as the distinction between ‘top-down’ conventionalisation processes and ‘bottom-up’ conventionalisation)⁴ interact.

The study of World Englishes offers a good opportunity to explore the interaction between cognitive conventionalisation and social conventionalisation in languages, because most of the world’s Englishes have developed in different contexts, and have come into contact with other languages to varying degrees, giving rise to a number of varieties in which variation in English is highly structured and correlated to the particular speech community.

2.4 Variation in the world’s Englishes

The global spread of English due to the colonial expansion of the British empire, the industrial revolution, and more recently, the emergence of the United States of America as an economic and military power has led to the use of English as the language of international communication, politics, commerce, and the media (Schneider, 2007). English is now not only the global lingua franca, but it has also diversified into home-grown forms and uses in many locations (Schneider, 2007).

The differences among the varieties of English are the result of general processes of language change as they play out in very particular contexts. A number of frameworks and models have been developed to categorise the different varieties of English across the globe and to describe their development and functions (Onysko, 2016). Two such frameworks and models will be discussed in this thesis: Kachru’s World Englishes paradigm and corresponding Three Circles Model (Kachru, 1985, 1990, 1992, 1996), and Schneider’s PCE paradigm and corresponding Dynamic Model (Schneider, 2003a, 2007). They form the focus of the discussion here because both place normativity as a central concern in their models and because both recognise the importance of the development of local norms in the different varieties.

In 1985, Braj Kachru proposed a model for the study of the varieties of English that focused on the characteristics of and relationships between the different varieties. Kachru (1985) called for a paradigm shift that he argued would liberate the study of English from the traditional monolithic view of the language and would recognise the different varieties of English as varieties in their own right, drawing attention to their distinct identities and the variation in the form and function among them due to the different linguistic and cultural contexts in which the language is used. He argued that the English language belongs to all who use it and stated that the pluricentric view was necessary in order to capture the range and depth of the spread of English across the globe, not only in demographic terms, but also in cultural and linguistic terms.

⁴ Section 2.6.1 discusses the distinction between these two types of social conventionalisation in further detail.

To accommodate this he proposed the Three Circles Model in which the varieties of English used in the world are grouped in three circles according to their the types of spread, patterns of acquisition and the functional domains in which they are used in different cultures and in contact with different languages (Kachru, 1985, p. 12). According to Kachru (1985), the Inner Circle represents regions considered to be the traditional bases of English and where English is used as the primary language, such as in the United States, Britain, Australia, New Zealand and Canada. As such, Inner Circle varieties are described as norm-providing because their users are native users of English and because they are traditionally considered to serve as models for other varieties of English. Outer Circle varieties are those regions where English is used as a second language (ESL or L2) and where it is institutionalised in non-native contexts. In these regions, English is only one of two or more languages, and is used in a wide range of domains and to varying degrees of competence. Varieties in the Outer Circle are norm-developing and are characterised by both an exonormative linguistic orientation, drawing on the norms of Inner Circle varieties, and an endonormative linguistic orientation, in which norms internal to the variety gradually develop. Expanding Circle varieties include regions where English is used as a foreign language for purposes of international communication, such as China and Indonesia. As such, Expanding Circle regions are viewed as norm-dependent, and are therefore characterised by an exonormative orientation.

Kachru (1985) states outright that it is the non-native users of English (those within the Outer Circle) that initiate and control the differentiation of the language, and he therefore places more emphasis in his model on the varieties used in the Outer Circle. He argues that the norms and standards of Inner Circle countries cannot and should not determine the norms of the Outer Circle countries where the language is used by non-native users.

Despite the usefulness of the model, one of the primary criticisms against it is the use of national borders to identify varieties of English, which abstracts away from the complex contexts in which varieties are used (Schneider, 2007). For example, in bi- and multilingual countries such as South Africa, multiple sub-varieties of English are used, making it difficult to place the country within the model because it could belong to two or all three circles at once. Nevertheless, the importance of Kachru's (1985, 1990, 1992, 1996) work is that it foregrounds the role of norms in identifying different varieties of English, and raises crucial questions around issues of innovation, normativity and codification.

In 2003 and 2007, Edgar Schneider proposed and elaborated the Dynamic Model for the different varieties of English that share their origins in British colonisation activities, or PCEs. The Dynamic Model asserts that the differential language change of PCEs is related to the very different extralinguistic contexts in which they have been used, and this has resulted in the differences among them in their forms and functions (Schneider, 2007). Furthermore, the colonial diffusion of English brought it into contact with other languages in different contexts and therefore PCEs have developed also in contact situations. To account for the importance of these contexts and the accompanying contact situations, Mufwene's (2001) evolutionary perspective of language change is invoked in the development of the Dynamic Model. Rather

than focusing on identifying specific differences between the varieties, the model proposes an underlying fundamentally uniform evolutionary process in which both psycholinguistic and sociolinguistic processes underlie the emergence of all PCEs (Schneider, 2003a).

In adopting an evolutionary perspective that incorporates both individual psycholinguistic and sociolinguistic processes (as discussed in Sections 2.2.1 and 2.2.2), various sub-disciplines within linguistics are invoked and the importance of new language varieties emerging in a competition-and-selection process between features available to users in a ‘feature pool’ of possible linguistic choices is emphasised (Schneider, 2007, p. 23). This feature pool consists of the “sum total of the individual forms and variants that each of the users involved, with different language backgrounds and varying linguistic experiences, brings to the contact situation” (Schneider, 2007, p. 22). Within these contact situations the interaction between two parties (i.e. the settlers and the indigenous population) is crucial, because it is this that gives rise to changes in the language and the development of the varieties (Schneider, 2007).

Variation among the varieties of English therefore results from the addition of innovative features to each variety’s feature pool and the selection and conventionalisation of certain features within the different contexts in which the varieties develop. This makes the ecologies (as described by Mufwene (2008)) in which this process unfolds an important determiner of the competition-selection outcome. The development of the different varieties is described as their progression through five consecutive stages, ultimately culminating in a new variety of English with new linguistic norms (Schneider, 2007, p. 21). As discussed in Section 1.2, normativity is therefore central to the Dynamic Model, and it is the progression of different varieties towards endonormativity that characterises the stage of their development.

The five stages are Foundation, Exonormative Stabilisation, Nativisation, Endonormative Stabilisation, and Differentiation (Schneider, 2007). At Phase 1, Foundation, English is brought to a new territory where it has not been used before, by a group of British settlers. These settlers typically originate from different regions within Britain and so the English spoken by them is not homogenous (Schneider, 2007). Initially the relationship between the settlers and the indigenous population of the land in which they have settled can range from friendly to unfriendly. During this stage, both groups are aware of each other’s presence, but each group considers itself distinct from the other: the settlers still consider themselves as “outposts of Britain, deriving their social identity primarily from the common territory of origin and a feeling of culturally belonging there” (Schneider, 2007, p. 37). Two kinds of contact situations emerge: on the one hand there is dialect contact between the settlers, and on the other there is (limited) contact between the settler population and the indigenous population. The first type of contact is characterised by a trend towards homogeneity and results in koinéization within the STL strand. The second level of contact, that between the settlers and the indigenous population, is restricted in this phase; however, some contact is achieved via cross-cultural communication by a few members of the settler and indigenous populations (Schneider, 2007). The limited communication between the two groups means that the acquisition of the linguistic skills required for these interactions is left to the indigenous population and marginal bilingualism develops,

giving rise to an indigenous variety of English (Schneider, 2003a, p. 245). There is therefore little influence of the indigenous language(s) on English, with the exception of toponymic borrowing.

In Phase 2, Exonormative Stabilisation, the settler community stabilises in the new territory and the native-speaking settlers provide for most of this stable usage of the STL strand (Schneider, 2003a, p. 245). During this stage the members of the settler community still view the community as an outpost of Britain, and therefore share a conservative and unaltered external norm orientation. However, adjustments to the new environment and culture start to creep in and slowly modify the STL strand, most notably in the adoption of local vocabulary especially for fauna, flora and cultural terms (Schneider, 2003a, 2007). Within the indigenous population increased contact with the settler community means that bilingualism starts to spread, and while indigenous population views itself as a local community, members of this population who have learnt English view themselves as “having an extra edge of experience and competitiveness within their own native group” (Schneider, 2007, p. 37). This leads to beginnings of the segregational elitism (which characterises English in some PCE-speaking countries to the present day), and represents the start of structural nativisation, in which one population group shifts to the language of the other, invariably resulting in some transfer phenomena in terms of phonology and structure.

Phase 3, Nativisation, is the most important stage of the model because it is the central phase of cultural and linguistic transformation (Schneider, 2007, p. 40). Contact between the indigenous and settler groups increases due to the weakening of ties with Britain and growing political independence. As a result, both settler and indigenous populations realise that they and the other are permanent residents, leading to a narrowing of the gap between the two populations. Several factors, including an appreciation of English and its use in important functions, place pressure on the indigenous group to adopt English, leading to widespread second-language acquisition, with accompanying local linguistic idiosyncrasies, and sometimes an almost complete language shift to English (Schneider, 2003a, p. 248). During this stage, the settler community is usually divided into two groups: innovative users who are likely to adopt IDG strand features, and conservative users who “reject the idea of linguistic innovation and local adjustment altogether and keep believing and insisting that the only acceptable way of using English is the metropolitan, conservative [external] norm” (Schneider, 2007, p. 43). This leads to a complaint tradition in which conservative users view features that deviate from this external norm as errors rather than as innovations. Despite this, the heaviest restructuring of the English language occurs during this third stage, most notably in the innovations and structural nativisation that spread from the IDG strand to the STL strand. Consequently, Schneider (2003a) states that this phase is of most interest to theoretical work in language change because it demonstrates how a linguistic system may be modified and how innovative features gradually stabilise and become selected at the communal level.

Phase 4, Endonormative Stabilisation, is characterised by the adoption of an indigenous local identity and the acceptance of a local linguistic norm. This acceptance is driven by political independence and the reconstruction of the settler community’s identity. The shift to Phase 4 is typically marked by some sort of

“quasi-catastrophic event”, an event X, in which the settlers realise that there is an inverse relationship between the importance they used to place on Britain and the (considerably lower) importance that the colony is afforded by Britain (Schneider, 2003a, pp. 48–49, 250). This effectively leads to a kind of convergence between the settler and indigenous populations and results in the disappearance of or reduction in the cultural and linguistic differences between the two population. This gives rise to a variety that is perceived and emphasised as “remarkably homogenous” (Schneider, 2003a, p. 251), which is symbolised in the use of “X English” (i.e. Australian English or Singaporean English) to refer to the variety. Questions regarding the acceptance of the new linguistic norm are usually social group struggles. However, the codification of the new variety (usually first in dictionaries and then later in usage guides, style guides and grammars) leads to acceptance of the local linguistic norm.

Phase 5, Differentiation, is described as a stage in which internal differentiation in the newly-established variety emerges. Following the political, cultural and linguistic independence of the nation, this stage marks a turning point in which the “focus of an individual’s identity construction narrows down, from the national to the immediate community scale” (Schneider, 2003a, p. 253), and sees individuals defining themselves as part of distinct subgroups, differentiated by various sociolinguistic parameters. This leads to the diversification of the language into distinct regional and social dialects, and in bi- or multilingual communities, to first- and second-language varieties.

The importance of norms, in particular the issue of norm orientation, is central to both Kachru’s (1985) and Schneider’s (2003a, 2007) models. The importance of norms in theories of language use, structure and change, as well as the importance of norms in the development of the different varieties of English offers an opportunity in which to situate the role of editorial intervention in the processes of ongoing linguistic variation and change in the varieties of English. Before the role of editorial intervention will be discussed, it is first necessary to briefly discuss a further type of variation that is relevant not only to language broadly but also to the different varieties of English, as well as to editorial work: register variation.

2.5 Register and language variation

Corpus-based approaches to studies of language use have made it possible to study the actual language used in naturally occurring texts in terms of the systematic ways in which linguistic features are used in association with other linguistic and non-linguistic features (Biber, Conrad, & Reppen, 1998, pp. 1, 5). In this regard, Biber et al. (1998) state that studies of language use that utilise corpus-based approaches tend to focus on particular linguistic features or the characteristics of texts or varieties (such as registers, dialects or styles). According to Biber and Conrad (2009, p. 6), a register (as discussed in Section 1.2) is a variety of language that is defined in terms of a particular set of situational characteristics. They add elsewhere that although registers are defined in non-linguistic terms (including the language user’s purpose in communication, the topic, the relationship between the addressor and the addressee and the production circumstances), there are important linguistic differences across registers that are related to differences in

situational characteristics (Biber & Conrad, 2013, p. 3). The primary assumption of register studies is that linguistic features are functional and, as a result, particular features are commonly used in association with the communicative purposes and situational context of texts (Biber & Conrad, 2009, p. 2). Therefore, the linguistic choices that give rise to the linguistic differences across registers, and consequent variation among them, are described as highly systematic and functional because they are linked to particular non-linguistic factors. This links up with the sociolinguistic concept of interaction between extra-linguistic (social and situational) and linguistic forces (Wasserman, 2014b, p. 67).

Registers can be classified and studied at different levels of generality (or specialisation) in relation to how they are specified situationally (Biber, 1995). At the broadest level, they are classified based on mode (written or spoken), with most studies that take this view focusing on the differences between written and spoken language, such as for example Biber (1988) and Biber, Johansson, Leech, Conrad, and Finnegan (1999). Such studies focus not only on synchronic differences between written and spoken language, but also on the changing ways in which written and spoken language interact, and the (historical or ongoing) evolution of written and spoken language (Leech et al., 2009; Leech, Smith, & Rayson, 2012; Mair, 2006). However, the classification of registers can also be highly specified and defined, and studied in terms of very specific, situational characteristics (for example, the methodology sections of academic articles) (Biber, 1995, p. 7). These types of studies focus on specific registers, such as newspaper writing or academic writing (either individually or through comparative investigations) to describe how registers differ and how they have changed and continue to change.

Two of the most important findings from register studies is related to the historical drift of written registers, on the one hand, towards more colloquial linguistic usage, and on the other, towards increased lexical density. As discussed in Section 1.2, these two broad trends of change are referred to as the processes of colloquialisation and densification, respectively. Colloquialisation is often described as one of the most noticeable linguistic developments in diachronic investigations of the English written registers (Biber & Gray, 2012), but Leech et al. (2009, p. 245) point out that “there are facets of grammar where anti-colloquialization – a movement further away from spoken English norms – appears to be ascendant”. Thus, colloquialisation, which is the movement towards more oral styles, is counterweighted by trends associated with a movement towards a more written style (Kruger & Smith, 2018). Densification is one type of trend that is regarded as a counterweight to colloquialisation (Biber & Gray, 2012; Kruger & Smith, 2018).

Early work suggested that these broad trends applied to all registers generally; however, recent work has shown that the generalisation of these trends of change to the English language as a whole is problematic, because language is used in many different situational contexts and for vastly different communicative purposes. This highlights the fact that variation in language can occur on different planes, such as variation in registers as well as the cross-varietal variation referred to in Section 2.4, and shows how the broader trends of change play out differentially across registers and varieties. For example, Collins (2012) shows that Inner Circle varieties of English, such as American English (henceforth AmE) and AusE, are

significantly more advanced in taking up features associated with colloquialisation than Outer Circle varieties of English, such as Hong Kong English (henceforth HkE).

In terms of the written registers, Biber and Gray (2012) argue that trends like colloquialisation and densification play out in different ways across different registers, linked to particular situational characteristics of the registers in questions. The drift of registers towards more colloquial forms is due to the development of mass literacy and near-universal education, which has opened up certain registers to wider and more diverse audiences. In contrast, densification of information is viewed as a consequence of the information explosion, and has led to the need to convey more information in the same (or less) amount of text (Biber & Gray, 2012; Leech et al., 2009). Therefore, communicative purpose and intended audience are crucial factors in understanding the tension between the broader trends of change, relative to specific register.

Different written published registers (which are of interest in this thesis) display divergent behaviour in terms of their openness to features associated with these ongoing trends of language change (Biber & Gray, 2012, 2013; Hundt & Mair, 1999). The degree to which a particular register is receptive to ongoing change also extends beyond trends like colloquialisation and densification, and includes innovations associated with language contact. Hundt and Mair (1999) state that because innovations in a language arise and spread at differential speeds through various registers (and also different varieties), registers can be described on a cline of openness to innovative forms, and may therefore be seen as ‘agile’ or ‘uptight’ in relation to their acceptance of these features. Biber and Gray (2012) add that the description of registers as either innovative or conservative is complex, particularly because different registers participate in changes in different ways, with some written registers actually serving as sources of change, relative to their underlying functional motivations. Support for this can be found in several studies, such as the study by Biber (2012), which demonstrates differences in the patterns of linguistic variation and use between face-to-face conversation and academic writing, or Biber and Gray (2012), who find differences in the dissemination of features associated with colloquialisation and densification among newswriting, academic prose and popular writing. Specifically, Biber and Gray (2012) find that academic writing, which serves an informational function but which is read by a specialist audience, is particularly receptive to the features associated with densification but has largely resisted colloquialisation. News writing, which also serves an informational function but which has come to be read by a very wide audience, is especially receptive to features associated with densification and (less strongly) receptive to the features associated with colloquialisation. Fiction has resisted features associated with densification, but has been the most receptive to the features associated with colloquialisation (Biber & Gray, 2012).

Taking into account the work of Biber (2012), Biber and Conrad (2009, 2013), Biber and Gray (2012, 2013) and Hundt and Mair (1999), one dimension of written published registers that has not received much attention in studies of register variation (and also the variation among different varieties of English) is the role of editorial intervention in shaping the language of published written texts. Kruger and Van Rooy

(2017) point out that all written published texts will, at some point during their production processes, have been subjected to editorial processes that are usually guided by the style guides, house styles or style manuals of the publishing organisation. The final versions of these texts are therefore the products of both writing and professional editing processes (Nelson, Wallis, & Aarts, 2002), and reflect the linguistic contributions of their writers, but also their editors. Since the contributions by editors shape the language of these published written texts, which play an important role in propagating the features associated with change, it is necessary to investigate the role of editorial intervention in shaping the language of published written texts more broadly. Furthermore, it is necessary to investigate how this intervention forms part of the processes of ongoing language variation and change, and within the context of WEs, how this intervention helps to move a variety towards endonormativity.

2.6 Editorial intervention, norms and (endo)normativity

As demonstrated in the preceding sections, linguistic norms are central to cognitive and social theories of language use (see Sections 2.2.1 and 2.2.2), theories of language change (see Section 2.3) and models of World Englishes (see Section 2.4). The discussions in these sections highlight the complexity of the notion of linguistic norms, which is construed in different ways from within different theoretical frameworks. This study is interested in the way the different theories of language use, language change and World Englishes intersect, to provide a framework within which to investigate the role of editorial intervention, not only in processes of ongoing change, but specifically in how these processes unfold in different varieties of English and how these processes help to drive the varieties of English towards endonormativity. The work of editors centres on linguistic norms, and so before moving on to a discussion of editorial work, it is first necessary to weave together the different notions of norms as set out in the different theories discussed so as to provide a firm basis from which to explain the effects of editorial intervention.

2.6.1 Norms and normativity

In the cognitive linguistic perspective (see Section 2.2.1), a norm is understood at the level of the individual, and is viewed as the cognitive entrenchment of features in the minds of users as a result of exposure to these features in usage experiences. These usage experiences give rise to the individual cognitive representations of the norms of a language. In this sense, a norm is the product of the cognitive processes that help to make sense of the linguistic usage events that people encounter in their everyday lives. Croft (2000) refers to this as competence and defines it as the knowledge that a language user has of the conventions of a language community (see Section 2.3). This definition links the individual cognitive representations of language to the social dimensions of language use.

In the social view of linguistic norms, norms are understood at the level of the speech community (see Section 2.2.2). They are considered norms because they constitute a regularity of behaviour that forms the common ground of a community and which is used as a coordination device among users to communicate meaning (see Section 2.3). Croft (2000) refers to this as a convention. However, the notion of a linguistic

convention is multiplex and includes the distinction between overt prestige norms (the norms that are usually proclaimed by authorised institutions, and which are ‘top-down’), and covert solidarity norms (the norms that emerge in communicative acts and which are ‘bottom-up’). As discussed in Sections 2.2, 2.3 and 2.4, the tension between overt prestige norms and covert solidarity norms, and individual cognitive norms and communal norms plays a central role in language variation and change more generally, and in the development of PCEs specifically.

The process of change starts when a new usage is introduced into the language either as a result of a deviation from an established norm or when there is a shift in the frequency of use of an existing variant (at the expense of other competing variants) (see Section 2.3). This process is a fundamentally individual psycholinguistic process (Van Rooy, 2011). If this new usage diffuses through the community (that is, if more and more members of the community use it) then it spreads and becomes a communal norm that emerges through use. It thus becomes conventionalised. Therefore, new norms are established through the process of conventionalisation, which is the diffusion of the innovative feature(s) through the community of language users

However, conventionalisation has many dimensions and is conceptualised in different ways in different theoretical frameworks. In World Englishes, for example, a further dimension is added: norms may be external to the variety or they may be internal to the variety (see Section 2.4). This additional dimension of normativity is captured in the notions of varieties as being norm-providing, norm-developing or norm-dependent (Kachru, 1985), or in the case of PCEs as varieties developing towards endonormativity (Schneider, 2007). As discussed in Section 2.4, as PCEs move through the stages of the Dynamic Model, they move away from an exonormative orientation and towards an endonormative orientation in which local norms emerge in contact settings.

This thesis is concerned with one particular aspect of endonormativity in World Englishes: the way in which editorial intervention interacts with multiple factors (both internal to a language as well as those external to a language) in the development of varieties of English towards endonormativity. The next section will therefore focus on editorial intervention, and specifically on the role of editorial intervention in the processes that give rise to new norms and which lead to endonormativity.

2.6.2 The nature of editorial intervention and the development of norms in PCEs

The endonormative status of a variety of English in the Dynamic Model is determined by several factors, including its widespread use, or conventionalisation in the community, and the acceptance of a new local norm in the community. There are many layers to the acceptance of innovative forms in the community, but it has been argued that acceptance by linguistic gatekeepers working in the local publishing industry of a new variety is one such form, and therefore acceptance by these gatekeepers provides a measure of endonormativity (Bamgbose, 1998; Kruger, forthcoming; Kruger & Van Rooy, 2017; Van Rooy, 2011). In certain PCE contexts the acceptance by the publishing industry is complicated by the fact that the

codification of new variety is not well established, which results in competition between covert solidarity norms and the extraterritorial English norms that are codified in different norm-providing sources (Kruger, forthcoming). In this regard, Kruger and Van Rooy (2017) argue that acceptance by the linguistic gatekeepers of a variety does two things: first, such acceptance legitimises the use of the feature, and second such acceptance feeds the feature back into the feature pool since it permits the use of the feature in writing, increasing the feature's frequency and chance of further selection and therefore facilitates the diffusion of the feature in the community. Together these two processes lead to the conventionalisation of the feature, in terms of both its widespread use in the community but also in terms of the legitimisation of this use by local norm-giving authorities.

One such authority is the corps of local editors who work on the written published forms of the language in the variety. As pointed out in Section 1.1, editorial intervention has not received much attention in investigations of the progression of varieties towards endonormativity and is usually regarded as an obstacle to change rather than as a mechanism integral to change. Observations in this regard abound, as demonstrated in the observation by Cameron that “copy and sub-editors are the foot-soldiers of hyperstandardization”, which is “the mania for imposing a rule on any conceivable point of usage, in a way that goes beyond an ordinary understanding of what is needed to ensure efficient communication” (Cameron, 2012, pp. 47, 53).

The underlying assumption in this view is that editors *consciously* remove from the texts that they edit the features that deviate from the overt norms sanctioned in the norm-providing sources that they use. There is no doubt that this is partly true; part of editors' work is to consciously match the text that they edit with usage sanctioned in normative sources. But editors' removal of these features must be a matter of degree since studies of published written texts have shown that innovative features or variants do actually occur in published, written texts that have undergone editorial intervention. Thus, the empirical evidence for the assumptions about the contribution of editors has not really been uncovered, which raises the question: Why do some features make it through the editing process, while others do not?

The answer to this may be found in the integration of social and cognitive factors within a usage-based approach to understanding language use (see Sections 2.2.1 and 2.2.2) as well as the realisation that editors do not actually consult a normative source for every single editorial decision that they make – this would be extremely time-consuming and unproductive. Instead, editors rely on their expertise in language, which is their entrenched knowledge of language, or competence (see Sections 2.2.1 and 2.3). This knowledge is built up in the same way as it is for other language users – through exposure to use – and is subject to the same cognitive processes, mechanisms and constraints that affect other language users, as discussed in Section 2.2.1. However, editors are also unique in comparison to other language users, as a consequence of their work experience and their frequent exposure to prescriptive overt norms, which also feeds into their knowledge of language. At the same time editors are exposed to and affected by the same sociolinguistic normative pressures of any other person in the language community, and therefore it could be argued that

editors' minds serve as an arena where it is possible to investigate the complex interplay between overt and covert norms. Kruger and Van Rooy (2017) argue that the degree of acceptability of a new feature by editors may therefore provide a measure against which to determine the endonormative stabilisation of different varieties of English.

Research on editing is generally delimited to the types of editorial tasks specific to different contexts and has not often focused on the actual changes that editors make to texts. Kruger (2017) points out that “measured against the description of what editors and revisers can (potentially) do in amending texts, there is little comprehensive evidence of what they do (typically) do”. Within the context of World Englishes, Mesthrie (2009) points that “[w]e rarely have information on the editing processes accompanying the written efforts cited in some World Englishes studies” (p. 274). It is therefore clear that in both the field of editing and World Englishes we know very little about what editors actually do when amending texts and how these changes may influence the language of texts. Furthermore, existing research on editing demonstrates that the nature of editorial work is highly variable, and although all written published texts undergo some degree of editorial intervention, this intervention varies in scope depending on the type of text and the context in which the text is produced (Kruger & Bevan-Dye, 2013; Law, 2011, 2014).

In general, editorial tasks and the dimensions of editorial work are classified into two broad categories: copyediting and substantive editing (Bisaillon, 2007; Butcher et al., 2006; Mackenzie, 2011; Mossop, 2014). Copyediting is widely viewed as a core component of the editorial process and is defined by Mackenzie (2011) and Mossop (2014) as the heart of the editorial process that comprises tasks that bring a document into conformance with pre-set rules, including rules related to spelling, punctuation, syntax and idiom, and good usage, as sanctioned in authoritative sources such as house styles, dictionaries, style manuals and usage guides. Thus, copyediting is normative in orientation. As the heart of the editorial process, copyediting is normally carried out on all texts prior to publication, and is therefore a dimension of editing that affects all written published registers.

Substantive editing is much more interventionist because it aims to improve the overall scope, level, length, presentation and organisation of a text (Butcher et al., 2006; Kruger, 2017). During substantive editing, the editorial focus is on the communicative success of the text from the perspective of the reader. Substantive editing is therefore regarded as fulfilling a communicative function (Bisaillon, 2007). Although substantive editing does not invoke the kinds of prescriptive rules that form the basis of copyediting, it still has a normative dimension because of the normative expectations associated with register and audience (Kruger & Van Rooy, 2017). Substantive editing may also be further sub-classified into different types. Mossop (2014), for example, distinguishes between stylistic and content editing. Stylistic editing is concerned with ensuring the readability and clarity of a text by tailoring the language to the audience and their use of the text, and smoothing the language to ensure it is suitable to the text type (Mossop, 2014). The editor therefore assesses the language in terms of not only the text's register and purpose, but also in terms of who the audience is, and what their needs and expectations of the text are. This may be a point at which editorial

intervention invokes covert solidarity norms. Content editing, on the other hand, is concerned with the coverage of the topic, and may sometimes necessitate additions to the text. The editor may thus be required to write or rewrite sections of text, and in so doing invoke their own normative preferences and writing conventions. However, not all texts are subject to the different kinds of substantive editing and this has been shown to be related to several factors: such as the importance of the reader for the profitability of the publication, the formality of the publishing environment and whether or not the editor is employed on an in-house or freelance basis, and the textual features and functions of the text (Law, 2014). For example, within the context of academic editing, editorial intervention is largely restricted to copyediting tasks (and some minor stylistic editing tasks), as a result of the fact that substantive revisions to academic work raise questions of ethics (Kruger & Bevan-Dye, 2010, 2013). However, substantive editing is a crucial component in the production of creative texts and also some types of instructional writing, such as school books (Law, 2014). Thus, not only is there variability in the kinds of editorial tasks carried out, but there is also variability in the extent to which these tasks are carried out across different registers.

Given the above, and framed within the key role of normativity in the World Englishes paradigm (as discussed in Section 2.4), copyediting, as a highly norm-dependent activity, may play a particularly important role in promoting or constraining the dissemination of innovative linguistic features in published written texts in varieties of English. The kinds of norm-providing sources invoked for this type of editing by editors in different varieties, and within the different editorial contexts, may provide some information on where editors turn to for their standards: do they draw on norm-providing sources external to the variety that they edit, or do they draw on norm-providing sources that reflect a norm local to the variety that they edit? The actual changes that editors make (or do not make) to texts compared to the recommendations of their selected norm-giving sources, may shed light on the influence of their covert norms (gained through personal exposure and experience) on the language of these texts. If, for example, it is found that the choices of editors reflect the communal usage patterns of the communities that they edit, then it may serve as evidence of endonormativity, and would provide evidence that editorial intervention plays an important role in a variety's progression towards endonormativity.

Editors are not unaware of linguistic variation and language change, nor of their role in this. For example, Mossop (2014) states that once linguistic innovations begin to spread, editors need to decide whether to accept these innovations or to edit them out of texts. In a similar vein, Mackenzie (2011) notes that these decisions are tough for editors, especially because certain language rules are equivocal and because correct usage is often a matter of debate. Despite the generalised views of editorial intervention as fulfilling a primarily normative function, editorial work also has a communicative dimension in which editors are required to reflect on the communicative function and effectiveness of the texts that they edit, and by implication how language is used by the readers of the text. This is most saliently demonstrated for the World Englishes context in the following editing guideline that advises editors to be aware of the distinctions between the norms associated with Inner Circle varieties and their impracticality in Outer Circle contexts:

[A]s the number of people who write in English as their second language increases, editors may become less fussy about correctness because these writers, being members of other cultures, will not have any particular allegiance to traditions of correctness; they will be concerned only with communicative effectiveness. This will of course also be true for the constantly increasing number of readers of English who are not native speakers. (Mossop, 2014, p. 57)

Editors are clearly faced with and aware of the challenges associated with resolving the issue of what constitutes correct or appropriate usage when making editorial decisions. While these editorial decisions are guided by overt normative recommendations, editorial choices also invoke the covert norms that emerge in the context of PCEs. Furthermore, while some editorial choices are indeed *conscious* and based on clear prescriptive guidelines, the nature of the publishing environment and the time pressures under which editors work, also means that editors do not consult norm-providing sources for every editorial change that they make. Editors thus also rely on their own competence, which is their individual cognitive representations of the norms of a language built up through exposure to language use in the communities in which they move (see Section 2.2 and 2.3), the unedited texts that they edit as well as their exposure to overt norms sanctioned in norm providing sources. Thus, if the texts that editors are exposed to within the context of the variety within which they work are more heterogenous (versus more homogenous) and not necessarily aligned with their own linguistic background, then this opens up opportunities for interesting shifts in linguistic representations for editors these contexts.

2.7 Conclusion

The first research question of this study (as set out in Section 1.3) seeks to determine how broad trends of language change, the different contextual forces of the varieties of English, register effects and editorial practice interact to influence language variation and change. To answer this question, this chapter has developed an innovative theoretical framework in which to position editorial intervention within processes of language variation and change. To do this, the discussion in this chapter has drawn on overarching linguistic theories of language use, and in particular usage-based theories (see Section 2.2). It has focused on the interaction between cognitive and social factors in language structure and language use, which provides the scaffolding for understanding processes of language variation and language change. The discussion of different evolutionary theories of language change in Section 2.3 highlighted how cognitive and social factors conspire to shape the processes by means of which variation arises in a language and through which change proceeds. These discussions laid the ground work for an exploration of how language use and language variation and change may be understood within the development of the varieties of World Englishes and the very important role that written language and written registers play in the processes of variation and ongoing change (as discussed in Sections 2.4 and 2.5). The last section of this chapter synthesised the main points of the various theories and models and positioned the role of editors within a usage-based approach that integrates cognitive and social perspectives for understanding language structure and variation within the context of World Englishes. This chapter has argued that the point of intersection

for the inclusion of editorial intervention in the different theories and models is to be found in the crucial role of norms and conventions in language use and language change.

Following Kruger and Van Rooy (2017), and drawing on the theoretical framework developed in this chapter, Section 2.6 argued for the importance of editorial intervention in the dissemination and conventionalisation of the features associated with ongoing change in different written published registers in different varieties of English. This chapter argued that in contrast to the traditional assumption that editorial work is a process in which all editorial choices are driven by conscious decision-making processes rooted in mindlessly applying the rules set out in overt norm-providing sources, editors very often rely on their own expertise in language (for reasons of time and productivity, but also for ensuring the communicative effectiveness of texts). This expertise is built up through exposure to use, including exposure to overt norm-providing sources as well as communal patterns of use. As such it was shown that editors' minds serve as an arena where it is possible to investigate the complex interplay between overt and covert norms.

The chapters that follow explore this and seek empirical evidence to test if the interaction of overt and covert norms is reflected in the actual choices that editors make in the texts that they edit. The next chapter, Chapter 3, sets out a discussion of the varieties, registers and linguistic feature investigated in the case study approach used to explore the role of editorial intervention in influencing ongoing language variation and change in different registers in two varieties of English: AusE and SAfE.

Chapter 3

Varieties, registers and grammatical feature investigated

3.1 Introduction

The aim of this thesis is to investigate how editorial intervention interacts with processes of language change in different varieties of English and across different register. Since it is not possible to investigate this broad question for all features associated with ongoing change in all varieties of English and across the range of texts that are edited within the scope of this thesis, some decisions for narrowing down the focus had to be made.

This chapter sets out a discussion of the varieties, registers and the grammatical feature selected for investigation in this study and provides a motivation for their selection. Section 3.2 explores the development and current state of the two varieties of English investigated (AusE and SAfE), with particular emphasis also on the contexts of editorial practice in the two varieties, while Section 3.3 outlines the registers investigated. Thereafter, Section 3.4 discusses the genitive alternation, the grammatical feature selected for analysis. Section 3.5 provides a motivation for the selection of these varieties, registers, and grammatical feature. The chapter is concluded in Section 3.6.

3.2 Varieties under investigation

Two varieties of English are selected for investigation in this study, namely AusE and SAfE. The two varieties were settled at around the same time and by similar founder populations from the British Isles, progressing through Phase 1 and 2 of the Dynamic Model within a similar period (Schneider, 2007). However, AusE progressed more quickly through Phase 3 and 4 of the Dynamic Model (and is currently entering Phase 5), compared to SAfE, which progressed to Phase 3 of the Dynamic Model some 50 years after Australia, and which has only recently started making inroads into Phase 4. Furthermore, given the particular circumstances under which English in South Africa has evolved and continues to evolve, the variety is unique in respect of its internal heterogeneity: several sub-varieties are developing within the same context and in close contact with one another.

In the two sub-sections that follow (Sections 3.2.1 and 3.2.2), the formation of the two varieties and their progression along the stages of the Dynamic Model are discussed before zooming in on the current position and status of English(es) in the two countries. The two sections each conclude by shifting attention to the potential role of editorial intervention in the progression towards endonormativity of the two varieties. Sections 3.2.1 and 3.2.2 therefore serve to provide a justification for the inclusion of the two varieties in this study.

3.2.1 Australian English

English is the dominant language in Australia and is used by various community groups in the country including native-born non-Aboriginal Australians, indigenous Aboriginal communities, and immigrants of non-English backgrounds (Collins, 2014). As a result of these multiple English-speaking communities, the definition of Australian English is sometimes problematic. Collins (2014) distinguishes between ‘English in Australia’ and ‘mainstream Australian English’ (or AusE), in which the former represents an inclusive term for all the social, regional and ethnic varieties used in Australia, while the latter refers to the variety rooted in the levelling of the dialects of the early settlers (which included dialects from all over the British Isles, most notably the south-eastern parts of Britain and Ireland) and which is recognised as the national language. In this study, AusE is used in the same sense as Collins (2014).

Within the context of World Englishes, AusE is recognised as a major variety of English (an ENL or Inner Circle variety) that has consolidated its own norms (incorporating distinctive linguistic characteristics of phonology, lexis, syntax, grammar and idiom) and emerged as an independent national standard and carrier of Australia’s national identity (Schneider, 2007). The variety has its roots in the evolution of a set of transplanted dialects of English that interacted with various indigenous and migrant languages, AmE and other varieties of English (Collins, 2014), and therefore represents the STL strand of English in Australia in the Dynamic Model. According to Schneider (2007), the endonormative stabilisation of AusE occurred during the last half of the twentieth century, when many commentators noted the remarkable homogeneity of the variety. However, recently, a great deal of highly structured variation has been noted in AusE, and divergent patterns of usage have been identified for Australians of varying socio-economic, gender, age, regional and ethnic backgrounds (Collins, 2014). The variety thus shows traces of moving into Phase 5 of the Dynamic Model since the 1980s (Schneider, 2007).

Present-day AusE is characterised by social and regional differentiation, contact with various adstrates and the emergence of the IDG strand, Aboriginal English (henceforth AborE) (Peters, 2014a, pp. 108, 121). Peters (2014a) notes that a large number of immigrant communities of non-English backgrounds have brought a number of languages into contact with AusE. This has given rise to various adstrate varieties of English in the country, but substantial shifts from first language usage are recorded in the second generation (Peters, 2014a). These immigrant adstrates are noted as being localised in time and place, and are unstable due to the immigrant communities’ shifts from bilingualism to English (Peters, 2014a). The contribution of these adstrates to the differentiation of AusE is therefore only temporary, as language shift towards English takes place in successive generations in these immigrant communities. Peters (2014a, p. 121) adds that AborE, which serves an important identity function among Aborigines, is likely to be a more significant aspect of AusE differentiation. For this reason, the discussion of the progression of AusE through the various stages of the Dynamic Model in this section focuses only on mainstream AusE and AborE (referring to the various adstrates when necessary).

The first British contact with Australia occurred in 1770 when James Cook sailed to the east coast of the continent. Eight years later the first British fleet arrived in Botany Bay in New South Wales and established the first penal colony, with the transportation of convicts to Australia continuing up to the mid-nineteenth century. The early settlers mostly included convicts and the prison warders (and their families) who watched over them; however, a number of free settlers from different social groups also started arriving, and class boundaries arose (Schneider, 2007). During this stage, dialect mixing and koinéisation occurred and two sociostylistic dialects emerged: a middle-class educated form and a vernacular working-class form (Schneider, 2007). These settlers were from all over the British Isles and included dialects from the south-eastern parts of England as well as Irish, Scottish and Welsh dialects (Leitner, 2004). The majority of these settlers were from the south-east of England, and it is widely agreed that the dialects of these settlers provided the main input for the development of the dialect of AusE that emerged from the ‘Sydney mixing bowl’ (Cox & Palethorpe, 2012; Kiesling, 2008).

The early settlers encountered a linguistically and culturally diverse indigenous population and although early contact tended to be restricted and peaceful, the increasing number of settlers to come into the country from 1830–1850 increased contact between the settlers and the indigenous populations, with calamitous effects for the indigenous population. The first effect on the indigenous population was their being driven off their land through the internal migration of the settler population outward to the west and south from 1820–1850 and the further internal migration and immigration of fortune-seekers in the gold rush from 1850–1901 (Schneider, 2007). The second effect was their near decimation through violence and exposure to the germs of the settlers, against which they had no resistance (Cox & Palethorpe, 2012; Schneider, 2007). The increase in the population of settlers, their regional expansion, and the granting of regional autonomy helped to stabilise the colony and the exonormative orientation of the variety (Kiesling, 2008). While contact with the indigenous population was not positive, Aborigines came into increasing contact with the settlers via employment, and later via education through the establishment of Christian missions and schools, leading to the development of ethnolectal forms of English among the Aborigines (Schneider, 2007).

In 1901, after a series of conventions and referendums motivated largely by an increasing sense of national pride and growing political and economic power among Australians, the former six colonies were federated into the Commonwealth of Australia, marking the onset of Phase 3 of the Dynamic Model, Nativisation (Korhonen, 2017; Schneider, 2007). Despite this, the country maintained close political, economic and cultural ties with Britain, and while nativisation was taking place and local linguistic forms in pronunciation, lexis and grammar started to develop, the external British norm remained unchallenged in formal contexts, reflecting the fact that a fully-fledged Australian national identity had not yet formed (Korhonen, 2017; Schneider, 2007). The marginalisation of the indigenous population continued during this time; however, regular contact between the settlers and Aborigines became normal, leading to bilingualism among the indigenous population and the emergence of AborE (Schneider, 2007). According

to Peters (2014a), AborE was used by Aborigines as a form of communication with the settlers as well as with other Aboriginal communities who spoke unfamiliar languages. This, coupled with the mobility of the Aborigines, served to spread AborE.

The endonormative stabilisation of AusE is linked to the period between 1942 and 1980, when the events of WWII acted as a catalyst for the severing of the affiliation and identity of Australia from Britain, and triggered the move towards endonormativity (Peters, 2014a; Schneider, 2007). The cultural changes in Australia during the latter half of the twentieth century were both audible and visible, and reflected in the Australian Broadcasting Commission's (ABC) move away from hiring only British news readers and the localisation of the Australian publishing landscape, which was helped by the first publication of the *Australian Government Style Manual* in 1966 (Peters, 2014a). With the publication of the *Style Manual*, a local reference on the written standard became available and was taken up by government and commercial publishers. In this burgeoning local publishing landscape, linguistic norms embodying AusE found their way into locally published texts, ultimately leading to the publication of further key linguistic reference works for AusE in the 1980s, such as the *Macquarie Dictionary* (1981), the *Australian National Dictionary* (1988), and the first AusE usage guide *Right Words* (Murray-Smith, 1987). With their expressly AusE content, these sources all served as instruments of codification and clear markers of the endonormative stabilisation of the variety.

In present-day Australia, AborE is the most marked of the ethnic varieties in the country (Peters, 2014a). Contemporary AborE is regarded as a variety that developed in high contact with English in the southern and central parts of the country (Peters, 2014a, p. 116). In the latter half of the mid-twentieth century, several events (such as Aborigines being allowed to vote in the federal elections in 1962 and the establishment of the Local Land Councils from the 1970s), which can collectively be viewed as Event X for AborE, helped to solidify the self-identity and solidarity of Aborigines. As a result, AborE is now a relatively homogenous ethnolect in Australia, that enjoys strong ethnic and cultural support in Australia\.

AusE is well-known for its persistent and pervasive informality, and its openness to the features associated with colloquialisation (Collins, 2014; Peters, Collins, & Smith, 2009). Peters (2009) specifically notes that AusE distinguishes itself for its tolerance for the features associated with colloquialisation and its willingness to employ them in writing, adding that the effect is to reduce register differentiation in the variety. Collins (2012) shows that AusE is one of the most advanced varieties of English in taking up features associated with colloquialisation, lagging only slightly behind the most advanced variety in taking, AmE. Another of AusE's most distinctive characteristics is its fluid heterogeneity in which a range of styles, registers and social varieties merge and give rise to structured variation (Collins, 2014, p. 449). AusE has therefore reached Phase 5 of the Dynamic Model; however, the ongoing differentiation in AusE is more closely related to sociolectal differences (conditioned by age, education and gender) in phonology and lexicogrammar, than it is to regional differences (Collins, 2014; Peters, 2014a). Peters (2009) states that having consolidated its own norms and forged its own identity, internal differentiation in AusE registers

and modes of discourse may emerge as part of the (normal) processes of ongoing change and the linguistic changes in the relationships between registers in modern English. Evidence for this in AusE is demonstrated in a series of contributions in Peters et al. (2009) that show how AusE patterns with the major northern hemisphere varieties, but also how it diverges from them.

The role of editorial intervention in the processes of ongoing change and the linguistic changes in the relationships between registers in AusE has, to my knowledge, not been explicitly investigated. As noted earlier, AusE is particularly receptive to the features associated with colloquialisation, and this might be a particular point at which editorial intervention plays an important role. One contribution in Peters et al. (2009) that points to the potential influence of editorial intervention in distinguishing AusE from other varieties is a study by Peterson (2009) on the use of three connective adverbs (*however, therefore, and thus*) to introduce a second main clause within a single orthographic sentence in current written English. Peterson (2009) compares the frequency of these ‘run-on sentences’ across unedited written registers and written published registers in AmE, British English (henceforth BrE), NZE and AusE⁵ and finds that the use of these ‘run-on sentences’ is much higher in AusE and NZE, compared to AmE and BrE. He finds that these sentences are much more frequent in unedited writing compared to edited writing, but notes that in terms of cross-varietal differences, the run-on sentences occur much more frequently in written published AusE and NZE, than in AmE and BrE, suggesting that the feature enjoys some acceptance by editors of AusE (and NZE).

Given that AusE has reached Phase 5 of the Dynamic Model, and given the fact that AusE is the dominant strand used in the country, it is to be expected that the linguistic profiles of editors and authors of English in this country are closely aligned, and that these covert representations are closely aligned to the overt AusE norms. Thus, editors in AusE work within a focused normative environment in which their editorial choices are most likely reflective of the covert norms of the authors that they edit.

3.2.2 South African English

South Africa is a linguistically diverse country in which 11 languages enjoy official status, as accorded by the constitution. Apart from English, these official languages are Afrikaans, seven languages from the two major groups of South African Bantu languages (Nguni, which includes Zulu, Xhosa, Swati and Ndebele; and Sotho, which includes Southern Sotho, Tswana and Northern Sotho), as well as Tsonga and Venda. English, which is used as a home language by 9,6% of the population (Statistics South Africa, 2012, p. 24), is by no means a numerically dominant language in South Africa; however, it is the most widely used language for written communication in high functions, including the media, publishing, education and government. The fact that South Africa has 11 official languages, coupled with the fact that English is the dominant language in formal functions, speaks to the wide-spread multilingualism in the country most often

⁵ Peterson (2009) extracts the data for his study from the Australian Corpus of English (ACE), the Wellington Corpus of Written New Zealand English (WWC), F-LOB and Frown and the ICE-Aus and ICE-NZ corpora.

studied within the context of sociolinguistic investigations (Mesthrie, 2010), and points to the varying use of English as first language and second language by different groups.

According to Van Rooy (2014b), English arrived late on the South African linguistic scene and followed the arrival of various Bantu languages in the last two millennia and the arrival of Dutch in the mid-seventeenth century. English in South Africa has therefore existed in a complex and diverse multilingual and multicultural environment since being brought to the country in the nineteenth century by British settlers after the annexation of the Cape from the Dutch by the British in 1806 (Silva, 1997). Over the course of its history in South Africa, English has come into contact with and come to be used by users of other languages, which has resulted in the development of multiple IDG strands of English (such as AfrE and BSAfE) alongside the STL strand. Although other strands have also emerged, such as South African Indian English, only three are investigated in this study: WSAfE, the native variety used by the white descendants of the original settler population(s) and which represents the STL strand; and two IDG strands, AfrE, a non-native variety used by white first-language users of Afrikaans and BSAfE, a mostly non-native variety used by the majority of the black South African population.

In plotting the development of SAfE along the five stages of the Dynamic Model, Schneider (2007) notes that the greatest challenge is the number of distinct speech communities that have entered the South African arena and interacted with one another at different points in time and under varying social circumstances. Perhaps the most important of these social conditions is the implementation of segregation policies in the twentieth century under the policy of apartheid that served to keep different ethnic groups apart, limiting contact among the different strands of English. This has had a considerable effect on the progression of the different sub-varieties along the stages of the Dynamic Model (Van Rooy, 2014a).

Schneider (2007) argues that SAfE has made deep inroads into Phase 4, Endonormative Stabilisation, of his Dynamic Model. Central to his motivation for placing the variety at this phase is the emergence of the 'rainbow nation' in which the collective, unified identity of all South Africans as members of a newly forged nation was emphasised following the first free government elections in 1994 that brought Nelson Mandela and the ANC into power. Despite this, different ethnic and social identities persist (which contributes to maintaining the distinctive sub-varieties). A further key characteristic of varieties that have entered Phase 4 of the Dynamic Model is the development of a local norm, which is, on the one hand, reflected in a growing acceptance of local usages and convergence between the different strands, and on the other, in the codification of the newly emerged norm in local dictionaries and grammars. With regard to the former, Schneider (2007) states that in South Africa some degree of convergence has been found for pronunciation. However, some scholars working on SAfE have demonstrated that there is a tolerance for a range of lexical features (Van Rooy & Terblanche, 2010), and elements of grammar (Kruger & Van Rooy, 2016, 2017; Spencer, 2011b; Wasserman, 2014a, 2014b) among the different sub-varieties used in the country. The publication of two substantial scholarly dictionaries of SAfE in the early and mid-1990s is, according to Schneider (2007, p. 188), a decisive step towards codification.

At Phase 4, the newly emerged variety is typically perceived of as homogenous, which is largely due to the convergence of the different STL and IDG strands, and while some linguistic heterogeneity will remain, mostly along ethnic and social lines, this will largely be downplayed (Schneider, 2007). In other words, as the different strands in a postcolonial language contact setting come into increasing and prolonged contact, the identities of the two groups are gradually realigned and brought closer together. The effect of this on the language of the groups is that they will be reshaped, until eventually the strands of the two groups converge and a new variety with a new local norm emerges and comes to be accepted. In this regard, Schneider (2007) concedes that converging tendencies among the different strands have been weaker in SAfE than in other PCEs, and that homogeneity is limited. Van Rooy (2014a, p. 35) confirms Schneider's (2007) assertion that ethnic and social identities persist, but takes the stronger position that the common national identity is largely symbolic and does not yet permeate the South African society – coming to the fore only at major (sporting) events (Spencer, 2011a). Van Rooy (2014a) adds that in countries where multiple, complex contact settings arise, convergence and homogeneity (and therefore endonormative stabilisation) should be investigated relative to the local settings for each sub-variety within the country, rather than for the country as a whole. Therefore, analysing the local contact settings of such countries will help to explain the lower degree of convergence and homogeneity at Phase 4 of the model than is found for countries where the contact settings are limited to an overarching one (Van Rooy, 2014a, p. 22).

With due consideration for the argument of Van Rooy (2014a), the formation of SAfE will be briefly sketched by zooming in on the contact settings that arose in South Africa with the arrival of the British settlers in the nineteenth century, so as to contextualise the development and current position of the different sub-varieties of SAfE investigated in this study.

The first wave of British settlers, who were primarily middle-class citizens from respectable backgrounds from the home counties in England, arrived in the Eastern Cape in 1820 (Lanham, 1996; Van Rooy, 2014b). Several frontier wars with the Xhosa left the descendants of these settlers impoverished, and as contact with Britain decreased, social levelling produced a homogenous dialect within the first two generations (Lanham, 1996, p. 21). Extensive contact with the Dutch (at that stage users of a form of Proto-Afrikaans (Bekker, 2012, p. 141)) who had been settled in the Cape for some 150 years, and to a lesser extent the Xhosa, introduced word and expression loans from Afrikaans into the English of these settlers (Branford & Cloughton, 1995). Some traces of nativisation can be found in the English of the 1820 settlers, mostly due to contact with Afrikaans (Van Rooy, 2014a, p. 26). Lanham (1996, p. 21) reports that British travellers to South Africa found the English of the settlers difficult to understand.

The second wave of British settlers, who were mostly of middle- or upper-class backgrounds from the northern counties of Yorkshire and Lancashire, settled in the Natal region from the 1840s (Lanham, 1996). This second wave of settlers, unlike those of the first wave, lived in more densely populated areas with access to urban centres, did not face any frontier wars, were not impoverished and had much less contact with Afrikaans. All of this served to maintain social class distinctions and led to the development of a

distinctive dialect that remained closer to the British input and was considered ‘purer’ than that of the first wave settlers in the Cape. This dialect inevitably came to be associated with social prestige (Lanham, 1996, p. 21). According to Van Rooy (2014a), before any significant nativisation could occur in the English dialect of the Natal settlers, a third wave of immigration and country-internal migration led to the establishment of significant new English settlements.

During the third wave, the discovery of diamonds in Kimberly in 1870 and of gold in the Witwatersrand in 1886 brought large groups of immigrants from all over England and Eastern and Western Europe, as well as the Afrikaners and settlers from the Cape and Natal colonies, to these areas, and densely populated industrial-mining cities emerged (Lanham, 1996, p. 23). In Johannesburg on the Witwatersrand in particular, Bekker (2012, p. 141) argues that a new SAfE dialect, or what is today known as WSAfE, was shaped by the input from the various dialects of the Cape and Natal settlers and from Britain, as well as the various non-native users, such as the English of Afrikaans first-language users. The output of this formation process resulted in a sociolectal continuum in which three sociolects in a class-stratified society emerged: Cultivated WSAfE, which had an external British reference and which was held in high esteem; the more local General WSAfE; and Broad WSAfE which was typically associated with attributes such as being manly, tough and patriotic and which was hardly distinguishable from AfrE (Bekker, 2012, pp. 141-142; Lanham, 1996, p. 24). During the twentieth century these sociolects dispersed, and General WSAfE spread at the expense of Cultivated and Broad WSAfE, so that what is known today as the native variety of SAfE (or WSAfE) has its roots in General WSAfE (Bekker, 2012, p. 142).

According to Van Rooy (2014a), Afrikaans users played a significant role in the nativisation of the new dialect in the 20th century. This role was made possible by increased contact between the STL strand and Afrikaans after the Act of the Union of 1910 and the adoption of an official bilingual language policy that sought to promote a better relationship among the Afrikaners and the English after the Second Anglo-Boer War (Lanham, 1996, p. 25; Van Rooy, 2014a, p. 27). During this time, the English retained economic power, which resulted in increasing Afrikaans–English bilingualism among groups of Afrikaners who lived in urban areas and who entered civil service (Lanham, 1996). Despite the bilingual language policy of the Act of the Union of 1920, English remained the dominant language in both education, and to a lesser extent, in radio. This resulted in a feeling amongst the Afrikaners that they were socially and economically disadvantaged, which, fuelled by the discrimination of the English, led to feelings of resentment and hostility towards English (Lanham, 1996; Van Rooy, 2014a). These feelings of resentment were actively promoted by Afrikaner nationalist organisations, most notably the Afrikaner *Broederbond* and the National Party (Lanham, 1996; Minow, 2010).

In 1948, one of these Afrikaner nationalist organisations, the National Party, came into political power “with avowed intentions to redress the subordinate role of Afrikaans in society” (Lanham, 1996, p. 26). According to Branford (1996, p. 39), this translated to the steady extension of the functions of Afrikaans in areas of religion, politics, culture and education, and led to the codification of Afrikaans by the

Taalkommissie (the Language Commission) and the bureau of *Die Woordeboek van die Afrikaanse Taal* (The Dictionary of the Afrikaans Language). Furthermore, segregation practices were transformed into official apartheid policies, with the result that contact between white South Africans and other population groups was severely restricted (Giliomee, 2003). This led to increased contact and consequently convergence between AfrE and Broad WSAfE at the lower end of the socio-economic scale, while contact between WSAfE users at the upper end of the socio-economic scale and Afrikaans was limited, with most of these WSAfE users maintaining a strong exonormative orientation (Lanham & MacDonald, 1979). Towards the end of the twentieth century, evidence of convergence between the STL strand and AfrE could be found most notably at lexical level and in some respects at the grammatical level; however, Van Rooy (2014a) states explicitly that this convergence is not the result of the mutual acceptance of identity realignment between the two strands, but is due rather to the prolonged and continual contact between the two strands.

BSAfE holds the special position of being the most widely used strand of English in South Africa. Makalela (2004) argues that it is a dialect of English in its own right, in which the indigenous first languages have played an important role in the nativisation of BSAfE through cross-linguistic influence from the first language on the second language. Contact between the English settlers and black South Africans was initiated via the activities of missionaries, and English spread to black, mostly Zulu, South Africans via the slow establishment of missionary schools in the nineteenth century (Hirson, 1981). According to Van Rooy (2014a), this resulted in an extended timeframe for the formation of BSAfE. During the last quarter of the nineteenth century and at the beginning of the twentieth century, many of these missionary schools, which accounted for almost all education offered to black South African children via the medium of English at the time, were staffed by teachers who were highly qualified native users of English. As such, the education they provided was of a reasonably high standard and produced an elite generation of Black South African lawyers, university lecturers and newspaper editors (de Klerk, 1999; Minow, 2010). At the beginning of the twentieth century, the population of learners in the mission schools grew substantially, necessitating state assistance (Hirson, 1981). The growing number of mission schools and increasing population of black children receiving their education through these schools, coupled with a growing number of job opportunities in urban contexts facilitated further interaction with English users and enabled the slow spread of bilingualism among black South Africans (Hirson, 1981).

The election of the National Party to power in 1948 and their clear intention to advance Afrikaans in society, resulted firstly in enforced policies of segregation and secondly in the passing of the Bantu Education Act of 1953. The enforced segregation policies, as noted above, resulted in restricting even further any meaningful contact between white South Africans (and by implication the STL strand of English) and other population groups. Furthermore, the removal of white teachers from schools due to the implementation of the Bantu Education Act of 1953 and the entrenchment of mother-tongue instruction up to the highest possible schooling level for black children had devastating consequences on the black South African

population (de Klerk, 1999, p. 312; de Klerk & Gough, 2002, p. 357). The education policy of the National Party resulted in feelings of suspicion among black South Africans that the apartheid government was trying to create a semi-literate, isolated labour force (de Klerk & Gough, 2002, p. 357). As a result, English became extremely attractive to black South Africans because it was viewed as the key to socio-economic advancement and power, but also because it was used as the language of the liberation movement (de Klerk, 1999; de Klerk & Gough, 2002). Demand for English grew, specifically among black school children, culminating in the violent Soweto uprising of 1976, in which many school children were killed. This led to the replacement of the Bantu Education Act of 1953 with the Education and Training Act of 1979 in which schools were allowed (in consultation with parents) to determine the medium of instruction after the first four years of schooling (Minow, 2010, p. 16). English emerged as the overwhelming choice of medium of instruction. According to Van Rooy (2014a, p. 31), the linguistic effects of the social engineering practices and policies of the apartheid government effectively ruled out the possibility of stabilisation in BSAfE between 1950 and 1990, and placed the development of BSAfE on hold.

Since the political transformation of 1994, South African society has started shifting to an egalitarian society in which economic opportunities are evenly distributed across all population groups. Yet despite this, the country has still not achieved national reconciliation and has not managed to rebuild itself as one nation, with distinct social and ethnic identities persisting (Schneider, 2007; Van Rooy, 2014a). In a country where the constitution affords official status to 11 languages, English still remains an attractive option as a lingua franca for many South Africans, even if it acts as a barrier for those South Africans who do not command it and attitudes towards it vary (Schneider, 2007, p. 186). As a result of the persistence of various social and ethnic identities, several sub-varieties of English are widely used throughout the country, with no single sub-variety emerging as the carrier of South Africa's (symbolic) national identity. The STL strand of WSAfE is a visible and influential sub-variety, but it is only used by a minority of South Africans, with the overwhelming majority of English users being BSAfE users. As noted previously, Van Rooy (2014a) has argued that in countries with multiple, complex language contact settings, the conditions of convergence and homogeneity for a variety to enter Phase 4 of the Dynamic Model require reassessment. In this regard, and with specific reference to the South African situation, Van Rooy (2014a) argues that in South Africa after 1994, while the different dialects have to some extent stabilised, English might not be all that central to the process of national identity rewriting. In disentangling convergence and homogeneity, Van Rooy (2014a, p. 36) explicitly states that identity rewriting might not be a sufficient condition for convergence, and argues instead that prolonged contact might be more important. Evidence for this can be found in the fact that despite the absence of a common identity in South Africa, some convergence can be found in a growing tolerance for (and therefore acceptance of) a range of features at the level of pronunciation and lexis, while recent work also points to some convergence in grammar (Van Rooy, 2014b).

One aspect that might serve as a measure of endonormativity is the extent to which various linguistic gatekeepers, such as language teachers and editors of written published texts, regard particular features or

usages as acceptable (Kruger & Van Rooy, 2017, p. 27). With regard to written language specifically, several studies that have variously investigated the three stands of SAfE surveyed in this section, provide some evidence for emerging endonormativity in SAfE on the basis of this type of acceptability. However, the question of acceptability is considerably more complex in the South African context, compared to the Australian context, since many different sub-varieties of English are used in the country – and so the question of the acceptability of innovative usages is not limited to a particular sub-variety, but applies across the sub-varieties.

Spencer (2011b), for example, exemplifies this kind of (cross-)sub-varietal acceptability by investigating the acceptability ratings among a group of AfrE and WSAfE teacher-trainee students of entrenched features of BSAfE that do not conform to Standard English in academic writing (such as the extension of the progressive and the omission of the article). She finds acceptability ratings of up to 26% for some of the features she investigates and concludes that the initial level of tolerance for these features among the WSAfE and AfrE trainee teachers provides evidence that acceptance of BSAfE is gaining momentum. She further argues that because certain features of BSAfE are tolerated even in formal, written, academic contexts at university level and penetrating segments of the South African population outside of BSAfE, there is evidence of endonormativity (Spencer, 2011b, p. 143).

Van Rooy and Terblanche (2010) investigate convergence between WSAfE and BSAfE in terms of innovative lexical features in a corpus of texts from a newspaper that targets BSAfE and WSAfE users. They argue that the acceptability of innovative lexical features points to convergence and provides evidence of progression into Phase 4 of the Dynamic Model. In their study, acceptability is operationalised in terms of the use of the features beyond the innovating segment of the community, which they determine by searching for the innovations in newspapers whose readership include large sections of the STL strand in South Africa. They find evidence of linguistic convergence between WSAfE and BSAfE for some of the innovative lexical forms they identify, and argue that this is an indication of the emergence of endonormativity in SAfE. Van Rooy and Terblanche (2010) note that the texts in their corpus have been copyedited by an AfrE editor, but they do not consider that acceptance of the features on the part of the AfrE copyeditor (as reflected by the retention of lexical innovations in the published, and therefore edited, texts), might also point to convergence between AfrE in respect of the particular innovative lexical features they investigate.

Explicitly focusing on editorial acceptance, Kruger and Van Rooy (2017) propose a model in which editors' acceptance of the linguistic features of a new variety of English legitimises the use of the feature and aids the conventionalisation of the feature by allowing it to disseminate in published texts. They argue that viewed this way, editorial acceptance might simultaneously help to facilitate the progression of a variety to endonormativity while providing a measure of the level of endonormativity reached (two points that are central to the investigation of the role of editorial intervention in this thesis). In testing their model, Kruger and Van Rooy (2017) investigate editors' acceptance of the use of the progressive aspect in a corpus of

published written texts produced by BSAfE authors across four registers (academic, informative, persuasive and reportage). They find that while the innovative extension of the progressive to an on-going timeframe is conventionalised in BSAfE, as demonstrated in its near similar frequency to the prototypical temporary timeframe in unedited BSAfE writing, the extended use also enjoys tacit acceptance from a corps of local editors (who are not BSAfE users themselves). They conclude that this provides evidence of endonormativity and serves to further legitimise and entrench the feature in BSAfE.

Building on the work in Kruger and Van Rooy (2017), Kruger (forthcoming) investigates how editorial intervention reshapes the academic writing of WSAfE and BSAfE authors. Using an inductive, exploratory quantitative method, Kruger (forthcoming) identifies linguistic features that distinguish unedited BSAfE and WSAfE academic writing, and edited BSAfE and WSAfE academic writing. Focusing specifically on downtowners and possibility modals (which are identified as two features in which editorial intervention differs across the sub-varieties), she finds that editors mostly leave BSAfE usage patterns unaltered (which she argues provides support for endonormativity) but sometimes change WSAfE usage by bringing it closer to BrE (and therefore argues that support for endonormativity is somewhat ambiguous). Kruger (forthcoming) therefore finds evidence to show that editorial intervention does play a role in the progression of the two sub-varieties of SAfE towards endonormativity, but this does not suggest convergence between the two. These findings provide support for the assertion by Van Rooy (2014a) that in countries where multiple sub-varieties of English have developed in complex and differing contact settings, the analysis of the endonormative status of a sub-variety should be studied relative to the sub-variety, rather than from within an approach that absorbs the different sub-varieties represented into one assessment of overarching endonormativity.

The more heterogenous linguistic landscape in South Africa, coupled with the fact that most editors of English appear to be both WSAfE and AfrE users (see Section 1.2) but the authors of English texts represent the full spectrum of the sub-varieties, means that there is likely a stricter acceptability requirement for innovative usages among editors because acceptability must be attained across sub-varieties. Furthermore, the linguistic representations of editors of English in South Africa are likely highly varied as a result of their exposure to a diffuse set of covert norms.

3.2.3 AusE and SAfE: Endonormativity and editing in sum

The brief outline of the two varieties in Sections 3.2.1 and 3.2.2 show that AusE is considered to be a classic Inner Circle variety in which the STL strand has prevailed, while in South Africa the complex sociolinguistic constellation of a comparatively high number of distinct, compartmentalised speech communities that have entered the arena at different points in time, and which have interacted with each other under varying social circumstances, has given rise to a complex internal variation structure in which multiple sub-varieties are developing in contact with each other at different rates in the same context. Furthermore, AusE represents a variety that has been codified in several national dictionaries and other

sources like style manuals and usage guides, reflecting the local norm from which editors can draw. SAfE, on the other hand, has not enjoyed the same level of codification, and the set of norms governing the use of English in South Africa appears to be exonormative, with the orientation mostly towards BrE. Furthermore, the kinds of linguistic usage events that authors and editors are exposed to differ in the two countries as a result of the distinction between the more homogenous linguistic environment in Australia and the more heterogenous linguistic landscape in South Africa. The two varieties therefore each present a different case for the potential influence of editorial intervention in the progression of the variety towards endonormativity. Their selection in this study, along with the registers and feature investigated, allows for a meaningful comparative investigation of how editorial intervention plays out in different varieties of English (and within these varieties, across different registers), and what role this intervention plays in the movement of the varieties towards endonormativity.

3.3 Registers investigated in this study

As discussed in Section 2.5, in this study register is not defined based on mode of production (such as written or spoken language), nor is it defined on the more specified level of the parts or sections of texts. Instead, this study views a register as a variety of language that is defined in terms of a particular set of situational characteristics, which include communicative function, the relationship between the author and reader, and production circumstances. The selection of registers included in this study was guided by several considerations. The first consideration related to the primary aim of the study, which is to investigate editorial intervention. As such the registers investigated all needed to be written published texts that are subjected to professional editing as part of their production processes, and which are published through commercial or academic presses, including institutions of higher learning. The second consideration is couched in the fact that patterns of linguistic variation are known to be strikingly different across different registers, because different registers reflect ongoing language change in different ways, relative to their underlying functional motivations (see Section 2.5) (Biber & Gray, 2012; Hundt & Mair, 1999). The last consideration is related to the nature and degree of editorial intervention, which has been shown to differ across different types of texts and contexts (Law, 2014). Therefore, the selection of registers investigated in this study include a range of written published registers that are known to differ in their openness to the features associated with ongoing change, and that are subjected to varying levels and types of editorial intervention.

Five general written published registers were selected: academic, instructional, creative, popular and reportage. The categorisation of text types into the five different registers was based on the categorisation of printed written texts used in the International Corpus of English (ICE) corpora.⁶ These five registers can

⁶ The ICE corpus is a large collection of corpora representing different national and regional varieties of English. Each ICE corpus consists of approximately 1 million words of spoken and written English produced after 1989. More information on the ICE corpus can be found at <http://ice-corpora.net/ice/>.

be represented on a continuum in terms of their openness to innovations and their tendency to retain conservative forms (Hundt & Mair, 1999). The academic register represents the more formal end of the continuum, having participated in the historical shift toward a more colloquial style far less than other written registers (Biber & Gray, 2012). Reportage represents the opposite end of the continuum, having been shown to be particularly receptive to features associated with change (Hundt & Mair, 1999), with the instructional, popular and creative registers taking up the intermediate positions on the continuum (see Section 2.5 for further discussion). The discussion that follows sets out a description of each register, describing the communicative purpose and function of each, as well as how the nature and degree of editorial intervention is known to differ across each.

The primary function of academic texts, such as academic articles, dissertations and theses, and academic books, is to convey information. This informational function is specified by the particular purpose of academic texts, which is to contribute new knowledge to the academic field and to convince other academics by way of argumentation and explanation that this knowledge has scientific merit (Biber & Conrad, 2009). Academic texts are typically produced within the institutions of academe, presenting highly complex content to a very specialised audience. In the course of their historical evolution, academic texts have become highly conventionalised, and are characterised as extremely formal and dense, mostly due to the distant and impersonal social relationship that holds among the author and the specialised reader. Assertions with regard to the discourse style of academic texts differ. Biber and Gray (2010), for example, state that the stereotypical view of academic writing is that it is grammatically complex, drawing on elaborated structures and expressions in which meaning relations are explicit. But recent work has shown that academic registers have, in fact, developed a discourse style that is not structurally elaborated and is characterised by a dense, compact use of non-clausal, nominal phrasal structures (Biber & Gray, 2011). Biber and Gray (2010) argue that this discourse style is rooted in the ‘information explosion’ and the accompanying need to compact more information into fewer words (which, they argue, is highly efficient for the specialist academic audience) and changes in technology that facilitate revision and editing.

The kind of editorial intervention to be found in academic texts is shaped by the fact that academic texts deal with highly specialised content produced by highly knowledgeable authors whose work need not be edited for content, but which does undergo a form of language editing. Academic texts, such as books and articles, are published by various publishing institutions that have very specific expectations of the language, style and presentation of the content in their publications (conveyed through, for example, style guides and manuals). These types of text are therefore produced within ‘traditional’ publishing contexts. Dissertations and theses, on the other hand, are usually produced in non-formalised publishing contexts: they are written by postgraduate students who then submit their work for editing. In both cases, these types of texts are usually edited by freelance editors who are contacted by the publishing institution or student to edit the work.

A key characteristic of editorial intervention in academic texts is that it usually focuses on two dimensions: copyediting, in which the editorial focus is on ensuring that the text adheres to the preferred spelling, punctuation and usage conventions of the publishing or higher education institution; and stylistic editing, in which the editor ensures that the text adheres to the distinctive and conventionalised discourse style of the particular register. Editors therefore do not typically engage in substantive editing practices, such as removing or rewriting sections of texts, and consequently do not assume a writing function when editing academic texts (Kruger & Bevan-Dye, 2010, 2013). The expectation for the editing of academic texts has largely arisen due to the fact that, increasingly, many of the texts produced for this formal register, where strict expectations of standard usage and conventions are set, are produced by (highly educated) second-language writers of English (Kruger & Bevan-Dye, 2010, 2013). This opens up a space where the normative dimension of editing in relation to the acceptance of second-language innovations and usages is particularly salient, and therefore the kinds of choices that editors make when editing these texts might be a measure of what is regarded as acceptable in formal writing. The inclusion of both AusE and SAfE in this study creates the opportunity to investigate this dynamic. For example, as discussed in Section 3.2.2, in South Africa a number of second-language varieties are used alongside the first-language variety of WSAfE. Since SAfE has only been codified in a handful of dictionaries, and because the norm-orientation of this country is still largely directed towards the external British norm, it may be predicted that editors of academic texts in South Africa might make extensive changes to academic texts in order to bring texts produced by second-language writers in line with the perceived standard. The situation is different in Australia, where English is used as a first language by the majority of the population and where the variety has been codified in a number of norm-providing sources. It is therefore probable that the production of texts by AusE authors would be more closely aligned with AusE norms, requiring less extensive intervention on the part of editors. However, one exception to this may be the high number of international higher research students in Australia, most of whom come from Expanding Circle countries where English is acquired as a second-language for its value as an international lingua franca. For example, Universities Australia (2017) reports that the education of international students at Australian universities is the country's largest service export and that in 2016 international students accounted for just over a quarter of all enrolled students at Australian universities.

Instructional texts also serve an informational function, but they are distinguished from academic texts in their purpose and audience. Instructional texts seek to convey factual information in an instructive, rather than an argumentative style. Furthermore, the readers of instructional texts are a less specialised and more general audience, compared to the academic register where readers are experts with highly specialised knowledge of the particular topics of the texts. The types of texts included in the instructional register include regulatory or administrative texts typically produced within governmental or corporate contexts, instruction manuals, and educational materials such as textbooks, teachers' guides and learners' books. The register is therefore rather broad. What distinguishes the types of texts included in the instructional register is that they are non-fiction texts written with the purpose of conveying information. Regulatory and

administrative texts target a general readership, while skills texts and educational materials target a narrower audience, though not necessarily a more specialist audience than the academic register. Furthermore, all types of instructional texts attempt to facilitate the conveying of information and the ease with which the texts are understood by the reader, and therefore emphasise accessibility, clarity, and explicitness.

Because of the informational and instructive purpose of instructional texts, the nature and scope of editorial intervention in these texts is much more substantive than for academic editing. Furthermore, because the instructional register includes a broad range of text types, the channels through which these texts are published may differ. For example, educational materials are typically produced in traditional publishing contexts by educational publishers, while administrative writing is most likely produced in a less formalised publishing environment. With the exception of perhaps governmental texts, most of the texts in this register are probably edited by freelance editors, who are likely to work on multiple texts from different clients at the same time.

Editors of instructional texts engage with the language, style, content and structure of the text, ensuring that the text adheres to the preferred spelling, punctuation and usage conventions of the publisher or publishing institution, but also that the distinctive and conventionalised discourse style of the particular text type is met. Furthermore, given the importance of accuracy of information in instructive texts, editors are usually required to engage in substantive editing practices, such as removing, rewriting, adding and rearranging sections of texts, and consequently might need to assume a writing function. The implication of this is that editors' own usage practices and preferences, which might differ from those of the authors, might find their way into these texts. Furthermore, given the fact that most editors working in this register are employed on a freelance basis, it is likely that these editors work with multiple sets of normative guidelines from their different clients, creating a much more diffuse normative representation and opening up the opportunity for a kind of cross-normative transfer in these texts.

The function of creative texts, which includes works of fiction such as narratives, novellas and short stories, is to entertain by way of reading pleasure. Creative texts are written for a broad general audience and are distinguished from other registers in that the language conventionally used in creative prose draws on narration and the representation of speech in the form of written dialogue (Biber & Conrad, 2009; Biber et al., 1999). Typical linguistic features found in creative prose are past tense verbs and third person pronouns (as linguistic features typical of narration) and the extensive use of features associated with spoken language (to reflect written dialogue). Since most creative texts are published through book publishers (although the advent of self-publishing has changed this somewhat), creative texts are edited in traditional publishing contexts, but usually by freelance editors (often specialising in the editing of creative texts).

The editing of creative texts is usually carried out over an extensive two-stage process: first, the texts go through a developmental editing stage in which a developmental editor assesses the text's content,

interrogating aspects such as the development of the story or narrative as well as the characters therein. During this stage, editors will also engage with the accuracy and logic of the content (such as ensuring the accuracy of historical references in historical fiction, or assessing the (a)chronological progression of the text's story line). Once developmental editing is complete, the text will be copyedited by a copyeditor during which matters such as the correctness and consistency of spelling, grammar and punctuation; consistency of tone; the representation of dialogue; and style receive attention. Creative texts therefore undergo extensive editing and it is not unusual for the final edited version of a creative text to differ substantially from the unedited manuscript. Because different kinds of editors work on creative texts during different stages of the production process, the text is subjected to the acceptability judgements of (at least) two people. This means that for creative texts, two different editors' linguistic competence (in the sense discussed in Section 2.3) are brought to bear on the text, which further adds to the (more) diffuse normative conditions under which these texts are edited.

The popular register represents non-fiction texts, such as magazine articles and books, that are intended for use by the general public. Although popular non-fiction texts are informational, the focus of these texts is less on the comprehensive reporting of this information and more on the conveying of different kinds of information for the interest of the general public. Furthermore, some popular texts may include extensive sections of narration, while others, such as biographies, may include the representation of dialogue.

The production of popular texts, such as magazines and books, is typically carried out via traditional publishing channels at book-publishing or media houses, which have very specific expectations of standard usage and conventions. Magazine texts are published on a regular basis, and this usually means that the text is edited by an in-house copyeditor who edits all texts that appear in the publication – this editor therefore works in a focused normative environment, drawing on the preferred usages and conventions of the publication for which she works. Modern book publishers, on the other hand, mostly outsource editing work to freelance editors, who might work on several kinds of publications all of which might require different norm-providing sources, creating a more diffuse normative environment. Furthermore, the scope of editing work done on popular magazine and book texts is usually substantive, requiring the editor to assess aspects of language (such as spelling, grammar and punctuation), as well as aspects of style, presentation, structure and content. In this process, the editor may need to (re)write sections of text and in so doing may impose her own writing conventions.

Reportage includes news reports and newsletters that function to convey general, though sometimes specialised, information of current interest. Both kinds of texts serve a reporting function, in which the focus is on the comprehensive reporting of information related to current events. Texts in the reportage register are known to be particularly receptive to the features associated with ongoing change and ongoing stylistic shifts in language, such as colloquialisation and densification (Biber & Gray, 2012) (see Section 2.5). News reports are typically published in newspapers on a daily or weekly basis. However, recent developments in technology mean that the publication of these types of texts has been shifting to online

news platforms, with the publication of texts online as soon as they have been edited. Newsletters published by various organisations are published less often (such as monthly, or quarterly).

Given the frequency with which newspapers are published, subeditors working on news reports are employed on an in-house basis and work under an enormous amount of pressure to complete their work before the publication goes to print. The editors of news reports do a substantive amount of editing in a very short space of time, and the work usually includes editing the language, style, structure and content of the text. Often this work also includes a substantive writing component, such as the writing of headlines and captions, and the rewriting of the news report itself. Because the editors of these texts are employed in-house, their work is guided by the preferred usages of one publication, and they therefore work in a focused normative environment. However, the fact that these editors work under an enormous amount of time pressure most likely means that these editors rely extensively on their knowledge of language which opens up the possibility that their covert norms might influence their editorial choices. Newsletters are published less often, and so editors working on these types of texts have more time to edit, compared to their counterparts working at newspapers. Because newsletters are published less frequently and in non-publishing institutions, they are usually edited by freelance editors. Such editing may also be substantive, but normally does not require as much rewriting as is required in news reports. Given that these editors are typically employed on a freelance basis, they probably work in a more diffuse normative environment, which creates the opportunity for innovations to creep in and for cross-normative effects to influence their editorial choices.

The registers represented in this study therefore include a diverse range of registers that are distinguished from each of other in several respects, such as their functions and purposes as well as their intended readers. These registers are also known to be variously receptive or resistant to the features and trends associated with ongoing language change. In addition to all being edited, the registers reflect a diversity in the nature and scope of editorial work as well as production circumstances, providing a good opportunity to investigate the influence of editorial intervention as it applies to different registers and within different contexts. Having discussed the varieties and registers selected for investigation in this study, the next section shifts the focus to the grammatical feature to be investigated.

3.4 Feature under investigation: The genitive alternation

The choice of the grammatical feature to investigate in this study was guided by several considerations, the most important of which included the following: it is a feature that is known to be undergoing change that is differentiated across varieties of English; it is a feature that shows variation across registers; it is a feature that can be linked to broader patterns of language change, such a colloquialisation and densification; and lastly it is a feature that has some normative salience, but which does not have such strong proscriptions that there is little variability. In other words, the feature selected for investigation in this study needed to be

one where the interplay between social norms and the covert norms of authors and editors, as proposed in Chapter 2, is productive and varied.

Genitive alternation, that is the choice between different constructions to express similar genitive relations (as in *the Missouri river's bank* versus *the bank of the Missouri river*), is probably the best researched of all syntactic alternations in the English language (Rosenbach, 2014). The significant body of research on this alternation in English has investigated historical changes in the distribution of the constructions from Old English to Late Modern English as well as the various semantic–pragmatic, phonological, syntactic and stylistic factors that influence the choice between the constructions, with a fair amount of work focusing on the influence (and interaction) of especially linguistic factors in the competition between the variant constructions (Rosenbach, 2002, 2014). Nearly all of this work has proceeded from the view that this alternation is made possible because two major nominal devices, the *s*-genitive and the *of*-genitive constructions, offer different ways of saying the same thing. Thus, genitive alternation is a linguistic variable, in the variationist sense (see Section 2.2.2) and the ongoing change of this feature in contemporary English is rooted in the increasing use of the *s*-genitive at the expense of the *of*-genitive (Labov, 1972; Rosenbach, 2002, 2014; Rosenbach & Vezzosi, 2000).⁷

Recently researchers have shifted their attention away from the development of this alternation in the older forms of English and have instead started to investigate this alternation in Present-day English. While most of the recent work has focused on AmE and BrE, some work also been done on cross-varietal differences in varieties of English other than BrE or AmE, such as HkE, Canadian, Indian, Jamaican, New Zealand, Nigerian, Singapore and Philippine English⁸ (Akinlotan, 2016; Heller, Bernaisch, & Gries, 2017a; Heller et al., 2017b; Szmrecsanyi, Grafmiller, Heller, & Röthlisberger, 2016b). Despite the increase in cross-varietal studies of genitive alternation in World Englishes, AusE and SAfE have largely been excluded from such studies and have each featured in only one study of the alternation between the two constructions (Ford & Bresnan, 2015; Rosenbach, 2017). Consequently, very little is known about genitive alternation in these two varieties.

In addition to this shift to a broader cross-varietal focus, analyses of the alternation between the two constructions have also been extended to other language-external factors, such as modality and register (Grafmiller, 2014; Heller, 2018a; Heller et al., 2017b; Jankowski, 2013; Jankowski & Tagliamonte, 2014);

⁷ Recently, some researchers have suggested that there is a third major construction, variously referred to as the NN-genitive (Szmrecsanyi, Biber, Egbert, & Franco, 2016a), nominal premodifier (Feist, 2012), noun modifier (Rosenbach, 2007) or noun+noun sequence (Rosenbach, 2006). Feist (2012, p. 265) argues that these NN-genitives are true genitives comprising an uninflected nominal (possessor) that occurs before the head noun (possessum) and which may be represented as 'N₁ N₂'. These NN-genitives are frequent in Present-day English (Feist, 2012; Rosenbach, 2007, 2014) and can (and do) alternate with the *s*-genitive and *of*-genitive constructions, as demonstrated in the alternation between the *s*-genitive *the study's findings*, the *of*-genitive *the findings of the study* and the NN-genitive *the study findings*.

⁸ These varieties will henceforth be referred to as CanE, InE, JamE, NZE, NigE, SngE, and PhiE, respectively.

however, the majority of this work has focused on the distribution of the constructions in press writing as compared to speech, and there is very little work on genitive alternation across different written published registers.

The work by Jankowski (2013) and Jankowski and Tagliamonte (2014, p. 324) is a particularly important motivating factor for the selection of the genitive alternation as the feature to investigate in this study. Scholars disagree on the causes of the ongoing shift to the *s*-genitive.⁹ Jankowski (2013) and Jankowski and Tagliamonte (2014) attribute this shift to an ongoing loosening of the animacy constraint with non-human possessors, which they argue is most advanced in written registers and thus suggest that the changing distributional patterns of the major adnominal constructions might be an example of change in which written language is leading the way. In light of this, it may be argued that written language, and especially written published language plays an important role in the ongoing change in genitive alternation and therefore the different types of written published registers are worthy of further investigation. Situating this within the context of the current study, to the best of my knowledge, no study has investigated the role of editorial intervention as a possible (language-external) factor that might account for the occurrence of the variant constructions in different types of written published registers.

The selection of the genitive alternation as a feature in this study therefore seeks to contribute to the existing body of knowledge by describing the distribution of and variation between the two major genitive constructions (the *s*- and *of*-genitives) in five different written published registers (academic, creative, instructional, popular and reportage) in two varieties of English for which little is known about genitive alternation (AusE and SAfE). The primary contribution is in the investigation of editorial intervention as a potential language-external factor that influences the changing distributional patterns and use of genitive constructions in different contexts.

A further motivating factor for the selection of this grammatical feature as a case study in this thesis relates to the fact that while there is advice in various normative sources on the use of the two genitive constructions (and more specifically, the use of *s*-genitives with animate possessors), the feature is not so saliently normatively marked that it crops up in complaints about language in English (compared to, for example, the split infinitive or subject-verbal agreement), and is possibly less strongly represented in the minds of editors as something to look out for. This opens up the opportunity to explore the interaction between overt norms (as sanctioned in norm-providing sources) and covert norms in editors' choices with regard to this feature.

Before proceeding, it is necessary to address the issue of grammatical variation and specifically how syntactically different genitive constructions might offer different ways of saying the same thing. This question is a fraught one, with much debate on whether the different genitive constructions really do

⁹ These disagreements will be discussed in more detail in Section 3.4.3.1.

constitute a true linguistic variable (as set out in Section 2.2.2), for two reasons. First, it is difficult to determine what is meant by the ‘sameness’ of the genitive constructions and second, it is difficult to delimit the range of contexts in which the constructions can alternate (Rosenbach, 2002, p. 25). Gries and Stefanowitsch (2004) and Stefanowitsch (2003), for example, argue that *s*-genitive and *of*-genitive constructions essentially encode different semantic relations, are therefore semantically distinct and do not constitute variants. Rosenbach (2002, 2007) argues for a less rigid approach to meaning equivalence by stating that under such a strict view there would be no such thing as grammatical variation because there will always be one or some factor(s) that, in specific contexts, will favour one or the other construction. Instead, she argues for an approach in which the constructions do not have to be equivalent in every semantic–pragmatic and/or functional aspect so long as there is enough overlap to allow for the alternation between the variants in different contexts (Rosenbach, 2007, p. 152). To this end, Rosenbach (2002) distinguishes between categorical contexts in which only one construction can occur, and choice contexts in which either could be used to express a (near) similar meaning. It is these choice contexts that allow for alternation between the variants and it is the alternation of the constructions in these contexts that forms the basis for investigations of genitive alternation.

This thesis adopts the latter approach and proceeds from the view that genitive variation can occur in certain contexts and that it is in these choice contexts that the variants under investigation express approximately the same set of relations among their constituent parts. It further assumes that the alternation between the variants is conditioned by (the interaction of) certain language-internal and language-external factors.

3.4.1 The constructions

The consensus in most research into genitive alternation in English, as pointed out in Section 3.4, is that two constructions, the *s*-genitive and the *of*-genitive, are the major adnominal devices available to express possessive relations (Rosenbach, 2007). The two constructions each comprise two noun phrases (called the ‘possessor’ and ‘possessum’) that are linked via a relational marker (Rosenbach, 2002, 2014). The *s*-genitive construction links the possessor and possessum via the clitic ‘s relational marker, and may be represented as [NP₁ ‘s NP₂], where the possessor occurs in the prenominal position and is followed by the possessum (as in (3.1a)). The *of*-genitive construction, on the other hand, uses the preposition *of* as relational marker, and may be represented as [*the* NP₂ *of* NP₁], where the possessor occurs in postnominal position (as in (3.2a)). The alternation between the two arises in the potential for either construction to be used in the same context to express approximately the same meaning, as demonstrated in (3.1b) and (3.2b).

(3.1a) She programmed **the clinic**_{possessor} ‘s **address**_{possessum} into her GPS. [SA_C-006-E]¹⁰

¹⁰ In this thesis, exemplifications of the genitive constructions are referenced to their identifying label in the corpus. This abbreviated label contains four sets of information: the variety from which the example is taken (*Aus* for examples from the Australian sub-corpus and *SA* for those from the South African sub-corpus); the register in which the genitive occurred (*A* = academic, *C* = creative, *I* = instructional, *P* = popular and *R* = reportage); the number of the unedited and edited text pair in the sub-corpus; and lastly whether the example occurred in the unedited text (identified with *O*

- (3.1b) She programmed **the address**_{possessum} *of* **the clinic**_{possessor} into her GPS.
- (3.2a) ‘I’ll be here,’ <character name deleted> said as <name deleted> climbed into **the back seat**_{possessum} *of* **the station wagon**_{possessor}. [Aus_C-005-E]
- (3.2b) ‘I’ll be here,’ <character name deleted> said as <name deleted> climbed into **the station wagon**_{possessor}’s **back seat**_{possessum}.

Not all instances of *s*-genitives and *of*-constructions are interchangeable. Because each construction forms part of complex families of construction types that differ in terms of their semantics and syntax, they do not all enter into proper genitive alternation and therefore some are restricted to categorical contexts (Rosenbach, 2006, 2007, 2014).

Grammars of the English language broadly identify two types of *s*-genitives: specifying or determiner genitives (as in (3.3)), and classifying or attributive genitives (as in (3.4)) (Biber et al., 1999; Huddleston & Pullum, 2002).

- (3.3) ... assist with general healthcare ensuring that [[**the children**]_{NP}’s **health needs**]_{NP} are taken care of ... [SA_I-084-E]
- (3.4) ...a farm stall near the bridge, is another good place for provisions; it has pancakes on the menu and [**a children’s**]_N **playground**]_{NP}. [SA_P-018-E]

Although determiner and attributive genitives appear to be very similar, the two differ in a number of ways, especially in how the possessor relates to the possessum. In determiner genitives, the possessor has a determiner function and expands the possessor into a noun phrase (Huddleston & Pullum, 2002). It occupies the determiner slot of the whole noun phrase to which it belongs and specifies the definiteness and establishes the reference of the noun phrase (Biber et al., 1999, p. 294). The possessor of a determiner genitive therefore renders the whole noun phrase definite (even if the possessor is indefinite) (Rosenbach, 2006, p. 80). For example, *a country’s border* is a determiner *s*-genitive with an indefinite possessor. When translated into the *of*-genitive, the construction would appear thus *the border of a country*.¹¹ Furthermore, since definiteness can only be marked once within an *s*-genitive, the possessor cannot be preceded by a separate determiner of the head noun. As such, the possessor “anchors” the whole noun phrase and narrows down its referent (Rosenbach, 2006, p. 80).

The possessor of attributive genitives, on the other hand, is not a noun phrase but a noun or nominal. In the case of attributive genitives, the possessor functions to modify the possessum by classifying what type it is (Biber et al., 1999, p. 294). Attributive genitives function in ways similar to adjectives and (most) noun

for original) or in the edited text (identified with *E* for edited). Thus, example (3.1a) is taken from the South African creative register and is from the edited version of text pair number 006.

¹¹ Indefinite possessors can be difficult to classify because they are often ambiguous in terms of their referential and non-referential meaning (Rosenbach, 2005, p. 22).

premodifiers because they denote properties of the head noun and are therefore not independently referential. Unlike determiner genitives, attributive genitives are neutral as to the definiteness of the noun phrase: any determiner in the noun phrase belongs to the head noun and not to the genitive (the determiner *a* in (3.4) belongs to the head noun *playground* and not to the nominal *children's*). They do not confer definiteness on the whole noun phrase and can be preceded by determiners as well as modifiers of the noun phrase to which they belong (Biber et al., 1999, p. 295).

Studies of genitive variation generally limit their analyses to determiner *s*-genitives because it is only the determiner *s*-genitive that can enter into proper alternation with *of*-constructions (Rosenbach, 2002, 2014). For example, while the determiner genitive in (3.3) may be paraphrased as *the health needs of the children*, the attributive genitive in (3.4) can only really be paraphrased using another periphrastic construction as in *a playground for children* but not as *a playground of children* (which would carry an entirely different meaning).

As with *s*-genitives, *of*-genitive constructions can relate the possessor to the possessum in two ways: either as complement to the possessum (as in (3.5)) or as modifier of the possessum (as in (3.6)) (Rosenbach, 2002, p. 18; 2014, p. 223).

(3.5) **The roof of the silo...** [Aus_C-002-E]

(3.6) He was **a man of great honour...** [SA_P-007-E]

In (3.5), the *of*-phrase functions as complement to the possessum by specifying whose roof it is. In contrast, the *of*-phrase in (3.6) acts as modifier of the head noun *man* by ascribing an attribute to it. In the case of (3.5), the *of*-construction is interchangeable with the *s*-genitive construction and may be rephrased as *The silo's roof* while retaining the same meaning. On the other hand, the modification function fulfilled by the *of*-construction in (3.6) means that it is not interchangeable with the *s*-genitive construction (*a great honour's man*). Rosenbach (2002, p. 18) notes that in both of these functions the whole *of*-construction, unlike the determiner *s*-genitive, could be either definite or indefinite (*A roof of the silo* and *the man of great honour*), and that it is only a definite *of*-genitive construction in which the possessor functions as a complement to a definite possessum that can alternate with the *s*-genitive, since the *s*-genitive always expresses the possessum as definite. In other words, *of*-phrases such as *A roof of the silo* cannot be rephrased as *A silo's roof* since the *s*-genitive would render the possessum *roof* definite, and would therefore translate to *The roof of a silo*. Therefore, *of*-genitives with a possessum that cannot be interpreted as definite cannot be expressed by a corresponding *s*-genitive and constitute categorical contexts of the *of*-genitive. In a similar vein, *of*-genitives that begin with a demonstrative (e.g. *this, that, these, those*) as well as those that begin with a possessive pronoun (e.g. *mine, yours, his, hers*) are not interchangeable with the *s*-genitive since possessors and central determiners cannot co-occur in English (Rosenbach, 2005, p. 622). For example, the phrase *this section of the chapter* cannot be rephrased as *this the chapter's section*. However,

definite *of*-genitives in which the possessor is preceded by a demonstrative can alternate with *s*-genitives. The example in (3.7) demonstrates this interchangeability:

(3.7) ...**the villages of this region** [this region's villages] and, of course in the north.... [Aus_C-005-E]

3.4.2 Historical changes in the genitive alternation

There is a substantive body of research on the historical development of the genitive alternation from Old English to Late Modern English. Although the current study draws on synchronic data, it is necessary to (briefly) sketch the diachronic development of genitive alternation to provide some contextualisation for the investigation of this alternation in Present-day English in this thesis, particularly in respect of patterns of ongoing change in genitive usage.

One of the most significant changes in the distribution between the two constructions occurred firstly in the transition from Old English to Middle English when the periphrastic *of*-genitive drastically replaced the old inflectional *s*-genitive. In Old English, the *of*-genitive started out as a marginal construction, accounting for only 1% of all adnominal genitive constructions, and was therefore not yet established as a major grammatical variant of the inflected genitive (Rosenbach, 2002, p. 177). According to Rosenbach (2002, p. 177), the *of*-genitive was predominantly used to indicate source (see (3.8)) or partitive relations (see (3.9)) between the possessor and possessum and developed from the original meaning of the preposition *of*, 'out of'. The examples in (3.8) and (3.9) reflect these two relations and are taken from Rosenbach (2002, p. 177):

(3.8) hē...hēt ȝetrimbian cyrican of treowe
'he had a church of wood built'

(3.9) sume of pām sundorhāl ȝan
'some of the saints'

Within the much more prevalent inflected genitive in Old English, two word-order options for the realisation of the possessor were available: the prenominal inflected genitive, as shown in (3.10) and the postnominal inflected genitive, as shown in (3.11). (Both examples are taken from Rosenbach (2002, p. 178)).

(3.10) pæs cyningespegnas
the-GEN king-GEN thanes
'the king's thanes'

(3.11) heafod ealra haligra manna
head all-GEN holy-GEN men-GEN
'the leader of all saints'

Beginning in Old English, the prenominal inflected genitive began to rise in frequency at the cost of the postnominal inflected genitive, which ultimately disappeared from use at the beginning of the Middle English period (Rosenbach, 2002, p. 179). The disappearance of the postnominal inflected genitive meant that the prenominal genitive was the only option for the inflected genitive in Middle English, and word order options such as (3.11) were no longer available. However, from the twelfth century onwards the inflected genitive started to be replaced by the periphrastic *of*-genitive, and by the fourteenth century the inflected genitive was largely restricted to use with animate possessors, topical possessors and in expressions of possessive/subjective genitive relations (Rosenbach, 2002; Wolk, Bresnan, Rosenbach, & Szmrecsanyi, 2013).

In Early Modern English, the *s*-genitive underwent a revival and began to increase in relative frequency vis-à-vis the *of*-genitive, so that it became the major competing nominal genitive construction with the *of*-genitive (Rosenbach, 2002; Rosenbach & Vezzosi, 2000). The distribution of these two genitive constructions fluctuated in the Late Modern English period: from 1650–1849 the *s*-genitive started to decline vis-à-vis the *of*-genitive, so that by 1800–1894 it accounted for 11% of interchangeable instances with the *of*-genitive (Wolk et al., 2013, p. 392). From 1850, the *s*-genitive gradually started to recover. Wolk et al. (2013, p. 392), who draw on the BrE news and letters sections of the ARCHER corpus, find that in the 1950–1999 time period, the *s*-genitive increased in frequency and accounts for 38% of interchangeable instances.

3.4.3 Factors influencing genitive alternation

Apart from studies on the historical development of the genitive alternation, a substantial amount of research has been done on the factors that condition the alternation between the major constructions. This research has shown that the choice between the two is not random but is determined by (the interaction of) a number of factors (Hinrichs & Szmrecsanyi, 2007; Rosenbach, 2002, 2008, 2014). The factors typically surveyed are linguistic (that is, those internal to language) and include a range of factors that are, on the one hand, regarded as general determinants of word order variation, such as syntactic weight, and on the other, seem to be idiosyncratic to genitive alternation, such as semantic relation and final sibilancy of the possessor (Rosenbach, 2014, p. 226).

Nearly all studies of genitive alternation include one or some combination of these linguistic factors (typically studied within the context of the two major standard varieties, AmE and BrE); however, recently some studies have extended the scope of factors to include those that are external to language, such as variety, modality and register (although research into these external factors is by no means as well established as that of language-internal factors). The investigations into the different factors that condition genitive alternation have led to the conclusion that none is deterministic (even if some have been argued to carry more weight than others). Rather, genitive variation is conditioned by multiple factors that interact to

influence the selection of one construction over the other (Hinrichs & Szmrecsanyi, 2007; Rosenbach, 2002, 2008).

The discussion of the different factors that follows is split along the lines of whether the factors are considered to be language internal or language external. Section 3.4.3.1 provides a brief exposition of the linguistic factors known to condition genitive alternation, while Section 3.4.3.2 discusses language-external factors and how such external factors interact with linguistic factors to condition the choice between the two constructions across varieties, modalities and registers.

3.4.3.1 Linguistic factors conditioning the genitive alternation

The linguistic factors known to play a role in genitive alternation can broadly be grouped into two types: those that operate on local, inherent properties of the possessor and those that are dependent on the context in which the possessor is used (Rosenbach, 2014, p. 231). Examples of the latter include animacy of the possessor, definiteness of the possessor and whether or not a final sibilant is present in the possessor. Examples of factors that are dependent on the context include semantic relation, syntactic weight, givenness, thematicity and persistence.

According to Szmrecsanyi (2010), language-internal factors that condition genitive alternation can be clustered into four major groups: semantic and pragmatic factors, phonological factors, economy-related factors, and processing and parsing-related factors. The first group include the semantic and pragmatic properties of the genitive constructions, such as animacy, definiteness of the possessor and semantic relation. Phonological factors are those factors in which phonological constraints condition the choice of construction, such as the final sibilancy of the possessor. While Szmrecsanyi (2010) notes that economy-related factors are language internal, he states that such factors pertain to preferences associated with contexts and registers where brevity is crucial or preferred (and so might actually be more accurately classified as language-external factors). This links up with the notion of densification as discussed in Section 2.5. The last group, processing and parsing-related factors, reflect cognitive and processing constraints (which are assumed to be the same for all language users). An important principle related to cognitive and processing constraints is the ‘Easy first principle’ which states that language users tend to place those constituents that are easy to retrieve first (Heller et al., 2017b). The ease with which such constituents can be retrieved is influenced by a number of factors, such as for example the principle of end weight or the frequency of a construction or constituent part. Linguistic factors that are processing driven include syntactic weight, givenness, thematicity and persistence. The discussion of linguistic factors in this sub-section will be arranged according to these four groups of factors (Szmrecsanyi, 2010).

Animacy is arguably the most important factor in genitive choice, and while it is not considered the most decisive factor, it has been shown to be extremely influential and will consequently receive more attention here than the other factors (Rosenbach, 2002, 2007, 2014, 2017; Szmrecsanyi, 2010; Szmrecsanyi et al., 2016a; Wolk et al., 2013). Most research into the effect of animacy has analysed the animacy of the

possessor and has shown that possessors high in animacy are more likely to occur in the prenominal position and therefore take the *s*-genitive, while possessors low in animacy are more likely to be realised postnominally in the *of*-genitive. Animacy is difficult to define and different studies, depending on whether or not animacy is the primary factor of the analysis, have taken different approaches to how it is conceptualised.

On the one hand, some studies draw on the binary classification of the possessor as either human, a category also taken to include higher animals (therefore animate) or nonhuman (therefore less/not animate) – a view that is also taken in (prescriptive) English grammars (see more detail in Section 3.5) (Heller et al., 2017b; Rosenbach, 2002). Others take a more nuanced approach, classifying the noun type of the possessor by plotting it along a gradient animacy scale in which nouns are classified in terms of an animacy hierarchy comprising two extreme poles (animate and inanimate). Table 3.1, which is taken from Rosenbach (2007, p. 154), reflects this scale.

animate		inanimate			
human N	animal N	collective N	temporal N	locative N	common N
>	>	>	>	>	>
<i>The boy's bike</i>	<i>The dog's collar</i>	<i>The company's director</i>	<i>Monday's mail</i>	<i>London's suburbs</i>	<i>The building's door</i>

Table 3.1: Animacy scale for plotting possessor noun type

This more nuanced approach to the classification of the noun types allows for a gradient categorisation of the animacy of possessors, where the two extremes of the scale are linked via intermediate categories: collective nouns, temporal nouns and locative nouns. The use of this scale to categorise possessors has demonstrated two important trends related to the animacy of the possessor in conditioning the alternative realisation of the *s*-genitive and the *of*-genitive. First, *s*-genitives do not occur exclusively with human possessors and second, *s*-genitives have in the course of Late Modern English and in Present-day English come to be used with some subsets of inanimate nouns, such as collective, temporal and locative nouns (Heller et al., 2017a; Rosenbach & Vezzosi, 2000; Wolk et al., 2013). The increasing use of the *s*-genitive with non-human possessors has formed part of the explanation for the construction's increase in frequency in Present-day English, but there is less agreement on whether this extension is motivated by a loosening of the animacy constraint for particular subsets of nouns lower down on the animacy scale (which would be evidenced by an increase in the use of the *s*-genitive with, for example, only collective, temporal and, to a lesser extent, locative nouns), if this is due to a general spread of the form to inanimate noun possessors (which would be evidenced by an increase in the use of *s*-genitives with all types of nouns lower on the animacy scale) or if this is due to changing discourse practices (such as the trend towards colloquialisation or economisation which favour the less formal and more compact construction respectively) (Hinrichs & Szmrecsanyi, 2007).

The first inanimate noun category on the scale, collectives, is a particularly interesting case. Rosenbach (2008, p. 153) notes that there is variation in how collectives can be conceived: either as animate or inanimate. Using *company* to demonstrate this, she argues that *company* may be interpreted to mean

institution, in which case it is an inanimate noun and would most likely take an *of*-genitive, or it may be interpreted as the group of people who are employed by the company, in which case it would have an animate interpretation and therefore take the *s*-genitive.

There is also evidence that temporal and locative possessors, which are ranked as inanimate on the animacy scale, have come to be used with *s*-genitives rather than with *of*-genitives, as would be expected with inanimate possessors. Rosenbach (2008) argues that the realisation of temporal and locative possessors in prenominal position might be due to two reasons unrelated to their animacy status (and therefore their increasing realisation with *s*-genitives is not due to a loosening of the animacy constraint). First, temporal and locative nouns bind head nouns with respect to time or space, making them good referential anchors (a core function of *s*-genitives), and second the increasing use of locative and temporal possessors with the *s*-genitive might also be due to the fact that temporal and locative nouns can also be proper nouns (and therefore rank highly in terms of definiteness).

The definiteness of the possessor, which like animacy can be plotted on a scale, also influences which genitive is selected. Referents high in definiteness favour the *s*-genitive, while referents lower in definiteness tend to favour the *of*-genitive. Rosenbach (2014) points out that while definiteness correlates highly with animacy, to the extent that some researchers coalesce the scales, definiteness and animacy are not reducible to each other and should therefore not be considered as the same factor. For example, most proper nouns are human names and are therefore high in animacy and definiteness. Other types of nouns, such as temporal or locative nouns, which are usually also proper nouns, are not high in animacy but are high in definiteness, and therefore when temporal or locative nouns are used in prenominal position in the *s*-genitive, this might be due to their definiteness, rather than a relaxation of the animacy constraint.

The semantic relation that holds between the possessor and possessum is, alongside animacy, one of the most important factors to condition the alternation between the genitive constructions. The identification of the range of semantic relations encoded by the genitive is extremely difficult and different approaches may be taken to this classification (Akinlotan, 2016; Feist, 2012; Hinrichs & Szmrecsanyi, 2007; Jankowski, 2013; Rosenbach, 2002, 2014; Szmrecsanyi, 2010; Szmrecsanyi & Hinrichs, 2008). Wolk et al. (2013) state that because there is no generally accepted approach to the categorisation of the semantic relations that can be expressed by the possessor and possessum, many studies dismiss this as a factor to be investigated. Some recent work to include semantic relation as a factor seems to accept the prototype approach adopted by Rosenbach (2002). It is beyond the scope of this study to give a detailed account of the different approaches to the classification of the semantic relations between possessors and possessums, and therefore in the interests of comparability with recent approaches, the approach outlined by Rosenbach (2002) is adopted here.

Rosenbach (2002) distinguishes between prototypical and non-prototypical semantic relations. The prototype category collapses several classes of relations that express legal ownership (3.12), body part

relations (3.13), kinship relations (3.14) and part–whole relations (3.15). Typically, semantic relation is coded based on the head noun of the possessum (Bresnan, Rosenbach, Szmrecsanyi, Tagliamonte, & Simon, 2017).

- (3.12) ...black children would go to **Erna's flat**, where she would teach them. [SA_A-013-E]
- (3.13) The morning sun shone on **<character name deleted>'s back** as he sat on his back porch, eating breakfast. [Aus_C-002-E]
- (3.14) **Malusi's uncle** wrote a list of the businesses in the Wynberg Industrial Park. [SA_I-018-E]
- (3.15) **The engine of the Volvo** rips into the silence, startling me. [SA_C-002-E]

Any semantic relations that do not fall into one of these four subcategories are classified as non-prototypical (demonstrated in (3.16) and (3.17):

- (3.16) In conclusion, public policy thus requires that both **the employee's right** to freedom of trade and **the right of the employer** to protect his interests, be protected in our society. [SA_A-019-E]
- (3.17) **<character name deleted>** studied the cakes and **<character name deleted>** looked at **the contents of the hamper** ... [Aus_C-002-E]

Rosenbach (2002, p. 64) notes that the prototype account of semantic relations among English genitives is the only taxonomy to provide predictions regarding the productivity of the different genitive constructions. In this regard, it has been argued that *s*-genitives are favoured by those constructions that express prototypical relations, while the *of*-construction is favoured by those that express non-prototypical relations (Grafmiller, 2014; Jankowski, 2013; Rosenbach, 2002; Wolk et al., 2013).

Final sibilancy of possessors is a phonological variable that also seems to influence which genitive is selected. Previous scholarship indicates that possessors ending in a sibilant, such as [s] as in *house*, [z] as in *news*, [ʃ] as in *Bush*, [tʃ] as in *avalanche*, [dʒ] as in *bridge* and [ʒ] as in *garage*, are less likely to attract the *s*-genitive because the sound sequence of the final sibilant followed by 's is considered harder to pronounce (Grafmiller, 2014; Heller et al., 2017a). Final sibilancy is therefore a production-driven factor. Being a phonological variable, it is expected that final sibilancy of the possessor would exert a stronger effect in spoken language than in written language, but the effect of final sibilancy across different modalities has been shown to vary in different varieties of English (Heller et al., 2017b).

Lexical density is an economy-related factor associated with the informational density of the context in which a construction occurs. This information density is usually measured in terms of the type-token ratio of a text. Because the *s*-genitive is the more compact and economical of the two genitive constructions (Szmrecsanyi, 2010), in contexts where there is a need to code information economically, the *s*-genitive is usually the preferred construction.

The syntactic weight of the possessor and possessum has also been shown to be important in conditioning which genitive variant is selected in a choice context. Syntactic weight is a processing-driven factor influenced by the principle of end weight. According to the principle, language users tend to arrange lexical phrases in order of increasing weight, so that heavier constituents tend to follow shorter ones. This ordering facilitates planning, production and processing (Heller et al., 2017b). Applying this to genitive alternation, heavier possessors are more often realised in postnominal position (in the *of*-genitive), while shorter possessors are more often realised in prenominal position (in the *s*-genitive) (Hinrichs & Szmrecsanyi, 2007). One of the challenges with syntactic weight relates to how it is operationalised, which may be either in terms of the number or length of words in the possessor and possessum or the syntactic complexity of the possessor and possessum. Most studies seem to take the former approach, which is generally accepted as the best (and least complex) way of operationalising syntactic weight (Rosenbach, 2014). A further problem arises in how syntactic weight is represented. On the one hand, it may be represented as a ratio of the number of words or tokens in the possessor and possessum, as in Bresnan and Ford (2010) and Grafmiller (2014), or the weight of each constituent may simply be coded as a separate measure and analysed separately, as in Heller et al. (2017b). Like animacy, the effect of syntactic weight is not deterministic, but it has been shown to have a strong influence in increasing the likelihood of one genitive construction being selected over the other.

Givenness and thematicity are two language-internal factors related to the information status of the possessor (Hinrichs & Szmrecsanyi, 2007), but they are not a property inherent to the possessor (Rosenbach, 2014, p. 228). They are factors related to the discourse context, and while they might seem to be similar (and would in certain respects overlap) they essentially refer to two different conditions. A possessor is considered to be ‘given’ if it has been mentioned in the preceding context of the text, otherwise it is considered ‘new’. Given possessors are more easily retrievable from memory and, it has been argued, favour the *s*-genitive. Conversely, new possessors are more likely to be realised via the *of*-genitive. The preceding context in which the givenness of a possessor can be established is usually limited to a certain number of words or lines prior to the occurrence of the possessor. Researchers may either conduct an automatic search for string identity of the possessor within this context or they may do so manually (which has the disadvantage of being extremely time-consuming). Rosenbach (2014) points out that the use of an automatic search tool is problematic because some possessors might be expressed by synonyms in preceding text, which would not be picked up through the use of automatic searches (Rosenbach, 2017, p. 229). Grafmiller (2014) notes that the influence of the possessor’s givenness is a controversial point in the literature, with some researchers finding that there is no significant effect of possessor givenness on which genitive construction is selected, while others find that givenness does influence genitive alternation.

Like givenness, thematicity is related to the discourse context of the text and refers to whether or not the possessor is a central topic or theme of the text. Thematicity is a processing-driven factor because highly thematic possessors are considered to be highly accessible in language users’ minds, and are therefore more

likely to occur early on in constructions (in the *s*-genitive). While the givenness of the possessor is established by determining whether it has been mentioned in the previous context of the genitive construction, the thematicity of a possessor is established by determining the frequency with which it occurs in the whole text (Grafmiller, 2014; Heller et al., 2017a; Heller et al., 2017b). Thematic possessors are therefore also given, with the exception of the first occurrence of the possessor in the text. The literature shows that highly thematic possessors favour the *s*-genitive, even when they would otherwise be disfavoured (in terms of other internal constraints). For example, possessors low on the animacy scale and therefore less likely to be realised by the *s*-genitive, such as common nouns like *bicycle*, are more likely to be used in *s*-genitive constructions if cycling is a central topic of the text in which the possessor occurs (Grafmiller, 2014; Heller et al., 2017a; Heller et al., 2017b; Hinrichs & Szmrecsanyi, 2007).

A further language-internal factor related to the discourse context and one that is in certain respects very closely related to givenness and thematicity, is persistence. Persistence refers to the influence that the choice of genitive variant will have on the likelihood that the same form will be used in subsequent contexts (Grafmiller, 2014). Persistence therefore relates to the genitive construction as a whole, rather than to one of its constituent parts, and is considered to be a processing-related factor. Because language users tend to recycle material that they have used or heard before, the *s*-genitive is more likely to occur if an *s*-genitive has been used in the preceding context, while the *of*-genitive is more likely to occur if it has been used in the previous context (Rosenbach, 2014, p. 232; Szmrecsanyi, 2010). The occurrence of a particular genitive may therefore serve as a prime for the selection of the same construction in the next context (Szmrecsanyi et al., 2016a, p. 7). Persistence is generally regarded as a minor factor that carries less weight than other factors like animacy and syntactic weight (Szmrecsanyi et al., 2016a).

There are therefore several language-internal factors that are known to condition genitive alternation; however, none of these factors is deterministic in conditioning which variant will be selected. While some (like animacy, syntactic weight and final sibilance) exert a greater influence on which genitive is more likely to be used, many of these factors have been shown also to interact with language-external factors, such as register and variety.

3.4.3.2 Language-external factors conditioning the genitive alternation

3.4.3.2.1 Genitive alternation across varieties

Researchers interested in the role of language-external factors in genitive alternation typically investigate this by analysing how the strength of (a certain subset of) linguistic factors changes across variety, modality and/or register. As noted in Section 3.4, a substantial amount of work has been done on genitive alternation in AmE and BrE, and it has been established that in the twentieth century the *s*-genitive has increased in frequency vis-à-vis the *of*-genitive, and that this increase in frequency is more advanced in AmE than in BrE (Hinrichs & Szmrecsanyi, 2007; Szmrecsanyi & Hinrichs, 2008). Drawing on the reportage and editorial sections of the Brown family of corpora, Hinrichs and Szmrecsanyi (2007, p. 448) and

Szmrecsanyi and Hinrichs (2008, p. 296) show that between the 1960s and the 1990s, the relative proportion of *s*-genitives increased significantly for both AmE and BrE and that by the 1990s, the *s*-genitive accounted for 53,2% of all interchangeable instances in AmE and 45,8% of all interchangeable instances in BrE. Hinrichs and Szmrecsanyi (2007) find that in AmE, the *s*-genitive is used with inanimate nouns, thematic possessors and in contexts of higher lexical density more often than in BrE, and therefore attribute the increase in the use of the *s*-genitive to a weakening of the effect of animacy and a strengthening of the effects of thematicity and lexical density, particularly evident in AmE.

More recently, researchers have shifted their attention to the genitive alternation in varieties of English other than AmE and BrE. Hundt and Szmrecsanyi (2012), for example, investigate the effects of animacy on genitive alternation in (early) written NZE and compare this to similar written material from a corresponding period in BrE. They find a marginally significant effect for variety, but a significant interaction effect for variety and animacy of the possessor: early NZE writers were significantly less likely to use the *s*-genitive with inanimate, and especially collective and locative, possessors than BrE writers, and therefore the effect of animacy is thought to be much stronger in NZE than in BrE (Hundt & Szmrecsanyi, 2012, p. 256). These findings demonstrate that the linguistic factors presumably carry different weight in different varieties of English.

As part of a project that seeks to explore probabilistic grammars in World Englishes, Szmrecsanyi et al. (2016b) investigate the genitive alternation in BrE, CanE, InE and SngE. Drawing on the ICE sub-corpora for each of the four varieties, they include several linguistic factors in their study (thematicity, givenness, type-token ratio as a measure of lexical density, syntactic weight, final sibilancy and noun phrase expression type) in order to determine how the different linguistic factors interact under the conditions of different varieties. They find that while variety is not the most important predictor for which genitive is selected, it does outperform some language-internal factors (such as thematicity and type-token ratio) and does (potentially) interact with other factors. They also find a distinction in the genitive alternation in terms of whether the variety constitutes a native variety or a non-native variety (the ENL/ESL distinction). Specifically, they find that the two native varieties, CanE and BrE, are more likely to realise proper noun possessors that are highly thematic and that occur in contexts with high type-token ratios in the *s*-genitive than the two non-native varieties, InE and SngE. Although Szmrecsanyi et al. (2016b) do not include animacy as a predictor in their data, they do find that noun phrase expression type of the possessor (coded as Proper noun > gerund > common noun) is by far the most important factor in genitive alternation in their data, and attribute the importance of this factor to its overlap with the predictive potential of the animacy of the possessor.

Following this, Heller et al. (2017b) have investigated the scope and limits of genitive alternation within and across nine different varieties of English (which they very roughly classify along ENL and ESL lines). Using both the spoken and written components of the ICE sub-corpora, the authors investigate genitive alternation in BrE, CanE, IrE, NZE, HkE, InE, JamE, PhiE and SngE. They find that *s*-genitive proportions

are lower in some ESL varieties (JamE, InE and PhiE) compared to the ENL varieties (British, NZE, Irish and CanE), and argue that this is due to a tendency among language learners to avoid inflectional marking and to prefer analyticity (and therefore use the more analytic *of*-genitive in contexts where ENL users might be more likely to use the *s*-genitive) (Heller et al., 2017b, p. 20). While they note that the influence of linguistic factors is fairly stable across the varieties they investigate, they do find some fluidity across the varieties in terms of the strength of the factors that favour the *s*-genitive. The authors conclude that the constraints that favour the *s*-genitive, and in particular animacy, are generally weaker in ESL varieties than in ENL varieties. They do find that SngE and HKE (which are ESL varieties) have higher *s*-genitive proportions compared to the other ESL varieties they investigate, and tentatively hypothesise that the strength of the animacy constraint appears to be proportional to how advanced a variety of English is in terms of the Dynamic Model. They therefore interpret the preference for the *of*-genitive among ESL varieties as a feature of second language acquisition and proficiency, and do not consider that these weaker animacy effects might derive from contact with substrate languages (Rosenbach, 2017).¹²

Heller et al. (2017a) focus on Asian varieties and investigate genitive alternation in South Asian and Southeast Asian English varieties (HkE, InE, PhiE, SngE and Sri Lankan English), which they compare to BrE. In contrast to the findings by Heller et al. (2017b), they find that when the possessor refers to a human, animal or inanimate noun, the Asian varieties (which represent non-native or indigenised varieties) select the genitive construction that is also most likely to be selected by BrE users. However, temporal, locative and collective possessors are realised in the *s*-genitive more often by the Asian varieties than they are in BrE, suggesting that the Asian varieties are more advanced than BrE in terms of the extension of the *s*-genitive to inanimate possessors (Heller et al., 2017a, p. 127). They further find that four factors (animacy of the possessor, frequency of the head noun, syntactic weight of the possessor and possessum, and possessor thematicity) account for the differences between the Asian varieties and BrE. They propose that the Asian varieties have actually developed their own unique preferences for the use of the genitive and note that syntactic weight (a processing factor) is a less dominant predictor compared to animacy (Heller et al., 2017a, p. 135).

Each of these cross-varietal studies demonstrates that while the influence of the different language-internal factors is relatively stable in different varieties of English, there is evidence that the strength of these factors differs across varieties. Rosenbach (2017, p. 14), for example, notes that in cross-varietal studies it is striking that animacy turns out to be the most important factor in genitive alternation, but that it is also the one factor in which there is the most variation in effect strength across the varieties. Despite the substantial body of work on genitive alternation in English, and the growing interest in cross-varietal differences in

¹² Rosenbach (2017) investigates genitive alternation in AfrE (an ESL variety), and WSAfE and BrE (two ENL varieties) using a forced-choice test in which the stimulus is creative writing. She finds that the ESL users in her study use the *s*-genitive far *more frequently* than ENL users, and attributes this a possible transfer effect of the substrate language for the second language AfrE users, rather than due to the processes of second language acquisition (see further discussion of this below and in Section 3.5).

genitive alternation, it is noteworthy that nearly all cross-varietal work to date has focused on the Asian varieties of English as well as CanE, NZE, AmE and BrE. Other varieties, such as AusE and SAfE, are conspicuous in their absence from cross-varietal work on genitive alternation and have each only featured in one study of the genitive alternation.

Ford and Bresnan (2015) conduct a study in which they build a mini-database of genitive and dative structures by asking a group of 40 AmE and 40 AusE users to carry out a series of sentence completion tasks. The focus of their study is on the methodological innovation of using sentence completion tasks as a method for gathering data that can be used as a proxy for unavailable corpus data; however, the study does provide a glimpse into genitive alternation in AusE. They report that they obtained 1509 genitives from the sentence completion tasks, which they annotated for animacy of the possessor, definiteness of the possessor, givenness of the possessor, length of the possessor and possessum, presence of a final sibilant in the possessor and type of semantic relation. Using a mixed-effects regression model, they find no significant main effect for variety, but do find a significant interaction effect for variety and semantic relation:

The Australians are more likely than the American speakers to use an *of*-genitive for nonprototypical genitives. For the Australians, 61.3% of their 475 nonprototypical genitives were *of*-genitives, while 58.2% of their 225 prototypical genitives were *of*-genitives. For the Americans, 57.1% of their 574 nonprototypical genitives were *of*-genitives, while 62.4% of their 237 prototypical genitives were *of*-genitives. (Ford & Bresnan, 2015, p. 213)

Because Ford and Bresnan (2015) do not find a significant main effect for variety, they do not report the frequency of the *s*-genitive and *of*-genitive in the two varieties, but it is possible to calculate this using the frequencies reported in the quote above, which yields the following (relative proportions are reported in brackets):

Variety	<i>s</i> -genitive	<i>of</i> -genitive	Total
AusE	291 (40%)	439 (60%)	730
AmE	335 (41%)	476 (59%)	811

Table 3.2: Distribution of genitives for AusE and AmE (calculated from Ford and Bresnan (2015)).

The data from Ford and Bresnan (2015) show that there is very little difference in the proportion of the *s*- and *of*-genitives across AmE and AusE. When these findings are compared to the frequency data for BrE and NZE from Heller et al. (2017b) (who found that the *s*-genitive is used in 23,3% of all interchangeable instances in BrE and 25,5% in NZE), it is clear that AusE patterns more closely with AmE (in terms of proportional frequencies alone) than BrE or its closest Southern Hemisphere neighbour, NZE.¹³ This suggests that AusE is much more advanced than the two varieties in the increasing use of the *s*-genitive.

¹³ I do not assume here that the two findings from the two studies are wholly comparable, but rather cite the frequencies from Heller et al. (2017b) as a way to demonstrate potential differences between BrE, AmE, NZE and AusE.

One recent study has investigated genitive alternation in SAfE, from both a cross-linguistic and cross-varietal perspective. Rosenbach (2017) investigates cross-linguistic patterns for the effects of animacy on genitive alternation in Afrikaans and SAfE (and specifically WSAfE and AfrE), which she compares to BrE. In addition to a small corpus-based investigation, she conducts an experimental study in which she asks 44 users of WSAfE, 43 users of AfrE and 43 BrE users to complete a contextualised, forced-choice task in which the respondents have to select one of two alternative genitive constructions in 11 sentences over several passages. She specifically includes only constructions with inanimate possessors and controls for other variables by including only common noun possessors, possessors not ending in a final sibilant and only constructions comprising short possessors and possessums. The findings show that users of WSAfE select the *s*-genitive with inanimate possessors least frequently (in 26,45% of all cases), followed by the BrE users (30,2%). The AfrE users use the *s*-genitive most frequently with inanimate possessors (in 37,4% of cases), and significantly more frequently than the two first language groups. In drawing parallels between these findings and the findings of Heller et al. (2017b), she points out that while there is clearly a weaker animacy constraint in second-language varieties, the effect of this weaker animacy constraint in the case of AfrE results in the more frequent use of the *s*-genitive with inanimate possessors rather than a less frequent use of the *s*-genitive with animate possessors (Rosenbach, 2017, p. 15). She further argues that because there is structural parallelism in the inventory of genitive constructions in English and Afrikaans and because inanimate possessors are realised more freely in prenominal position in Afrikaans than in English, the significantly higher use of the *s*-genitive with inanimate possessors by her AfrE participants points to potential substrate influence of the Afrikaans L1, rather than to an impact of second language acquisition (Rosenbach, 2017, p. 15).

3.4.3.2.2 Genitive alternation across modality

Genitive alternation has also been shown to vary across spoken and written language. Szmrecsanyi and Hinrichs (2008), for example, compare data from the Corpus of Spoken American English (CASE), the Freiburg Corpus of English Dialects (FRED) and the reportage and editorial sections of the Brown family of corpora. For BrE, the *s*-genitive occurs more frequently in spoken data (59,6% of all interchangeable instances with the *of*-genitive) than written data from the 1990s (45,8% of all interchangeable instances). In contrast, in AmE, the *s*-genitive occurs more frequently in the written data (53,2% of all interchangeable instances with the *of*-genitive) compared to the spoken data (where the *s*-genitive occurs in 48,2% of all interchangeable instances). They also show that while the *s*-genitive continued to increase in frequency in both written and spoken AmE and BrE in the last half of the twentieth century, the reasons for this increase in frequency are very different for the two varieties, and that these differences can be traced to dissimilarities in the relative weight of the various linguistic factors in spoken and written language. For example, Szmrecsanyi and Hinrichs (2008) find that thematicity is a strong factor in determining which genitive is selected in written texts (where highly thematic possessors favour the *s*-genitive), but it is not an important factor in spoken language. Final sibilancy, on the other hand, has a much stronger effect in

spoken language than in written language, and has grown increasingly influential in the course of the twentieth century. In terms of syntactic weight, Szmrecsanyi and Hinrichs (2008) find that possessor length does not exert any more or less strong influence in either spoken or written language, but possessum length does: longer possessums are more likely to occur with *s*-genitives in written language, but possessum length is not an important factor conditioning the choice in spoken language. Similarly, lexical density is an important predictor in written language (the higher the type-token ratio of the text, the more likely it is that the *s*-genitive is selected), but appears to be unimportant in spoken language. Szmrecsanyi and Hinrichs (2008) therefore conclude that the individual linguistic factors have a different effect on spoken and written language, and that for press language it is the register-internal process of economisation that drives the increase in the use of the *s*-genitive, rather than the colloquialisation of the written norm. They argue that journalists have come to prefer the more compact *s*-genitive with highly thematic possessors and in the increasingly lexically dense press writing environment (Szmrecsanyi & Hinrichs, 2008, p. 307).

Jankowski (2013), too, investigates differences in genitive alternation in spoken and written English, but focuses on CanE. To do this, she constructs two real-time corpora spanning the twentieth century, one representing written press language (drawn from *Maclean*, a national Canadian news magazine that is based in Toronto) and the other representing spoken language (drawn from the Federal Hansard transcripts of the Canadian House of Commons – although she does acknowledge that the Hansard corpus straddles the boundary between spoken and written language (Jankowski, 2013, p. 26)). In analysing all interchangeable instances of the *s*-genitive and *of*-genitive constructions coded for animacy, possessor and possessum length and final sibilancy of the possessor, Jankowski (2013) shows that during the course of the twentieth century, the *s*-genitive rises in frequency in both the written and spoken corpora, with the spoken corpus being somewhat more conservative than the written corpus. She also finds that in the course of the twentieth century, the *s*-genitive comes to be used increasingly frequently with inanimate possessors in her written corpus compared to her spoken data and therefore shows that while the expansion of the *s*-genitive to inanimate possessors is not unique to written language, it has progressed more quickly in her written language data. Based on this she concludes that written language (and in particular written news language) is most likely the locus of the ongoing change in genitive alternation (Jankowski, 2013, pp. 109–110).

3.4.3.2.3 Genitive alternation across written registers

As is clear from the discussion on genitive alternation across modalities, nearly all work on written texts has drawn on corpora comprising journalistic writing, with very little work having been done on other kinds of written published registers. Journalistic writing has been shown to be particularly receptive or open to change, much more so than other genres (Hundt & Mair, 1999), and it has certainly been shown to lead the way in the increasing use of the *s*-genitive as well as the extension of the *s*-genitive to inanimate possessors. But the question of how the ongoing change in genitive alternation extends to other written published registers is less clear. One study that suggests that the extension of the use of the *s*-genitive to inanimate nouns in written registers is not limited to press writing only, is Rosenbach (2003). Drawing on

experimental data in which fiction is used as the stimulus, Rosenbach (2003) finds that the increasing use of the *s*-genitive with inanimate nouns is also a development to be found in fiction. Grafmiller (2014) addresses the gap in register differences by examining genitive alternation across spoken and written language, and across different written published registers in order to determine the relation of genitive choice in journalistic writing to other written registers and to spoken language. He argues that an investigation of genitive choice is necessary because written registers vary along numerous dimensions that might also influence the weight of different linguistic variables (Grafmiller, 2014):

For instance, given the slower, careful construction of professional writing, it may be that the influence of factors related to real-time processing, e.g. end weight, is diminished in certain styles. Other factors, e.g. rhythmic structure or semantic relation, might play a larger role, due perhaps to greater room for artistic playfulness in some genres (e.g. fiction). (p. 474)

Focusing on AmE, Grafmiller (2014) draws on the Penn Treebank portion of the Switchboard corpus of American English (to represent spoken language from the early 1990s) and the Boston University Noun Phrase Corpus which comprises genitive tokens from five different sections of the Brown Corpus of written AmE (representing written language from the 1960s): Press, Non-fiction, Learned, General fiction and Western fiction (Grafmiller, 2014, p. 474). The two corpora are therefore not aligned in terms of timeframe. Only interchangeable instances of the *s*-genitive and the *of*-genitive are considered and all instances are annotated for animacy, semantic relation, information status, thematicity, final sibilancy of the possessor, end-weight, persistence and type–token ratio (Grafmiller, 2014, pp. 476–479).

Overall, Grafmiller (2014) finds that the proportional preference for the *s*-genitive is lower than for the *of*-genitive across both spoken and written language (see Table 3.3). With regard to differences across the written registers, he finds that the difference in the proportional frequencies of genitive constructions is highly significant for the general fiction, non-fiction and learned registers and that there is very little difference in western fiction and no difference in the press register (Grafmiller, 2014, p. 476).

Genre	<i>of</i> -genitive		<i>s</i> -genitive		Total Count
	Count	Percentage	Count	Percentage	
SPOKEN	659	59,1	456	40,9	1115
WRITTEN	1612	64,6	885	35,4	2497
<i>Learned</i>	314	77,8	90	22,2	404
<i>Non-Fiction</i>	623	67,8	296	32,2	919
<i>General fiction</i>	227	63,4	131	36,6	358
<i>Western fiction</i>	221	55,3	178	44,7	399
<i>Press</i>	227	52,4	206	47,6	433

Table 3.3: Distribution of genitive constructions across modalities and genres (adapted from Grafmiller (2014, p. 476)).

To understand how individual factors might contribute to the differences between the written registers, Grafmiller (2014, p. 487) uses a multilevel, mixed-effects logistic regression model in which press is treated as the baseline of a five-level register factor. The results show that the following factors, in order of

significance, account for the most significant difference in the proportions of the two constructions in the written registers: animacy > register > presence of final sibilant > length ratio > possessor givenness > text frequency of the possessor. In other words, apart from animacy, register is a clear factor in influencing the proportional preferences of the two constructions across the different registers: the *s*-genitive occurs with a higher proportion in press writing, compared to non-fiction, general fiction, western fiction and learned writing. Final sibilance of possessor and length ratio significantly disfavour the *s*-genitive, while a high text frequency of the possessor and given possessors favour the *s*-genitive (Grafmiller, 2014, p. 490). Semantic relation, persistence and type–token ratio have no significant main effects. In terms of the interaction of factors, Grafmiller (2014) finds significant positive interactions between length ratio and general fiction, learned and western fiction, suggesting that syntactic weight is an important factor in these registers and that longer possessors are much more likely to be realised in *s*-genitives in these three registers than in non-fiction and especially press. In addition, he finds a significant positive interaction for all five registers and animacy, but notes that the influence of animacy increases for every register relative to press and that the effect of animacy is weaker in press than in any of the other written registers, and most strong in spoken language (Grafmiller, 2014, p. 491). To demonstrate this decreasing strength of animacy across modalities and registers, Grafmiller (2014, p. 491) ranks the registers along the following scale:

Spoken > Fiction, Narrative non-fiction, Academic > Press

The last significant interaction effect he finds is between non-press writing and press writing and final sibilancy of the possessor: in non-press registers (fiction, narrative non-fiction and academic) the presence of a final sibilant on the possessor significantly increases the likelihood that the *of*-genitive will be used, while in press writing the presence of a final sibilant on the possessor has no effect on which genitive is selected (Grafmiller, 2014, p. 492). The findings of the study suggest that in written language, animacy, genre, final sibilancy of the possessor, thematicity and givenness are the strongest factors conditioning the genitive alternation across different written registers, but that there is variability across different registers with respect to the importance of some linguistic variables (Grafmiller, 2014, p. 490).

3.5 Motivation for the selection of the varieties, registers and feature investigated in this thesis

The different sections of this chapter have so far set out discussions of the different varieties, registers and feature selected for investigation in this study. The aim of this thesis is to investigate the interaction between broader trends of language change, the different contextual forces of the varieties of English, register effects and editorial intervention in influencing language variation and change. The motivation for the selection of the two varieties, five registers and one feature for this study is couched in this aim.

In terms of broad trends of ongoing language change, the selection of the genitive alternation is a particularly useful feature to investigate for several reasons. First, the genitive alternation is a feature that

is currently undergoing change as reflected in the ongoing increase in the use of the *s*-genitive at the expense of the *of*-genitive. This pattern has been found in several varieties of English, and the feature therefore constitutes a type of “global feature” (Hundt, 2007). There is much disagreement on the reasons for causes of this ongoing change in written language, with some scholars arguing that the shift in preference is due to the broader stylistic trend of colloquialisation, while others argue that the increasing use of the *s*-genitive is linked to the broader trend of economisation. Since different written published registers have been shown to be variously receptive to the features associated with ongoing change and variously influenced by the broader stylistic trends, different written published registers differ in how they take up features associated with ongoing change. Given this it is likely that there will be differences across different written published registers in terms of their receptivity to the incoming *s*-genitive. To date, only one study has investigated genitive alternation across different written published registers (Grafmiller, 2014), but this study draws on data from the Brown corpus, and the data might therefore not reflect current usage. The selection of the five written published registers in this study is therefore motivated by the dearth of research on genitive alternation across different written published registers in Present-day English.

Since a key focus of this thesis is the influence of editorial intervention, the registers included are all registers that are known to undergo editorial intervention as part of the processes through which they are produced. Furthermore, English grammars provide guidance on the use of the two constructions, and advise that *s*-genitives should be used with animate (human) possessors, while *of*-genitives are preferred with inanimate possessors because the apostrophe *s* attributes the power of possession to the head noun of the possessor and inanimate nouns do not have the power to possess. For example, Butterfield (2015, p. 721) states that *s*-genitives are commoner with nouns that represent humans or animals while the *of*-genitive is preferred in all other cases. However, the normative advice given is not particularly widespread, nor is the feature strongly normatively marked, which opens up the space for variation and particularly for editors’ (and authors’) own linguistic experiences (and resulting grammatical representations) to influence their choices.

Given that different registers participate in the processes of ongoing change in differential ways, adherence to the recommendations in English grammars is obviously a matter of degree. Furthermore, despite these recommendations, *s*-genitives have been shown to be increasingly used with inanimate possessors (which occur at a higher rate in some registers, like the academic and instructional registers, compared to other registers, like reportage and creative writing). Since editors are linguistic gatekeepers who enforce selected usage practices and since editors are likely to have a well-developed sensitivity to the stylistic preferences of the different types of texts that they edit, an investigation of the interplay between prescription and editors’ own linguistic representations (that is their covert norms) might point to the extent to which the extension of the *s*-genitive to inanimate possessors has become conventionalised in Present-day English and what the stylistic constraints on those extensions are.

In addition to the different registers' receptivity to features associated with ongoing change (and editors' potential influence in this openness), globally used grammatical patterns have also been shown to display divergent behaviour in local contexts (Hundt, 2007). As demonstrated in Section 3.4, there is a substantial body of research on the historical development of genitive alternation, as well as the various language-internal and language-external factors that condition the use of the different variants. The work on language-external factors, and especially investigations of genitive alternation across different varieties of English has generally focused on differences between AmE and BrE, but recently the cross-varietal approach has been extended to include other varieties of English. These cross-varietal studies have shown that some varieties are more advanced than others in their openness to the *s*-genitive, with some scholars arguing that the difference may be accounted for in the distinction between ENL and ESL varieties, as well as how far a variety has progressed along the stages of the Dynamic Model. For example, Heller et al. (2017b) and Rosenbach (2017) all find significant differences in the proportion of *s*-genitives across the ENL and ESL varieties that they investigate. However, in Heller et al. (2017b), the proportion of *s*-genitives is consistently lower in the ESL varieties they investigate compared to the ENL varieties investigated, and they conclude that this is due to a preference for the more analytic, transparent *of*-genitive in learner varieties. Rosenbach (2017) finds the opposite to be true in her study of AfrE (an ESL variety), and WSAfE and BrE (two ENL varieties). Rosenbach (2017) finds that the ESL users in her study use the *s*-genitive far *more frequently* than ENL users, and attributes this a possible transfer effect of the substrate language for the second language AfrE users, rather than due to the processes of second language acquisition. The stronger preference for the *s*-genitive in AfrE may therefore be linked to the fact that it is fairly advanced in terms of the Dynamic Model (see Section 3.2.2), which Heller et al. (2017b) propose as a possible reason for the higher proportion of *s*-genitives in the data for some of their second-language varieties, or it may be due to the substrate influence of the Afrikaans L1 (see Section 3.4.3.2.1).

Hardly any of the cross-varietal work on genitive alternation done to date has included AusE and SAfE. The dearth of work on these two varieties alone constitutes a good enough reason to conduct a cross-varietal investigation of the two varieties; however, further motivation is to be found when considering the different status of English in the two countries and the (conflicting) findings of Heller et al. (2017b) and Rosenbach (2017). The motivation for the selection of AusE and the (sub-varieties of) SAfE in this thesis therefore hinges on two assertions. First, as discussed in the sub-sections of Section 3.2, AusE and the different sub-varieties of SAfE have all progressed through the stages of the Dynamic Model at different rates, with AusE being more advanced than any of the SAfE sub-varieties, and within SAfE, WSAfE and AfrE being more advanced than BSAfE. Based on the assertion by Heller et al. (2017b) that the proportion of *s*-genitives is higher in varieties of English that have progressed further along the Dynamic Model, one would expect AusE to have a higher proportion of *s*-genitives than WSAfE and AfrE, with BSAfE lagging even further behind. But Rosenbach (2017) shows this not to be the case in her investigation of genitive alternation in BrE, WSAfE and AfrE, where her data reflect a significantly stronger preference for the *s*-genitive (with inanimate possessors) among AfrE users, compared to BrE and WSAfE. Rather than attributing this to a

feature of second language acquisition, Rosenbach (2017) argues that substrate influence might actually account for different proportions of the two constructions in second-language varieties. This assertion can be very closely linked to Van Rooy's (2014a) argument that in contexts of high contact, different strands of English come into contact with each other at different times and in different settings and this differentially influences the progression of these varieties towards endonormative stabilisation. In a country like South Africa (where multiple second-language varieties are represented alongside the first-language variety) the local contact setting, substrate influence and differential progress of the sub-varieties along the stages of the Dynamic Model, as well as their status as either first- or second-language varieties, might all point to (and account) for different usage preferences in the genitive constructions.

This complex web of different sub-varieties opens up the possibility of contact effects of other substrate languages in South Africa on genitive alternation, but it also allows for an investigation of the effect of a sub-variety's progression along the stages of the Dynamic Model on genitive alternation. With regard to the possible influence of substrate languages, the first languages of AfrE and BSAfE users are particularly interesting. Like English, Afrikaans offers two devices for the expression of genitive relations: one in which the possessor is realised prenominally and is linked to the possessum via the *se* relational marker (henceforth the *se*-genitive), and one in which the possessor is realised postnominally and is linked to the possessum via the *van* relational marker (henceforth the *van*-genitive). There is therefore structural parallelism between English and Afrikaans in respect to genitive constructions (see Table 3.4), with strong overlap in the structure and semantics of the prenominal genitive constructions (Rosenbach, 2017). However, as noted previously, Rosenbach (2017) finds that Afrikaans uses the prenominal *se*-genitive more frequently with inanimate possessors than does English, and therefore concludes that the animacy constraint is weaker in Afrikaans than it is in English. According to Rosenbach (2017), it is this structural parallelism and the weaker animacy constraint in Afrikaans that leads AfrE users to select the *s*-genitive more often in contexts where native English users are more inclined to select the *of*-genitive.

The inventory of genitive constructions in BSAfE is somewhat different. Table 3.4 shows that neither Southern Sotho nor Zulu (two south-eastern Bantu languages representing one language from each of the two major Bantu language groups in South Africa) offer a device in which the possessor is realised prenominally. Only the postnominal construction is available. Furthermore, Southern Sotho only has *ya* available as the relational marker for the postnominal genitive construction, while the relational marker in Zulu varies depending on the noun class of the possessor (for example, all human nouns take *ka*, while head nouns that are loan words take *ye*). There is therefore limited structural parallelism between English and the South African Bantu languages in terms of the constructions that are available to express genitive relations. If substrate languages do influence the genitive choices of second-language users, then it can be expected that the genitive choices of BSAfE users would be overwhelmingly in favour of the *of*-genitive. Furthermore, because only the postnominal genitive construction is available in the two Bantu languages,

it is likely that the various linguistic factors known to influence genitive alternation in English will be substantially weaker in BSAfE.

	Prenominal genitive			Postnominal genitive		
	possessor	relational marker	possessum	possessum	relational marker	possessor
English	<i>Harry</i>	's	<i>heart</i>	<i>the corner</i>	<i>of</i>	<i>the street</i>
Afrikaans	<i>Harry</i>	<i>se</i>	<i>hart</i>	<i>die hoek</i>	<i>van</i>	<i>die straat</i>
Southern Sotho		∅		<i>pelo</i> [heart]	<i>ya</i> [of]	<i>Harry</i> [Harry]
				<i>khona</i> [corner]	<i>ya</i> [of]	<i>seterata</i> [street]
Zulu		∅		<i>inhliziyo</i> [heart]	<i>ka</i> [of]	<i>Harry</i> [Harry]
				<i>ikhona</i> [corner]	<i>ye</i> [of]	<i>stratu</i> [street]

Table 3.4: Genitive constructions available in Afrikaans, English, Southern Sotho and Zulu.

The analysis and comparison of genitive alternation in AusE and the different sub-varieties of SAfE therefore offer an excellent opportunity not only to extend cross-varietal investigations of this alternation to varieties that have largely been excluded from such studies, but also to investigate the (potential) influence of two substrate languages (with varying degrees of structural parallelism with English in the devices available to express genitive relations) on the genitive choices made by different kinds of second-language users, compared to first-language users of two different varieties.

The inclusion of a third, language-external factor, namely editorial intervention, is the final, major innovative extension of the analysis in this thesis. In AusE, given that English is the dominant language, it is probable that editors of AusE are native English users and therefore they are exposed to a homogenous normative environment – much more so than South African editors of English texts. As discussed in Section 1.2, a review of the South African Translators' Institute (SATI) database of English editors shows that native users of English constitute 24% and Afrikaans–English bilinguals 44% of this group (SATI, 2015). Furthermore, only 8% of this group of editors reports one of the nine official African languages as their first language, while the remaining 23% report a European language as their first language. It is also striking that none of the editors who reports an African language as a first language is accredited by the institute – suggesting that there are no accredited BSAfE editors in South Africa. This mix of first- and second-language English editors in South Africa, coupled with the possible influence of the Afrikaans substrate on the genitive choices of AfrE editors and the fact that converging tendencies among the different strands of English in South Africa have been weaker than in other PCEs, suggests that the editorial choices of SAfE editors might be quite different from those of AusE editors and that the strength of some linguistic factors in conditioning the alternation between the two constructions will also differ considerably, as demonstrated in the findings by Rosenbach (2017) for WSAfE and AfrE. Furthermore, the fact that there are no editors representing the BSAfE sub-variety places a particularly stringent bar for the acceptance of BSAfE usages, as editors who work on the texts produced by this community are either WSAfE or AfrE users. For AfrE writing, the possibility exists for usages particular to AfrE to find their way into printed texts – if these texts

are edited by AfrE users. In addition, the limited proportional number of WSAfE editors means that opportunities for the dissemination of usages associated with other sub-varieties, and in particular AfrE, is greater. Lastly, editors in South Africa are faced with a heterogenous linguistic environment, which may feed into their own linguistic representations, diluting and shifting their own probabilistic grammars in different ways over time. Lastly, if there is very little difference in the kinds of usages accepted by first- and second-language editors who edit first- and second-language authors' writing, then this might suggest convergence between the sub-varieties, and would provide evidence of endonormativity.

3.6 Conclusion

This chapter has discussed the varieties, registers and feature selected for investigation in this study. Section 3.2 explored the development and current state of the two varieties selected, AusE and SAfE. Specifically, this section established that the two countries have progressed differentially along the stages of the Dynamic Model. AusE has recently entered Phase 5 of the Dynamic Model, having consolidated its own norms and emerging as an independent national standard and carrier of Australia's national identity. Furthermore, AusE is widely considered to be a classic Inner Circle variety. The discussion of the South African context demonstrated that the linguistic landscape in this country is quite different. The country has still not achieved national reconciliation and has not managed to rebuild itself as one nation, with distinct social and ethnic identities persisting, alongside several varieties of English (Schneider, 2007; Van Rooy, 2014a). Although the STL strand of WSAfE is a visible and influential sub-variety, it is only used by a minority of South Africans, with the overwhelming majority of English users being BSAfE users, followed by AfrE users. To this end, this chapter argued that rather than viewing English in South Africa as one overarching variety, a more complete and nuanced analysis is one in which the local contexts of each sub-variety are considered. The discussion showed that while converging tendencies have been weaker in South Africa, compared to other countries, some convergence can be found across the three sub-varieties investigated in this chapter, as reflected in a growing tolerance for (and therefore acceptance of) a range of features at the level of pronunciation and lexis, with some recent work that points to some convergence in grammar (Van Rooy, 2014b). In keeping with the focus of this study on the role of editing, some evidence was presented to demonstrate the influence of editorial intervention in the progression of the varieties towards endonormativity. This discussion pointed to the potential influence of editorial intervention in ongoing language change in the different varieties.

Section 3.3 discussed the five registers selected for inclusion in this study: academic, creative, instructional, popular and reportage. The discussion in this section highlighted differences and similarities across the registers and established that the five registers included in this study represent a diverse range of registers that are distinguished from one another in terms of their functions and purposes, intended readers, their receptivity to the features associated with ongoing change as well as how they have participated in broader trends of language change. This section also considered how the nature and scope of editorial intervention

differs across the five registers and established that the five registers vary with regard to production circumstances and the nature and scope of editorial work.

The genitive alternation, the feature selected for analysis in this study, received a substantial amount of attention in this chapter. The discussion in various sub-sections of Section 3.4 first provided an explanation of the two constructions that alternate (Section 3.4.1), before moving on to a brief exposition of the historical development of the two constructions (Section 3.4.2). Section 3.4.3 explored the various factors known to condition the alternation between the two constructions and showed that various linguistic factors are well known to influence the alternative realisation of the two variants. Crucially, this section also explored how the strength of these linguistic factors varies across different varieties of English and across different registers.

These three sections laid the foundation for the motivation for selecting AusE and the different varieties of SAfE, the five written published registers and the genitive alternation as test cases to analyse how editorial intervention interacts with broader processes of language change, specific change in varieties of English and register effects to influence different opportunities for and constraints on the processes of dissemination and conventionalisation in a variety's progression towards endonormativity. The next chapter, Chapter 4, describes the various aspects of the research design and methodology used to investigate this interaction in this thesis.

Chapter 4

Methodology: Combining survey and corpus research to investigate the effects of editing in written varieties of English

4.1 Introduction

Research questions 2.1, 2.2 and 2.3 of this study have an empirical focus. Together, research questions 2.1 and 2.2 seek to gather and analyse information on the (social and linguistic) profiles of editors of English texts in different varieties of English as well as the kinds of norm-providing sources that these editors use to guide their editorial choices and whether or not these sources reflect the ‘norms’ of the varieties in which they edit. Research question 2.3 seeks to investigate how editors working in different varieties of English respond to the presence of genitive alternation, as a feature associated with language variation (see Section 3.4), and how this response serves to accelerate or inhibit variation (and change) in the varieties in which they edit. To this end, this study draws on a quantitative approach in which multiple methods are used to answer each of the three questions. To answer research questions 2.1 and 2.2, a survey focused on gathering quantitative data on the social and linguistic profiles of editors of AusE and SAfE and the norm-providing sources used by these editors was used. To answer research question 2.3, a corpus-based variationist approach was used to investigate the interaction of editorial intervention with various language-external and linguistic factors in conditioning genitive alternation in AusE and SAfE. This study therefore serves as a test case for how editorial intervention might form part of processes of variation and change, and how this might influence the progression of different varieties of English towards endonormativity. Although the approach used in this study is primarily quantitative in nature, drawing on distributional analyses and statistical modelling of the data, qualitative analyses are employed throughout, where appropriate, to further probe and unpack the findings from the quantitative analyses of the data.

This chapter discusses the research methodology used in the collection, analysis and interpretation of the questionnaire and corpus data used to answer research questions 2.1, 2.2 and 2.3, and provides a brief discussion of the ethics considerations and process for obtaining ethics clearance for the collection the data. The first two sub-sections of Section 4.2 outline the design and piloting of the questionnaire. Section 4.2.3 discusses the population and sampling procedures followed in the selection of the sample for the survey research in this study, as well as the distribution of the questionnaire and subsequent collection of responses. Section 4.2.4 describes how the collected questionnaires were processed and the data analysed. Section 4.3 shifts the focus to the corpus-based component of the investigation and outlines the design of the corpus used in this study in Section 4.3.1, and the construction and size of the corpus 4.3.2. Section 4.3.3 offers a discussion of the data extraction and annotation procedures followed for the corpus data, while Section

4.3.4 explains how the corpus data were analysed. Section 4.4 outlines the various ethics considerations relevant to the collection of the questionnaire and corpus data and provides information on the ethics clearance obtained for the collection of the data in this study.

4.2 Questionnaire

The first part of the study involved the distribution of a questionnaire among editors of English texts in Australia and South Africa. The purpose of the questionnaire was twofold: in the first instance, it sought to gather information on the demographic and language profiles of editors of English texts in Australia and South Africa, and in the second it sought to gather information on the kinds of sources that these editors draw on to guide their editorial choices as well as the norm-orientation of these sources. These data were gathered in order to answer research questions 2.1 and 2.2, which seek to investigate the linguistic and social background of editors in the two countries, and to determine what overt norm-providing sources editors of English texts in Australia and South Africa use in the course of their editing work. These two sets of data serve as useful indices of the overt and covert norm orientations of editors: the norm-providing sources that they use mark their explicit overt norms, while the information on their sociolinguistic backgrounds provide information on their covert norms (as discussed in Section 2.6).

4.2.1 Questionnaire design

The questionnaire drew on a combination of open- and closed-ended questions with fixed-alternative responses spread over two sections. The closed-ended question format was selected to facilitate data processing; however, open-ended questions were included to allow respondents the opportunity to provide additional information where the researcher anticipated such information would be available. Since the target population included editors that are based in two different countries, the questionnaire took the form of self-completion online questionnaire administered via Qualtrics (www.qualtrics.co.uk).

To answer research question 2.1, the first section of the questionnaire focused on the collection of respondents' demographic information, such as their age, gender and level of education, but also included questions on the respondents' language profiles, such as their first or home language, all the languages they know and all the languages in which they do editing work. This information is particularly important within the context of this study as it allows for an investigation into the language profiles of language users who are the linguistic gatekeepers in the two varieties, and provides opportunities to reflect on the grammatical representations of language of these linguistic. Furthermore, in variationist sociolinguistic investigations, social factors such as age, gender and education level are known to account for differences in language use. For example, there is a correlation between gender and a preference for standard or prestige forms, as well as between education level and age (Labov, 2001; Tagliamonte, 2012). Thus, in order to gain a full understanding of the profiles of editors, it is necessary to gather information on their sociolinguistic profiles.

To answer research question 2.2, the questions in the second section were designed to gather information on the kinds of norm-providing sources that the editors in the two countries use to guide their editorial choices. Four questions, each related to a particular kind of norm-providing source commonly used by editors, were included in this section. These four norm-providing sources were online dictionaries, print dictionaries, usage guides and style manuals. For each of the four questions, a list of sources was provided from which the respondents could select the sources that they used. In each instance, respondents could select more than one option, up to a maximum of five. Each question also included an ‘Other’ option, which could be used to list sources not included in the list of options provided. The sources included as options in these four questions were drawn from lists of sources recommended for editors by two editors’ organisations in Australia and South Africa: the Institute of Professional Editors (IPEd) in Australia (www.iped-editors.org) and the Professional Editors’ Guild (PEG) in South Africa (www.editors.org.za). Both lists are publicly available on the organisations’ websites. In addition to these four questions, an additional open-ended question was included at the end of the questionnaire so that respondents could list any other (kinds of) sources not covered by the first four questions.

An original version of the questionnaire included a third section that intended to gather information on the kinds of problem-solving strategies that editors use in the course of the work. Such strategies included the use of specific norm-providing sources, consultation with other editors, relying on own intuition or relying on an already-developed strategy. However, following the piloting stage, the decision was made to exclude this section from the questionnaire.

4.2.2 Piloting of the questionnaire

Piloting is considered a crucial stage in the development of a questionnaire because it is through the piloting of a questionnaire that the researcher is able to assess the usability, clarity and appropriateness of the questionnaire as well as whether it measures the complete content of the construct under investigation (Saldanha & O'Brien, 2013). To this end, the piloting of the questionnaire in this study was undertaken in order to determine the face and content validity of the questionnaire and to gauge how it would perform under the conditions of data collection.

Six editors, three each from Australia and South Africa, were contacted via email and invited to participate in the pilot. All of the participants were editors of English texts and represented editors with a diverse range of experience and expertise. One editor from Australia and one editor from South Africa worked in academia and are familiar with questionnaire-based research, while the remaining four participants were editing practitioners with extensive experience in a diverse range of editing contexts. The participants in the pilot therefore represented potential participants in the sample population.

All participants were informed of the purpose of the study and the questionnaire, and were provided with detailed information on what their participation in the pilot would entail. Specifically, the participants were

asked to complete the questionnaire using the online questionnaire link and to provide feedback on the design and layout of the questionnaire, the wording and order of questions, the content and formulation of the questions and the range of answer options provided. The participants were also asked to note down the time it took to complete the questionnaire. All six editors agreed to participate in the pilot and received the electronic link to the questionnaire as well as instructions for how to provide feedback in July 2017 (see Appendix A). One participant from Australia withdrew their participation after having received the questionnaire; however, the remaining five all completed the questionnaire and returned their feedback in August 2017.

Overall, the participants made recommendations for minor adjustments to the phrasing of some questions as well as recommendations for the addition of answer options to certain questions. For example, for the question on whether participants worked on a full-time or a part-time basis, the original question included only these two answer options, but one participant pointed out that some editors may be employed on a full-time basis while simultaneously doing part-time editing work after their full-time working hours. To accommodate such cases an additional option, 'Both', was provided.

The majority of the comments were directed at the questions that sought to gather demographic information. However, all of the participants commented on the third section of the questionnaire. This section, as mentioned above, asked the participants to indicate if they make use of particular problem-solving strategies in the course of their editing work, and if they do, to provide examples of the kinds of problems that they solve using each strategy. Ten strategies were listed. All of the participants commented that they struggled to complete this section with some stating that they were unsure what to do. Others felt that the completion of this section would take too long. After considering all the participants' comments as well as the purpose of the questionnaire, the decision was made to remove the third section from the questionnaire since the kind of data it would collect would not contribute to the primary aims of the questionnaire, which were to gather information on editors of English texts in Australia and South Africa and the kinds of norm-providing sources they use. The decision to remove the third section was also motivated by the fact that an extremely long questionnaire could potentially reduce the response rate. Thereafter, the questionnaire was finalised and distributed (see the final questionnaire, included as Appendix B).

4.2.3 Participants and sample

The target population for the questionnaire was all editors of English texts working in Australia or South Africa. No age range was imposed on the selection of participants; however, it was expected that all participants would have completed their schooling and would therefore be 18 years or older.

Since no centralised database of editors in Australia or South Africa exists, it was unfeasible to draw the sample of respondents for the questionnaire from one list of editors for each country. To try to ensure that the sample of editors was as representative of the population as possible, various sampling frames and

techniques were used. First, the mailing list co-ordinators of various editors' organisations, language services agencies, media houses and publishing institutions in Australia and South Africa were contacted via email with a request to distribute the invitation to participate in the questionnaire to their members. Second, various editors within the researcher's professional network were also contacted via email with the invitation to complete the questionnaire and a request to distribute the invitation to any qualifying editors that they may know. Each e-mail included a short description of the study as well as a request to distribute the invitation to the organisation's members or qualifying editors. The invitation was attached as a pdf and included a short description of the study and the purpose of the questionnaire as well as a link to the electronic questionnaire. If potential respondents clicked on the link, they were redirected to the online questionnaire, where they could access and read the Participation and Information Consent Form, before consenting to participate in the study. The invitation to complete the questionnaire was also posted to the researcher's Facebook account as well as the Facebook accounts of two of the study's supervisors, from where it was shared by other users in the researcher's and supervisors' networks.

The invitation to complete the questionnaire was sent out in September 2017 and was available online for completion during September and October 2017. A total of 163 responses was collected via Qualtrics. One respondent declined to participate and was therefore unable to complete the questionnaire, while 26 responses were incomplete. Thus, of the 163 responses, 136 were complete and could be included in the study.

4.2.4 Processing and analysis of questionnaire data

Qualtrics, the online platform used to disseminate the questionnaire, provides a tool in which all completed responses can be exported to a Microsoft Excel workbook. Using this tool, the raw data are collated into one master spreadsheet where each individual respondent's answer(s) to each question is recorded. The raw data were exported from Qualtrics into an Excel spreadsheet, which was manually inspected and cross-checked against each recorded response on the Qualtrics platform.

The raw data were cleaned up and each participant was allocated a code that identified them as either an Australian editor or a South African editor. Australian editors were allocated a code using the *AusE* identifier, followed by a number. Similarly, the South African editors were allocated a code using *SAfE* as identifier, followed by a number.

The respondents' answers to the demographic information were manually cleaned up. Since most of the questions in this section made use of fixed-alternative answers, this process was relatively straight-forward. Where participants were required to enter information in answer to a question, these answers were manually inspected and standardised. For example, the questions dealing with the respondents' strongest languages and strongest working languages required that the respondents type in their languages. Because of this, the answers to all of these 'fill-in' questions required some processing to check for and correct any spelling

errors and/or to standardise the responses to facilitate later processing. Once this manual processing was complete, the data from the demographic questions were analysed descriptively and processed as percentages per response per country. These calculations were all done in Excel, using the various data processing and pivot table tools.

The data gathered from the questions regarding the respondents' use of the different norm-providing sources were much more complex to process, for two reasons. First, the respondents were asked to provide edition information for the options that they selected from the list of available options. Different kinds of edition information were provided in different formats, and these needed to be standardised. Second, for each question that dealt with a particular source type, an 'Other' option in which respondents could list other sources not provided for in the list of options, was provided. In addition to this, one open-ended question was provided at the end of the questionnaire where the respondents could list any other types of sources not covered by the four source type categories. In both cases, the respondents were required to provide bibliographic information of the 'Other' sources that they listed, which required some cleaning up and standardisation.

The data on the different sources used by the respondents therefore required some pre-processing, which was carried out in two stages. During the first stage, the responses were cleaned up, cross-checked for accuracy and standardised. During the second stage, the range of sources and the kinds of sources listed in each 'Other' section were manually checked and the bibliographic information verified, whereafter the data were standardised and re-categorised.

With regard to the first stage, all edition information on the sources that the respondents selected from the list of options provided had to be checked and standardised because this information had been indicated using either the edition number or the year of publication. To standardise this information, the decision was made to standardise all instances to the year of publication as this would allow for the comparability of the different sources in terms of their age, something that would not have been possible if the edition number had been used. As such, each response had to be manually checked, firstly to confirm the year of publication where the respondent had provided this and secondly to determine the year of publication where the respondent had provided the edition number. All other bibliographic information, such as author(s), place of publication and publisher also had to be checked and standardised for each edition listed. This information is particularly important for sources that have a long publication history in which various editions have been revised by different people (as in the case of *Fowler's Dictionary of Modern English Usage* (Butterfield, 2015) and *New Oxford Style Manual* (Ritter, 2012)) and/or published by different publishers.

The second stage of processing the source data was related to the fact that in using the 'Other' options, many respondents indicated that they make use of different kinds of sources that extend well beyond the scope covered by the four categories. Furthermore, there were also numerous cases where respondents had

listed a source under the incorrect category. To try to make sense of this data, the decision was made to extend the four categories initially used to categorise the sources to nine categories and to manually (re)classify the sources listed by each respondent under one of these nine categories. The nine broad categories of sources are: dictionaries, glossaries, thesauruses, style guides, usage guides, editing and writing sources, language reference works, people, and various online and print sources.

Straaijer (2017) notes that the borders between different language (reference) works are fuzzy since there may be some overlap in the content between the texts of particular genres of reference works. To distinguish among different kinds of reference works, Straaijer (2017) turns to function in order to establish criteria for the identification of different genres. The same principle was applied to the categorisation of the different types of sources mentioned by the respondents in the questionnaire of this study. Four of the categories (dictionaries, thesauruses, glossaries and people) were relatively simple to delineate and identify.

In this study, dictionaries are defined as alphabetical compilations, available in print, online and/or electronic format, of the words within a language and which provide information on the words' spelling, pronunciation, meaning, word class as well as examples of use (Siegel, 2018). They may be monolingual or bilingual. As discussed in Section 2.4, in the context of World Englishes, and specifically the endonormative stabilisation of a variety, dictionaries serve as important codification instruments for a variety of English and are typically marked as reflective of that variety in their titles (Schneider, 2007). Glossaries, like dictionaries, are also online or print compilations of words, but unlike dictionaries they do not cover all the words of a language and are typically selective in terms of the words covered. As such, they provide and/or develop standard terminology within a particular topic and clarify spelling. They may provide definitions of words, and some may even provide equivalent terms in one or more other languages. Thesauruses are lists of words, available in print, online and/or electronic format, for which synonyms, associated words and antonyms are provided. They may be arranged alphabetically or by theme. The category 'people' included all sources where the respondents indicated that they consult another person to resolve a problem that they encounter in the course of their editing work. Examples of these include consulting colleagues, subject experts, other editors and online or the social-media chat groups of editors' or language practitioners' organisations.

The delineation of the categories 'style guides', 'usage guides', 'editing and writing sources', 'language reference works' and 'various online and print sources' was much more problematic than that of the four categories outlined above, because there is a higher degree of overlap in the content and functions of the sources in these categories. Style guides, which are developed for use either within a particular organisation (such as a book publisher or a scientific association) or for more general use, address a range of mostly orthographic aspects of language (Peters, 2014b, p. 584) and set out prescriptive guidelines for spelling, punctuation, the treatment of numbers, units and measurements, and references as well as specific rules for formatting as a way to ensure consistency within and across documents. As such, style guides are usually associated with the book publisher or organisation for which they have been developed and are used for in-

house purposes. However, some style guides, like the *Chicago Manual of Style*, can carry weight beyond their own house or organisation and are (often widely) used by organisations, people and publishers outside of the originating house (Peters, 2014b, p. 584). One important feature of style guides is that in their preliminary matter, they normally indicate a preferred dictionary which is to be used alongside the style guide in enforcing particular spelling conventions.

Usage guides are somewhat more difficult to classify, and whole books have been written on delineating what counts as a usage guide. The difficulty arises in the considerable overlap in content and form between usage guides and other language reference genres (Peters, 2017; Straaijer, 2017). It is not the aim of this thesis to explore the various issues associated with establishing a precise definition of the genre of usage guides, and therefore the category ‘usage guides’ in this study is taken to include sources that cover contentious issues of grammar and lexis (Peters, 2014b). They may adopt a prescriptive or descriptive approach, and reflect examples of usage created by the author or actual usage. They may be comprehensive, but more often than not they are not comprehensive and reflect a range of topics deemed important by their author. Such topics are usually framed as areas in which usage is problematic or incorrect. For example, *English, our English: Common errors in South African English and how to resolve them* (Linnegar, 2009) is an example of a usage guide that focuses on a select range of problems in SAfE identified by the author, and exemplifies these problems and how to correct them with examples provided by the author. *The Cambridge Guide to Australian English* (Peters, 2007), on the other hand, is an example of a descriptive usage guide that provides advice on contentious issues of grammar and lexis in AusE and which draws on actual examples of usage drawn from various corpora.

All texts whose function it is to instruct the reader on various aspects of writing (such as composition and structure) and editing (such as levels and types of editing) were categorised as editing and writing sources. The sources include handbooks on matters related to writing and editing processes and skills, as well as guides on writing and editing available via various institutions or websites. The category ‘language reference sources’ was created to include all sources that provide information on language and that do not fit neatly into the any one of the ‘dictionary’, ‘style guide’, ‘usage guide’ and ‘editing and writing sources’ categories. All of the sources in this category have in common that they deal with issues related to grammar and other aspects of language and therefore overlap, to a large degree, in content and function with the other categories, but do not quite fit into them. As a result, this category may be viewed as quite ‘fuzzy’, but it is useful for the discussion in this thesis because it allows for a more nuanced description and analysis of the kinds of sources that editors draw on. Examples of sources that were placed in this category are texts dealing with English grammar (such as popular grammar books and comprehensive English grammars), language reference websites and blogs (hosted either by well-known publishers of language reference works, such as *Oxford Reference Online*, or individuals, such as *Grammar Girl*), online corpora (such as COCA or the BNC), and operating-system grammar checkers.

The ‘various online and print sources’ category comprised an array of subject-specific online and print sources that do not deal specifically with issues of language but which are used by editors in making decisions about correct or appropriate language use in given contexts. Examples of these sources include organisations’ websites that are consulted to confirm the use of specific terms, and search engines which are used as tools to gather information about a particular aspect of language use or specialised topic.

In addition to categorising each source listed by the respondents into the different source type categories, the sources were also classified in terms of the particular norms to which they are oriented. In this regard, norm orientation is taken to mean the variety of English on which the source is based. To do this, each source was assessed in a step-wise fashion in order to ascertain its norm orientation. The first and most obvious step was to determine whether the norm orientation of the source could be gauged from its title. For example, *The Cambridge Guide to Australian English Usage* is very obviously oriented towards the AusE norm. If it was not possible to determine a source’s orientation from the title, then the preliminary matter, and in particular the foreword, was read in order to determine if the norm orientation of the source was expressed therein. If, after having read the title and preliminary matter, the norm orientation of the source was still not clear, then the source’s inclusion and treatment of “aging” and “ageing” was used as a litmus test for norm orientation. Although the two spelling variants help to distinguish between an AmE and BrE orientation respectively, at this stage of the classification process, this binary test was deemed appropriate because it was assumed that any source that was not oriented towards the AmE or BrE norm would state this very clearly either in the title or in the preliminary matter. This process resulted in the identification of 10 norm orientation categories, summarised in Table 4.1.

Orientation	Explanation
American English	Sources in which language entries are modelled on AmE norms.
American and British English	Sources in which both AmE and BrE norms serve as the model for language entries.
Australian English	Sources in which language entries are modelled on AusE norms.
Multinational	Sources in which multiple regions’ language norms are included in such a way that does not model entries around a particular variety.
British English	Sources in which language entries are modelled on BrE norms.
English and a foreign language	Bilingual sources in which language entries are modelled on English and another foreign language.
Indeterminate norm orientation	Sources for which the norm orientation could not be gauged.
South African English	Sources in which language entries are modelled on SAfE norms.
Afrikaans	Sources in which language entries are modelled on Afrikaans.
South African English and Afrikaans	Bilingual sources in which language entries are modelled on SAfE and Afrikaans.

Table 4.1: Categorisation of norm orientations

Once the respondents’ answers to the questions on norm-providing sources had been processed, the data were analysed quantitatively using descriptive statistics and presented in the form of percentages. Appendix C reflects the categorisation of all sources listed by the respondents as well as the number of respondents in each country who listed each source.

4.3 Corpus-based investigation

The second part of the study involved an investigation into how editors of English texts respond to genitive choices, as a feature associated with ongoing change, in the texts that they edit, and how this intervention differs across different registers and varieties of English. The second part of the study therefore sought to answer research question 2.3.

Because this study is interested in actual language use and the choices that editors make when editing texts (rather than relying on anecdotal or intuitive observations of what these choices might be), the decision was made to follow a corpus-based approach. Corpus-based approaches draw on large collections of natural texts that reflect patterns of real language use which are analysed by means of both quantitative and qualitative analytical techniques (Biber et al., 1998, p. 4). This section sets out a discussion of the corpus-based approach used to answer research question 2.3. In Section 4.3.1, an overview of the corpus design is presented. Thereafter, the composition and construction of the corpora are discussed in Section 4.3.2. Section 4.3.3 focuses on the extraction and annotation procedures employed for the dataset, while Section 4.3.4 explains the various methods applied in the analysis and interpretation of the data.

4.3.1 Corpus design

The most important consideration in the design of a corpus is what it intends to represent (Biber et al., 1998, p. 246). Since the purpose of the corpus in this study is to investigate how editorial practices in AusE and SAfE play a role in the alternation between two genitive constructions in different written printed texts, the most appropriate corpus design is one in which the factors of variety, register and editorial intervention can be investigated. To this end, a parallel, register-differentiated corpus of aligned unedited texts and their edited counterparts representing each of the two varieties, was selected as the most appropriate design for this study.

The main corpus is divided into two primary sub-corpora, with each representing one of the two main varieties under investigation, namely AusE and SAfE. As discussed in Section 3.3, five printed written registers are differentiated for each variety: academic, creative, instructional, popular and reportage. Each register is further sub-divided to include unedited texts and their edited counterparts, which are aligned. In other words, the design of the corpus used in this study is such that for each of the two varieties under investigation, aligned sets of unedited texts and their edited counterparts from the same five printed written registers are represented. This design allows for a parallel investigation into the occurrence of genitive constructions across the unedited texts and their edited counterparts for each register within each variety, but it also allows for a comparative investigation across the different varieties.

According to Biber (2015), a carefully designed corpus is one in which factors such as composition (as informed by sampling approach and diversity of the corpus) and size are considered in relation to what the corpus intends to represent. In the case of the corpus to be used in this study, the composition was one in

which the approach to sampling would need to ensure the representation of the diversity of English users in the two varieties and the diversity of texts that undergo editorial intervention. With due consideration for the need to construct a parallel corpus of printed written texts that reflects the range of editorial intervention in the two varieties of English, several decisions guided the design of the corpus.

In the first instance, the texts included in the corpus needed to reflect the printed written English of the range of English users in Australia and South Africa, but also the range of English editors in the two varieties. With regard to AusE, Collins (2014, p. 449) notes that the view of AusE as the dialect used by native-born non-Aboriginal Australians is narrow and does not take into account “Australia’s multiethnic society... [in which English is also used by] community groups of various non-English migrant backgrounds... [and] the English of Aboriginal communities”. SAfE is well known for its complexity and for the fact that it comprises a number of sub-varieties in which English is variously used as a first language and as a second language by language users of different ethnic groups (Bekker, 2012). As discussed in Section 3.2.2, the three sub-varieties investigated in this study are AfrE in which English is used as second language by first-language users of Afrikaans, BSAfE in which English is used as a second language by first-language users of one of the indigenous African languages, and WSAfE which represents the native variety of English spoken in South Africa. With due consideration for this, and given the usage-based view in which this study is grounded (as discussed in Chapter 2), the sampling approach to the representation of English in the Australian and South African corpus aimed at being inclusive of these different users of English. In other words, during the data-collection process, no texts were excluded if they had been written by non-native AusE and SAfE authors or edited by non-native AusE and SAfE editors. The only criterion for inclusion was that the English texts needed to have been produced and published in either Australia or South Africa.

Since this study is primarily interested in the influence of editorial intervention on the written English of the two varieties under investigation, the choice of register categories to be included was informed by the fact that the registers in the corpus needed to reflect the range of texts that undergo editorial intervention, which is known to differ across sectors and text types (see Section 3.3). In this regard, the categorisation of text types into different registers was based on the categorisation of printed written texts used in the International Corpus of English (ICE) corpora,¹⁴ of which five registers representing printed written texts broadly reflect the different contexts in which editors work (Nelson et al., 2002, pp. 7-8). These five registers are: academic, creative, instructional, popular and reportage. Texts included in the academic register are texts that are “written by academics for academics”, and which function as informational texts with a specialist, academic audience (Nelson et al., 2002, p. 7). Such texts include academic articles and books as well as dissertations and theses. Creative texts function to entertain a general audience and are

¹⁴ As discussed in Section 3.3, the ICE corpus is a large collection of corpora representing different national and regional varieties of English. Each ICE corpus consists of approximately 1 million words of spoken and written English produced after 1989.

used for leisure reading purposes. Creative texts are fictional in nature and include narratives such as short stories, novellas and novels. Instructional texts are informational texts that are instructive in nature, and include regulatory or administrative texts typically produced within governmental or corporate contexts as well as skills texts that offer instructions (Nelson et al., 2002, p. 8). In this study, instructional texts are also taken to include schoolbooks and pre-tertiary learning materials since these are informational texts that serve instructive purposes. Examples of instructional texts included in this register are schoolbooks, teacher's guides, instruction documents and manuals, administrative writing and 'hobby' texts. Texts in the popular register are intended for use by the general public, and include non-fiction writing, as well as popular books and magazine articles. Reportage refers to all forms of news writing, and may include news reports published by newspapers as well as newsletters published by various organisations.

The last aspect related to composition in the design of the corpus had to do with the timeframe in which the texts had been authored and edited. Since this study is interested in language change, a diachronic corpus would have been most desirable, but given the near impossibility of sourcing a substantial number of unedited texts and the edited counterparts from different time periods for each of the different varieties and five registers, it was not possible to construct a diachronic corpus for this study. Instead the decision was made to construct a synchronic corpus and focus on current usage and current editing practices (for which the sourcing of texts would be easier), and specifically how these editorial practices might help or hinder the processes of ongoing change. Since synchronic variation and diachronic change are related, and since language change proceeds through variation, as discussed in Section 2.2, the use of a synchronic corpus in this study was deemed acceptable (given the near impossibility of constructing a diachronic corpus). The decision was therefore made to construct a synchronic corpus that reflects current usage and editing practices in the varieties under investigation. To this end, only texts that were authored and published after 1990 were considered for inclusion, however, in nearly all cases, the texts included had been authored and edited after 2000.

The second important consideration in corpus design is related to the size of the corpus and the need for corpora to be large enough to reflect the distribution of the features under investigation (Biber, 2015). In regard to the size of the corpus this study, one consideration was crucial: since this study investigates editorial treatment of genitive alternation, and since editors do not change everything in the texts that they edit, the corpus needed to be sufficiently large to provide substantial number of cases to investigate the editorial treatment of the feature. The corpus was therefore designed to be as large as possible, including whole texts (and not samples of texts), and with the aim of including as many texts from as many different authors as possible.

This corpus design facilitates an investigation of genitive alternation within and across the varieties under investigation, as well as how this alternation plays out similarly or differently in different printed written registers. Furthermore, the inclusion of unedited texts and their edited counterparts allows for an analysis of the influence of editorial practices in the varieties and the five registers.

4.3.2 Corpus construction and size

As discussed in Section 4.3.1, the corpus used in this study consists of two main sub-corpora representing the varieties of English under investigation. The AusE sub-corpus was started from scratch, with construction commencing in 2017, while the SAfE corpus is the result of the expansion of an existing SAfE editing corpus. The original SAfE editing corpus, the construction of which was initiated in 2011 under the directorship of Haidee Kruger, was expanded in the course of 2016 and 2017 through the addition of approximately 1 million words per unedited and edited sub-corpus.

Because the corpus is designed to include edited texts as well as their unedited counterparts from a variety of registers, several strategies for the collection of texts were used. Unlike most corpora of written texts, the unedited dimension of the corpus used in this study presented specific and unique challenges, the most salient being that the unedited versions of printed written texts are not available for access in the public domain, and therefore such texts had to be provided by their publishers, authors or editors. The first step in the collection of texts was to contact various publishing organisations (such as publishers and news agencies), organisations representing editors, and language services agencies in the two countries with a request to contribute texts for inclusion in the corpus. In each case, an invitation to contribute textual material for inclusion in the study (or a request to distribute this invitation), was sent via email to the appropriate contact at each organisation. The initial points of contact for the collection of Australian data included the following organisations (in addition to various contacts within the researcher and her supervisors' own professional networks):

- Australasian Association of Writing Programs (AAWO)
- Australian Association of Language Companies (AALC)
- Australian Institute of Interpreters and Translators (AUSIT)
- Australian Publishers' Association (APA)
- Australian Society of Authors (ASA)
- Canberra Society of Editors
- Freelance Editors' Network
- Fairfax Media
- Institute of Professional Editors (IPEd) and its six branches (Editors NSW, Editors Queensland, Society of Editors (SA), Society of Editors (Tasmania), Editors Victoria and Editors WA)
- Magazine Networks
- Media, Entertainment and Arts Alliance (MEEA)
- News Corp Australia
- Schwartz Media
- Translators and Interpreters Australia

In addition to various contacts within the researcher and her supervisors' own professional networks, the initial points of contact for the collection of textual material for inclusion in the South African corpus included:

- Academic and Non-Fiction Authors' Association of South Africa (ANFASA)
- Professional Editors' Guild (PEG)
- Projourn
- Publishers' Association of South African (PASA)
- South African National Editors' Forum (SANEF)
- South African Science Journalists' Association (SASJA)
- South African Translators' Institute (SATI)
- Southern African Freelancers' Association (SAFREA)
- Various newspapers and magazines at Media24, Independent News & Media and Caxton & CTP Publishers and Printers

During this initial round of contact, several of the South African organisations responded to the request by either providing texts or, in the case of organisations, distributing the request to their members. However, the response rate from the Australian contacts was very low, and in the case of organisations who distributed the invitation to their members, many of these members indicated that due to confidentiality agreements with their authors they would be unable to supply texts. Furthermore, due to privacy agreements in place between these editors and their authors, many were unable to provide the contact information of their authors, but all advised that authors would be the best point of contact for the request to contribute textual material.

Given this, the decision was made to cast the net for data collection somewhat wider, and to send the invitation to contribute textual material for inclusion in the corpus to authors of English texts in Australia as well as to specific contacts at various publishers listed in the Australian Publishers' Association's membership database. With regard to Australian publishers, care was taken to ensure that as diverse a range of publishers, in terms of the types of texts published, were contacted. Since no centralised database of Australian authors exists, it was necessary to compile a list of authors. To do this, the author lists on various publishers' websites were consulted and all authors' names were copied into an Excel document. In addition, the client lists on a number of Australian literary agents' websites were also consulted, and the authors' names copied into the Excel document. Thereafter, the list was sorted to remove any repetition, resulting in a final list of 101 authors. Using Google's search engine, the contact information of each author was sought. In the case of some authors, an email address was located and the invitation sent via email; however, for many authors it was impossible to find a contactable email address. In all such cases, the authors' websites were consulted and the request to contribute material was submitted using the website's built-in online contact form.

Once the texts for the corpus had been collected they required some processing. All texts had been provided in electronic format (in Microsoft Word), either as one text in which editorial changes were marked as Track Changes, or as two separate texts (one unedited and one edited). For the texts that contained Track Changes, the unedited text was recovered by rejecting all changes in the document and then saving this. The edited version was recovered by accepting all changes, and then saving this version. Once the unedited and edited versions were recovered, each text was first cleaned up to remove any structural elements (such as tables or figures) or field codes (such as placeholders for captions) that might be corrupted during conversion to plain text format (txt). All such elements were marked up using a mark-up scheme similar to the content elements mark-up set out in the ICE Markup Manual for Written Material (Nelson, 2002). Once this initial round of processing was completed, the texts were converted to plain text format before being manually checked and edited to ensure accurate conversion to txt format. In addition, general mark-up was applied to various content elements of the text following the ICE Markup Manual for Written Material (Nelson, 2002), where relevant, and the metadata on each text's author and editor were added (see further discussion below). The texts were then aligned in two stages. All of the BSAfE had been previously aligned manually at sentence level in Excel and therefore did not require any further processing. All other texts were automatically aligned using an automatic alignment tool in *Sketch Engine*. An inspection of the aligned segments revealed that in some registers where editorial intervention is known to be substantive (such as in creative writing or instructional writing), the automatic alignment did not yield satisfactory alignment accuracy. As such, individual texts were inspected and where the accuracy of the automatic alignment was deemed unsatisfactory, these texts were aligned manually in Microsoft Excel at sentence level.

The cleaned-up, aligned texts were uploaded into Lexical Computing Limited's online corpus management and text analysis program, *Sketch Engine* (<https://www.sketchengine.eu/>), automatically tagged using the built-in POS tagger. Although *Sketch Engine* contains 500 ready-to-use corpora, it also allows licence holders to upload their own corpora and use the various tools available to analyse these corpora.

As discussed in Section 4.3.1, the representation of the range of users of English in each of the different varieties is an important consideration. During the collection of textual material, all contributors were asked to indicate if the author and editor were first-language users of English, and if not, what their first language was. Using this information, the corpora were constructed on the basis of the sub-varieties to which the authors belonged. In other words, five sub-varieties are represented: AusE for texts whose authors who are native speakers of AusE, AfrE for texts whose authors are AfrE users, BSAfE for texts whose authors are BSAfE users, WSAfE for texts whose authors are native user of SAfE, and Other (SA) for texts from South Africa whose authors are not AfrE, BSAfE or WSAfE users. This metadata was captured for each file. In addition, information about the editor's language background (for the edited texts) was also captured (see further discussion below).

Before moving on to a discussion of the composition and size of the final corpus, it is necessary to first discuss some of the challenges associated with the construction of the corpus, so as to provide some context for the composition and size of the final corpus. Because the corpus comprises unedited texts as well as their edited counterparts, the corpus exposes the unedited writing of authors and the editing practices of editors, which in 'normal' corpora would otherwise be less visible, placing both in a particularly vulnerable position. Although measures were taken to limit this exposure and to alleviate any concerns that would arise as a result of this exposure (see the discussion of these measures in Section 4.4, which deals with ethical considerations), many of the authors, editors and publishers who were contacted expressed their unease and/or simply declined to contribute texts.

In other cases, the nature of the text production and publishing process influenced the availability of texts. This was particularly the case for the reportage register, for two reasons. First, journalists and editorial staff at newspapers work under severe time constraints associated with the frequency with which newspapers are published (usually daily or weekly). This means that neither journalists nor editorial staff typically have time to dedicate to gathering the texts requested. This reason for the inability contribute texts was cited by most journalists and editorial staff contacted for the collection of textual material for inclusion in this corpus. Second, unlike at book publishers and language services agencies, the publishing and editing process at a newspaper is such that records of unedited manuscripts are not usually kept. Most newspapers make use of a centralised production and layout program, in which a skeleton framework for the layout of the designed pages of each issue is created. Journalists file their stories by uploading them into the relevant allocated slot on these designed pages of the system, from which editors can access the raw copy and then start editing *over* them. In this process, the raw unedited versions of the text are lost once the editor commences the editing process.

A last problem is related to the collection of textual material produced by Aboriginal English and BSAfE authors. With regard to texts produced by Aboriginal English authors, many of the authors, editors and publishers contacted expressed the view that investigations of the language use of such authors and editors are best carried out by researchers of Aboriginal heritage, and therefore declined to contribute textual material for inclusion in the study. This means that this very important and emerging group of authors and editors within the Australian publishing landscape (Freeman, 2010) are, unfortunately, not represented in the corpus in this study. While some texts produced by BSAfE authors were collected and are included in the South African corpus, no creative texts representing this sub-variety were obtained, despite several requests being sent to BSAfE authors and publishers of only BSAfE creative texts. Moreover, there are no texts from AfrE authors in the creative register. This is most likely due to the tendency in South Africa for Afrikaans authors of fiction to write in Afrikaans and then to have these texts translated into English.

Furthermore, contributors were asked to provide the name of the author and editor of each text. The collection of this information was solely to determine the number of unique authors and editors represented across the sub-varieties and registers. While some contributors were happy to provide this information,

others were either unwilling or unable to do this. This means that the total number of unique authors and editors per sub-variety is unknown. However, it is possible to determine the minimum number of unique authors and editors per register, per sub-variety, which is reflected in Table 4.2 and which shows that with the exception of AusE reportage, all sub-varieties and registers reflect the work of different authors and editors.

Because of the longer construction history of the SAfE corpus, the number of texts for especially the AfrE and BSAfE users is higher, compared to the AusE, WSAfE and Other (SAfE) authors, which skews the comparability of the sub-corpora (see Table 4.2). Only one text could be obtained for the Australian reportage register and none could be obtained for WSAfE authors in this register, despite several invitations to contribute textual material being sent to Australian and South African news agencies and newspapers.

	AusE			AfrE			BSAfE			WSAfE			Other (SAfE)		
	Text	± Auth	± Edit	Text	± Auth	± Edit	Text	± Auth	± Edit	Text	± Auth	± Edit	Text	± Auth	± Edit
Academic	15	15	5	74	55	8	25	22	5	12	10	2	29	27	3
Creative	8	8	8	0	-	-	0	-	-	6	6	4	0	-	-
Instructional	9	4	5	25	16	2	6	2	1	28	5	6	25	8	2
Popular	7	4	4	13	4	2	2	2	2	3	5	3	0	-	-
Reportage	1	1	1	29	19	3	98	11	20	0	-	-	5	-	-
Total	40	32	23	141	94	15	131	37	28	49	26	15	59	35	5

Table 4.2: Text and approximate unique author and editor count per register, per sub-variety

The number of texts within the South African sub-varieties is therefore much larger than the AusE corpus in terms of the number of texts. This is especially true for the AfrE and BSAfE sub-corpora (even if no creative texts are included for these two sub-varieties). The sub-corpora are therefore unbalanced across the registers and sub-varieties in terms of the number of texts, but also in terms of word counts (see Table 4.3). The challenges associated with the text collection process, specifically with regard to the difficulty with obtaining access to unedited, printed written texts, are the cause of this imbalance. This is the greatest limitation of the corpus used in the current study. Biber et al. (1998, p. 250) note that it is important not to underestimate the difficulties associated with constructing corpora, and advise that in the face of such difficulties and constraints, compromises often have to be made. Having said this, they note that ten texts of approximately 1 000 words each in each register should be sufficient for representing (variation among) many grammatical features (Biber et al., 1998, p. 249). Since genitives are relatively common features in writing (Rosenbach, 2017), and bearing in mind the difficulties associated with the collection of texts for the corpus in this study, the size of the corpus and the representation of the different registers and sub-varieties, while not being ideal, is viewed as acceptable.

	AusE		AfrE		BSAfE		WSAfE		Other (SA)	
	Unedited	Edited	Unedited	Edited	Unedited	Edited	Unedited	Edited	Unedited	Edited
Academic	36 994	38 223	926 328	914 753	302 589	305 091	45 604	46 536	253 522	263 260
Creative	315 859	333 158	0	0	0	0	405 994	365 912	0	0
Instructional	44 188	46 008	275 058	281 767	58 013	63 519	367 430	356 002	210 973	226 642
Popular	274 537	281 172	55 148	57 021	21 471	21 351	145 977	175 609	0	0
Reportage	4 339	4 481	14 158	14 032	31 020	27 961	0	0	18 421	18 251
Total	675 917	703 042	1 270 692	1 267 573	413 093	417 922	965 005	944 059	482 916	508 153

Table 4.3: Word count per register, per sub-variety

As pointed out above, the representation of the range of users of English in each of the two countries is an important consideration, and for this reason, the corpus is constructed along the lines of the authors' language status. Nonetheless, the current study is interested in the editorial practices of editors of English texts in the two varieties, and therefore information on the editors' language status was also collected.

The representation of the editors' English status in the two varieties is given in Table 4.4, which indicates the number of texts edited by editors belonging to different sub-varieties in each register. Nearly all editors represented in the Australian corpus are users of the native variety of AusE with the exception of one AmE editor in the creative register and one BrE editor in the popular register. All editors whose work is included in the Australian corpus are therefore first-language users of English. In the South African corpus, most editors in all registers are WSAfE users, with some AfrE editors represented in the academic, instructional and reportage registers. No BSAfE editors are represented in the corpus. As discussed in Section 3.2.2, South Africa represents a complex sociolinguistic constellation in which multiple sub-varieties are developing in contact with each other at different rates in the same context. Since the aim of this study is to investigate how editors as linguistic gatekeepers help to move a variety towards endonormativity, and since one aspect of endonormativity is convergence between the STL and the different IDG strands in the language contact setting, the representation of mostly WSAfE editors editing the work of AfrE and BSAfE authors opens up the opportunity to also determine if editorial choices reflect a converging tendency among the varieties or not. For example, if WSAfE editors of BSAfE or AfrE authors' texts introduce changes that move either of the indigenous varieties further away from WSAfE, then there is evidence that editors are preventing convergence between the varieties. On the other hand, if the choices of WSAfE editors bring the linguistic patterns of two (or all) of the varieties closer together, then there is evidence to show that editors are facilitating the process of convergence.

	Australian texts			South African texts			
	AusE	BrE	AmE	AfrE	WSAfE	BrE	Unknown
Academic	15	0	0	18	59	0	63
Creative	7	0	1	0	5	1	0
Instructional	9	0	0	29	39	0	16
Popular	6	1	0	0	14	0	4
Reportage	1	0	0	12	119	0	1
Total	38	1	1	59	236	1	84

Table 4.4: Representation of editors across registers

Despite the somewhat skewed representation of authors in the South African corpus, the representation of editors is (broadly) in line with the language profiles of editors of English texts in South Africa (see, for example, Law and Kruger (2008)) and Australia, and broadly in line with the data collected using the questionnaire in this study (see Section 5.2.2). With regard to the South African corpus, WSAfE editors are represented across all five registers, with WSAfE users making up the majority of editors in most registers. The only exception to this is the academic register, where the language profile of the editors is unknown for 63 texts. AfrE editors are represented in the academic, instructional and reportage registers, with one BrE editor in the creative register. The language status of some editors is unknown in the academic, instructional and to a much lesser extent, the popular and reportage registers. This is due to the fact that in many cases contributors of textual material did not have a record of who the editor for every text was (this was particularly the case for language services agencies). Since most language services agencies in South Africa outsource their work to a local corps of freelance editors, it is very likely that these texts were edited by local editors, the majority of whom are in all probability either WSAfE or AfrE users.

4.3.3 Data extraction and annotation

Using the concordance tool available in the *Sketch Engine* platform, all instances of 's and s' and of were extracted for each register and in each variety. The South African corpus used in this study exceeds 6 million words. Given the fact that genitive constructions are frequent linguistic features (Rosenbach, 2017), and a search of the South African corpus yielded 110 663 instances of 's and s' and 114 560 instances of of, it was deemed unmanageable to manually sort through every instance of 's, s' and of in the entire corpus. To circumvent this a two-pronged approach to the extraction of the data was used based on the fact that the editors would have done one of three things when encountering a genitive construction in the unedited texts: (1) the editor might retain the genitive construction as is, (2) the editor might replace the genitive construction with another appropriate genitive construction (as demonstrated in (4.1) below), or (3) the editor might remove the genitive construction from the edited text completely (as demonstrated in (4.2) below). However, editors might also add a genitive to the edited text where one does not exist in the unedited, and therefore a fourth type of intervention, added, is included (as demonstrated in (4.3) below).

(4.1a) ...controlled **the management of the company** until his death... [SA_A-001-O]

(4.1b) ...controlled **the company's management** until his death... [SA_A-001-E]

(4.2a) ...and Emma sitting in a tree, k-i-s-s-i-n-g..." **Emma's smile faded**. She hoped Gay hadn't seen anything. [SA_C-006-O]

(4.2b) ...and Emma sitting in a tree, k-i-s-s-i-n-g..." Emma hoped Gay hadn't picked up on anything between her and Craig. [SA_C-006-E]

(4.3a) ----- [Aus_C-001-O]

(4.3b) The killings, in themselves a retaliation for **the murders of six Catholic civilians** on the previous day. [Aus_C-001-E]

In the first step of extracting the data, the decision was made to extract a random sample of 1 000 instances of each variant (*'s* and *s'*, and *of*) from each of the unedited registers in each of the varieties (together with its aligned segment from the edited registers) for each of the two regional sub-corpora (Australia and South Africa). Where an unedited register did not contain more than 1 000 instances of *'s* and *s'* or *of*, the entire set of instances was extracted. This process yielded a sample of 6 586 tokens from the Australian sub-corpus and 9 230 tokens from the South African sub-corpus.

The second step was to extract a sample of instances of each variable (*'s* and *s'*, and *of*) that occurred in the edited sub-corpus (together with its aligned segment from the unedited sub-corpus) but which did not occur in the unedited sub-corpus. In the Australian and South African sub-corpora together, the number of cases that met these criteria did not exceed 1 000, and therefore all such occurrences were extracted and added to the sample of tokens. This process yielded a further 1 301 cases from the Australian corpus and 2 740 cases from the South African corpus, resulting in a total sample of 7 887 cases for AusE and 11 970 cases for SAfE included in the analysis.

Once the sample had been extracted, each of the cases was manually inspected to determine if it constituted a choice context. Rosenbach (2014, p. 244) notes that one of the greatest issues regarding studies of genitive alternation is the lack of comparability of the results of these studies due to the use of differing criteria for the inclusion or exclusion of genitive constructions as well as the tendency for such studies to make use of differing definitions for the factors that are analysed. In response to this, Bresnan et al. (2017) have made available an annotated dataset of alternating genitives in four spoken varieties of English as well as the inclusion criteria and annotation schema used in the annotation of the data. In line with the call by Rosenbach (2014) to ensure consistency in the inclusion criteria used and the annotation of data, the inclusion criteria and annotation schemata used by Bresnan et al. (2017) were adopted in this study.

The following criteria, as set out in Bresnan et al. (2017, pp. 2-3), were used to identify interchangeable *s*-genitives to be included in the dataset. As a first step, only those tokens that consisted of two full noun phrases separated by *s'* or *'s*, were included. This meant that the following types of observations were excluded: non-genitives containing the clitic *'s*; *s*-genitives that do not contain a possessum; double genitives; and *s*-genitives with pronominal possessors. Subsequent to this, the following criteria were applied:

1. *s*-genitives that were ambiguous between a classifying or a determiner function, and which could not be disambiguated as classifying genitives from the surrounding context were included.
2. Measure genitives indicating temporal length and value were included.
3. Fixed expressions and unambiguous classifying genitives were excluded.

Similarly, and following Bresnan et al. (2017, pp. 3–5), only those observations that comprised two full noun phrases separated by *of* were included as *of*-genitives. Thereafter, the following criteria were applied:

1. All remaining genitives that contained a possessum starting with a determiner other than *the* were excluded. However, *of*-genitives that contained a possessum that lacked a definite determiner but which had a unique and unambiguous referent (as informed by either world knowledge or the context in which the genitive occurred) were included.
2. The following were all excluded:
 - a. Appositive genitives, where the *of*-phrase is a post-modification whose head is co-referential with the head of the preceding noun phrase.
 - b. Classifying genitives, in which the *of*-genitive is a modifier qualifying the possessum, rather than a complement of it.
 - c. Double genitives.
 - d. Fixed expressions.
 - e. Completely fixed (non-alternating) idiomatic expressions (as assessed by the author, who is a native user of English).
 - f. Partitive *of*-genitives that do not express part–whole relations and that cannot be expressed in the *s*-genitive. For example, *one of my dogs* cannot be expressed as *my dog's one*.
 - g. Date-like constructions, where the possessum is an ordinal number representing a day or the possessor is a span of time.

In total, 2 948 interchangeable tokens were identified in the Australian text sample (*s*-genitive, $n = 1\ 203$; *of*-genitive, $n = 1\ 745$) and 6 182 interchangeable genitive tokens were identified in the South African sample (*s*-genitive, $n = 3\ 470$; *of*-genitive, $n = 2\ 712$).

Following this, each token was manually coded for a set of variables, based on the existing literature on the genitive alternation (see Section 3.4.3.1). Six extra-textual variables were coded. The first three are VARIETY (*AusE* or *SAfE*), AUTHOR LANGUAGE as an indication of sub-variety (*AusE*, *AfrE*, *BSAfE*, *WSAfE*, and *Other SA* for authors of South African texts that are not *AfrE*, *BSAfE* or *WSAfE* users, or whose English-speaker status is unknown) and EDITOR LANGUAGE (*AusE*, *BrE*, *AfrE*, *WSAfE* and *Unknown* for editors whose English-speaker status is unknown). The languages for each token's author(s) and editor were coded using the codes as set out in Table 4.5.

Code	Category
AusE	Australian English
AfrE	Afrikaans L2 English
BSAfE	Black South African English
WSAfE	White South African English
Languages coded as Other : Arabic, Asian, Dutch, German, Mixed provenance, Other African and Unknown	

Table 4.5: Language codes used for authors' and editors' languages

In addition, AUTHOR GENDER and EDITOR GENDER were annotated using five codes. *F* reflects females, *M* males and *Unknown* was used if the gender of the author or editor was unknown. It was sometimes the case that a text had been authored by multiple people of different genders. Where this was the case, the authors' genders were indicated using either *F + M* in cases where the first author was female and *M + F* in cases where the first author was male. Lastly, each observation was also coded to indicate the REGISTER in which it occurred using the following codes: *Aca* for the academic register, *Cr* for the creative register, *Ins* for the instructional register, *Pop* for the popular register and *Rep* for the reportage register.

Each interchangeable observation was then further coded to determine how it had been treated by the editor. This was determined by inspecting the aligned segment of the relevant observation, and then coding the observation using one of the four editorial intervention labels in Table 4.6.

Code	Category
Added	The editor has added the observation (i.e. it does not occur in the unedited text and has been added to the edited text). For example: Unedited: ... <i>made his way across the road towards <name deleted></i> . [Aus_C-002-O] Edited: ... <i>made his way past a young man slouched behind the wheel of a dusty Triumph Gloria and across the road towards <name deleted></i> . [Aus_C-002-E]
Removed	The editor removes the observation from the unedited text so that it no longer appears in the edited text. For example: Unedited: ... <i>and Emma sitting in a tree, k-i-s-s-i-n-g... "Emma's smile faded. She hoped Gay hadn't seen anything.</i> [SA-C-006-O] Edited: ... <i>and Emma sitting in a tree, k-i-s-s-i-n-g ... "Emma hoped Gay hadn't picked up on anything between her and Craig.</i> [SA-C-006-E]
Replaced	The editor replaces the observation in the unedited text with the alternative genitive construction in the edited text. For example: Unedited: ... <i>the family of her friend did not agree with the South African...</i> [SA_A-013-O] Edited: ... <i>and her friend's family did not agree with its policies.</i> [SA-A-013-E]
Retained	The editor retains the observation (includes minor changes to possessor or possessum, but does not result in change of the response variable type). Unedited: ... <i>while the Asian elephant's trunk might have 150 000 muscle fascicles...</i> [Aus_P-001-O] Edited: ... <i>while the Asian elephant's trunk might have 150,000 muscle fascicles...</i> [Aus_P-001-E]

Table 4.6: Editorial intervention labels used in the data annotation

The linguistic factors coded for each observation include general factors that are known to govern syntactic variation as well as factors that are more idiosyncratic to genitive variation (Rosenbach, 2014, pp. 225-226). Based on the survey of linguistic factors usually investigated in studies of genitive alternation (see Section 3.4.3.1) the observations in the data set were coded for six linguistic factors: semantic relation between the possessor and possessum, animacy of the possessor, definiteness of the possessor, syntactic weight of the possessor and the syntactic weight of the possessum (expressed as both number of characters and number of words), and the presence of a final sibilant in the possessor.

The SEMANTIC RELATION between the possessor and possessum of each observation was coded using one of five values set out by Bresnan et al. (2017, p. 11), the classification of which was based on the motivation in Rosenbach (2002). For each observation, and as advised by Bresnan et al. (2017, p. 11), the coding of the semantic relation was based on the head noun of the possessum. These five values are:

1. *K* for kinship terms which typically include all types of family relations and colloquialisms that refer to these relations. For example, *Ella's father*.
2. *B* for body parts, including the entire body and only including metaphorical uses in which the physical sense dominates. For example, *a person's heart*.
3. *O* for relations that express legal ownership of anything concrete and non-concrete that can be possessed, but excludes anything that is used regularly by someone but not owned by them. For example, *the farmer's own non-disposable equipment*.
4. *P* for part/whole relations that apply to concrete physical inanimate possessor. For example, *the page of my book*.
5. *N* for non-prototypical relations, which includes anything not falling into the other four categories such as derived nominals and part/whole relations of non-concrete physical objects and entities.

Following the annotation schemata set out by Bresnan et al. (2017, p. 6), the ANIMACY of the possessor was coded as one of the following five values:

1. *A* for animates including all nouns that refer to humans (dead or alive) and higher animals (such as *prisoner, children, student, father, Kandinsky, president, elephant*).
2. *C* for collectives including nouns that refer to organisations as well as nouns not referring to organisations but with potentially variable verbal and pronominal concord (such as *society, OKIMO, National Party*).
3. *T* for temporal nouns or adverbs with a time reference (such as *the afternoon, the twentieth century, today*).
4. *L* for locative nouns referring to places (such as *the Missouri River, Africa, Sparta*), including nouns such as 'state' or 'water'.
5. *I* for other inanimate concrete and abstract nouns, as well as gerunds/infinitives (such as *HbS, pain, vessels, test, multiculturalism*).

The DEFINITENESS of the possessor was coded using three values: *Definite* for any possessor with a definite determiner or a definite pronoun; *Indefinite* for any possessor with an indefinite determiner or no determiner; and *Definite pn* for all possessors that are proper nouns and names.

As in Bresnan et al. (2017, p. 8) both the POSSESSOR LENGTH and POSSESSUM LENGTH were coded. However, while Bresnan et al. (2017) base this on the number of orthographic words as separated by whitespaces and hyphens, in the current study the length of the possessor and possessum were coded using

the number of orthographic words (as separated by whitespaces and hyphens) and the number of characters excluding whitespaces.

The presence or absence of a FINAL SIBILANT in the possessor was also coded according to the codes set out Bresnan et al. (2017, p. 12). If the possessor ended in an /s/, /z/, /ʃ/, /ʒ/, /tʃ/ or /dʒ/, then it was coded as *Y* to indicate the presence of a final sibilant, otherwise, the observation was coded as *N* to indicate the absence of a final sibilant.

4.3.4 Data analysis and interpretation

The analysis of the corpus data was carried out in three stages. First, an overall analysis of the proportions and the normalised frequencies of the two constructions was done for the written published sub-corpus across the sub-varieties and registers. Second, the effects of editorial intervention on the overall proportions and normalised frequencies of the two constructions across registers and sub-varieties was investigated. Lastly, the interaction between the language-external variables (REGISTER, AUTHOR LANGUAGE¹⁵, AUTHOR GENDER, EDITED STATUS, EDITOR LANGUAGE and EDITOR GENDER) and the linguistic variables selected for inclusion in this study (SEMANTIC RELATION, ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, DEFINITENESS and FINAL SIBILANCE) in conditioning the choice of construction was investigated.

In the first stage, the overall results for the total estimated proportions and normalised frequencies of the *s*- and *of*-genitives in the written published (and therefore edited) texts are presented, first as overall proportional preferences per sub-variety and then as proportional preferences across sub-varieties and registers. Since not much is known about genitive alternation in AusE and (the different sub-varieties of) SAfE, the analysis of the overall proportions of the two constructions across the different sub-varieties investigated offers a good opportunity for a first corpus-based investigation into genitive alternation in these varieties. To understand how each of the varieties in this study ranks in relation to other varieties of English, the overall proportions are presented alongside other similar data for AmE and BrE in order to compare how the sub-varieties investigated in this study rank in terms of their openness to the incoming *s*-genitive with other varieties of English for which this data are available. Following this, the differences in the estimated normalised frequencies of the two constructions across the four sub-varieties and five registers are investigated (see Appendix D for an exposition on the calculation of the total estimated proportions and normalised frequencies).

Following this, a statistical evaluation of the differences across the sub-varieties and registers is carried out using conditional inference trees, as implemented in R (R Core Team, 2018) using the R-package ‘partykit’

¹⁵ In some analyses, AUTHOR LANGUAGE is coded as SUB-VARIETY to facilitate the analysis. Since the language of the author is a reflection of the sub-variety of the text, the codes are treated interchangeably.

(Hothorn & Zeileis, 2015). Conditional inference trees are non-parametric tree-structure models of regression and classification that make use of binary recursive partitioning (Levshina, 2015, p. 291). Levshina (2015, p. 291) explains that the process of binary recursive partitioning in the modelling of conditional inference trees follows three steps:

1. First, all independent variables are tested to see if they are associated with the response variable and the variable with the strongest association is selected.
2. A binary split is then carried out on this variable by dividing the dataset into two subsets.
3. Steps 1 and 2 are then repeated for each subset, until no further variables can be associated with the outcome at the level of statistical significance.

The visualisation of this process is presented as a tree structure in which the binary splits represent branches and which facilitates the interpretation of the data. Conditional inference trees offer a particularly useful method to investigate how multiple predictors interact in conditioning the selection of the response variable. Furthermore, they do not require distributional assumptions to be met and are considered to be robust in the presence of outliers (Levshina, 2015, p. 292). For this part of the analysis, the dependent variable is the proportional frequency of the *s*-genitive in the edited sub-corpus, while REGISTER and AUTHOR LANGUAGE (i.e. sub-variety) are set as the independent variables.

In the second stage of the analysis, the focus shifts to the effect of editorial intervention on the alternation between *s*- and *of*-genitives in the sub-varieties and across the five registers. The analysis of the effect of editorial intervention is presented in three steps: First, differences in the total estimated proportions of the two constructions are presented and analysed across sub-varieties, registers and unedited and edited texts. A statistical evaluation of the influence of REGISTER, AUTHOR LANGUAGE (i.e. sub-variety) and EDITED STATUS on the total estimated proportional frequencies of the constructions is carried out using conditional inference trees, as implemented in R (R Core Team, 2018) using the R-package ‘partykit’ (Hothorn & Zeileis, 2015). For this, the dependent variable is the proportional frequency of the *s*-genitive, with the three predictors the factors listed above (see Appendix D for an exposition on the calculation of these proportions and normalised frequencies). Second, the estimated normalised frequencies of the two constructions are presented and analysed across sub-varieties, registers and unedited and edited texts.

Since an overall view of proportions and normalised frequencies might mask different kinds of changes made by the editors, the third step in the analysis draws on a derived calculation in which the percentage of change across each register and sub-variety is calculated with reference to the percentage of *s*- and *of*-genitives that were added, removed, replaced or retained in each register in each sub-variety.

As is now well known from the substantive body of research on genitive alternation (see the discussion in Section 3.4), register and (sub-)variety are not the only factors known to influence the choice of genitive construction (Bresnan et al., 2017; Feist, 2012; Grafmiller, 2014; Heller, 2018a; Heller et al., 2017a; Heller

et al., 2017b; Hinrichs & Szmrecsanyi, 2007; Hundt & Szmrecsanyi, 2012; Jankowski, 2013; Jankowski & Tagliamonte, 2014; Leech et al., 2009; Rosenbach, 2002, 2005, 2008, 2014, 2017; Szmrecsanyi, 2010; Szmrecsanyi et al., 2016a; Szmrecsanyi et al., 2016b; Szmrecsanyi & Hinrichs, 2008; Wolk et al., 2013). Instead, register and variety have been shown to interact with various linguistic factors to influence the likelihood of one genitive construction being used over another. To investigate this interaction in the different varieties represented in this corpus, during the third phase of the analysis, the language-external factors of REGISTER and SUB-VARIETY, as well as EDITED STATUS, and the linguistic factors (POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, POSSESSOR DEFINITENESS, FINAL SIBILANCE and SEMANTIC RELATION) are modelled using random forests and conditional inference trees to explore how the different factors interact to condition the differential use of the two genitive constructions.

Random forests are non-parametric regression and classification models that are used to determine the importance of several individual independent variables (or predictor variables) in conditioning the outcome of the response variable (Levshina, 2015, p. 292). They provide an indication of the impact of each predictor variable on the response variable, given all other independent variables. The impact of each independent variable is reflected in its importance score, which can be visualised on a dotchart. The interpretation of the importance scores of the different independent variables is straight-forward: the higher the importance score, the stronger the impact of the variable on the response variable and the higher it ranks on the dotchart. In other words, random forests provide predictions about the importance of the different independent variables in conditioning the response variable and can be used to weed out variables that have no significant predictive power. Like conditional inference trees they do not require distributional assumptions to be met and are considered to be robust in the presence of outliers (Levshina, 2015, p. 292). As such, they can be used to identify the most important factors in conditioning the choice between the variants represented in the response variable – which in the case of this study, is the choice between the *s*- and the *of*-genitive construction.

The analysis of the interaction between the linguistic and language-external factors is carried out in three stages. During the first stage, the focus is on determining differences in strength of the various independent variables in conditioning genitive alternation across unedited and edited writing. In this part of the analysis, random forests are used to determine which of the various linguistic and language-external factors are important in conditioning genitive alternation across the unedited and edited tokens separately. This analysis will help to determine how editorial intervention shifts the interaction of the different factors in conditioning the choice of the two constructions. The analysis then zooms in on differences across unedited and edited tokens in each of the four sub-varieties. For this, one random forest per sub-variety is modelled for unedited and edited writing. This will be done in order to identify which factors are important predictors across the sub-varieties and registers, but also to probe how the strength of the various linguistic and language-external factors differs across sub-varieties, and within each sub-variety across unedited and

edited writing. This will lay the stage for the final analysis, which investigates how the most important linguistic and language-external factors in each variety interact to condition genitive alternation. This last analysis draws on conditional inference trees.

4.4 Ethics application and considerations

The nature of the data collected for this study necessitated the participation of editors in the survey and editors, authors and/or publishing institutions for the collection of textual material. Since the study involved human participants, ethics clearance prior to the collection of data was sought from the Ethics Committee for Language Matters at the North-West University, Vanderbijlpark, South Africa, and the Human Research Ethics Committee at Macquarie University, Sydney, Australia.

Several ethics considerations such as ensuring the anonymity of participants and contributors, ensuring that participants were fully informed of the nature of the study and assuring the participants that their participation was voluntary, informed the data-collection process. In the case of both the survey and the collection of textual material, all participants received a Participant Information and Consent Form (PICF) which outlined the purpose of the study, the expectations of their participation as well as the processing and reporting of the data. The PICFs used in the study are included in Appendix E. The participants were therefore fully informed of the nature of the study.

With regard to the collection of questionnaire data, participants were able to access the questionnaire on the secure online survey administration platform *Qualtrics*. One of the benefits of using *Qualtrics* is that it allows questionnaire respondents to complete and submit questionnaires anonymously, since the platform does not store any information on the respondent. The anonymity of the respondents was further assured by the fact that none of the questions in the questionnaire gather any identifying information. Furthermore, all data from the questionnaires are reported in the form of aggregate summaries, mitigating the risk of an individual being identified. Prior to completing the questionnaire, each respondent was asked to read through the PICF and then provide their consent to participate in the study.

The corpus design in this study is unusual, because it includes both unedited and edited versions of published texts. This unusual design raised potential concerns not typically at issue in corpus compilation. Such issues included the potential discomfort of authors by the visibility of their unedited manuscripts and of editors by the visibility of their editorial work and choices. Furthermore, due to implicit or explicit agreements of confidentiality between editors, authors and publishers, care had to be taken to ensure that such agreements were not violated. These concerns were mitigated by several steps. In the first instance, permission was sought from editors, authors and publishers for the use of texts, as far as was possible. All texts included in the corpus were anonymised during the process of preparing the texts. The anonymisation process included removing all identifying information, including author's, editor's and publishing institution's names; dates; contact information; addresses; etc. Furthermore, all contributors were asked if

extracts of the texts could be cited for exemplification purposes in the study. Where permission to do so was not granted, a note was made for the text and the text was not cited in the study. In these cases, texts were only used for quantitative analysis. Where permission was granted to cite texts, care was taken to ensure that the citation is not traceable using online sources (such as search engines).

Ethics clearance was obtained from the Research Ethics Committee for Language Matters at the North-West University, Vanderbijlpark, South African (ethics clearance reference number: NWU-00558-16-A8) and the Human Research Ethics Committee at Macquarie University, Sydney, Australia (ethics clearance reference number: 5201600843D).

4.5 Conclusion

This chapter set out a discussion of the research methodology used in the collection, analysis and interpretation of the questionnaire and corpus data for the empirical component of this study. The first section focused on the procedure for the design of the questionnaire used in this study. The questionnaire forms an important component of the data collection for this study because it sought to gather information on the language profiles and the kinds of norm-providing sources used by editors in the two varieties under investigation. Section 4.2 discussed the design of the questionnaire as well as the population and sampling procedure followed in the selection of the sample for this study. In addition, matters such as the distribution of the questionnaire and subsequent collection of responses were discussed. The processing of the data from the questionnaire received special attention in this section, and the several stages of this process were set out. Thereafter, a description of the quantitative approach used in the analysis of the questionnaire data was provided. The design and the construction of the corpus used in this study were discussed in Section 4.3. Specifically, this section focused on describing the compilation of the corpus and its usefulness in investigating the kinds of changes that editors make to the texts that they edit. In addition to the design and construction of the corpus, the data extraction and annotation procedures followed were also outlined. Lastly, the various descriptive and statistical analysis procedures employed in the analysis of the corpus data were explained. The next chapter presents a discussion and analysis of the questionnaire and corpus data.

Chapter 5

Presentation and analysis of the empirical data: Evidence for the role of editorial intervention in ongoing language change in AusE and SAfE and across written published registers

5.1 Introduction

This chapter presents the results of the empirical investigation of this thesis across two main sections. Section 5.2 reports on the findings for the questionnaire that was used to gather information on the biographical, work and language profiles of editors in Australia and South Africa, as well as information related to the kinds of norm-providing sources that they use and what the orientation of these sources is. The overall findings for the questionnaire are synthesised in the last sub-section of Section 5.2. The focus then moves to the presentation of the findings for the corpus-based investigation. The analysis is carried out over several sections. Section 5.3.1 sets out a distributional analysis of the proportions and normalised frequencies of the two constructions in the edited (and therefore published) written texts for each register in each of the sub-varieties. The analysis in Section 5.3.1 lays the foundation for a contrastive analysis of variation between the two constructions in the two varieties and across the different registers. Section 5.3.2 adds the role of editorial intervention as an additional factor in influencing genitive alternation. Thereafter, in Section 5.3.3, the interaction of the linguistic and language-external factors is evaluated statistically across the unedited and edited texts in the sub-varieties in order to explore if the various factors influencing the alternation have the same or different effects across the sub-varieties, and if editorial intervention shifts the effect of these factors, or not. The chapter concludes with a summary and synthesis of the findings from the questionnaire data and corpus-based investigation.

5.2 Presentation and analysis of questionnaire data

Research question 2.1 and 2.2 of this study seek to investigate the sociolinguistic background of and to determine what overt norm-providing instruments editors of English texts in Australia and South Africa use. Combined, these data will provide the first perspective on the interplay between overt and covert norm orientation of editors in these two countries. According to the Dynamic Model (Schneider, 2007), the codification of a variety of English through the compilation and publication of regional dictionaries, style manuals, usage guides and grammars serves as one marker of the variety's endonormative stabilisation and contributes to its shift from exonormativity towards endonormativity. These codification instruments, typically published by publishing houses and institutions vested with authority over the language (variety), serve as a kind of standardisation of the variety, reflecting distinctive aspects of the standard written form

of the variety. It is well known that these instruments are used by editors to inform their intervention in the written forms of language (Cameron, 2012; Peters, 2014b); however, the question that arises is whether or not editors in the different varieties draw on the norm-providing sources available for the variety of English they edit, and if not, what other sources they consult and why this may be. The collection of information regarding the sociolinguistic backgrounds of editors of English in Australia and South Africa provides information on the language profiles of these editors. As discussed in 2.6, editors' covert norms are their entrenched knowledge of language. This knowledge arises from their own linguistic backgrounds, but is also built up through their exposure to language use. Thus, the collection of information on the linguistic profiles of editors, and the interpretation of this with reference to the sociolinguistic circumstances of the variety within which they carry out their work and the overt norm-providing sources they use, opens up the possibility to understand how overt and covert norms interact to influence editors' choices.

Although the questionnaire was split into two main sections, the analysis of the data gathered using the questionnaire is presented over three sections in this chapter. First, the demographic data on the respondents' age, gender, work experience and context, qualifications and conceptualisation of the editing role are discussed in Section 5.2.1. Thereafter, Section 5.2.2 presents the data on the respondents' language profiles, and discusses the languages the respondents listed as their first or home language, their five strongest languages as well as the language(s) in which they do editing work. Section 5.2.3 then shifts the focus to the norm-providing sources that the respondents use by firstly providing an overview of the responses to the questions on sources and secondly analysing the particular kinds of sources, and the norm orientation of these sources, used by editors in the two varieties under investigation.

5.2.1 Demographic profile of respondents

The first section of the questionnaire comprised questions that sought to contextualise and compare the demographic, work and linguistic profiles of editors in the two countries. On the one hand, the collection of this data allows the researcher to gauge whether the respondents in each of the two samples represent the target population of editors in each country. On the other, the collection of this data also allows for an investigation into the social profiles of the editors in each of the two varieties, providing information on their exposure to linguistic usage events (which shapes their own grammars and gives rise to their covert norms) and the implications of this for the editors' approaches to their linguistic gatekeeping role.

As mentioned in Chapter 4, the link to the questionnaire was distributed in a range of forums to ensure as wide a reach as possible. Since it is likely that there is some overlap between the different forums it is not possible to determine the exact number of members of the target population that received the questionnaire. A total of 136 completed responses were returned. Of these, 78 (57%) were by Australian editors and 58 (43%) were by South African editors. Since there is a slight difference in the number of responses from the two countries (with 14% more respondents from Australia than South Africa), the respondents' answers to the demographic questions are presented and analysed as percentages per country to remove any skewing

effect related to reporting the raw numbers for each. In the discussion that follows in this section (as well as in Section 5.2.2), the data from each country for each question are presented alongside each other and are discussed individually before being compared.

The respondents were first asked to indicate their age by means of a fixed-alternative response, which included five options: 18–29 years, 30–39 years, 40–49 years, 50–59 years and 60 years or older. Figure 5.1 reflects the percentage of respondents per age range, per country. In both the Australian and the South African samples, the majority of the respondents are 50 years or older, with the number of respondents in each age group decreasing as the age range decreases. The distribution across age groups in both samples is similar to that reported by each country’s editors’ organisation. The 2016 member survey of the Institute of Professional Editors (IPEd) found that “43% [of Australian editors] were over 55 years old and only 11% were under 35” (IPEd, 2016, p. 2), while the 2018 rates survey of the Professional Editors’ Guild found that the majority (more than 55%) of South African editors are 55 years old or older (PEG, 2018, p. 2). In both countries, editors are therefore older language users. As a social factor in language use, age has been shown to correlate with language users’ adaptability to changing usage and preferences for standard or prestige forms for the language (Labov, 2001; Tagliamonte, 2012). In this regard, Baxter and Croft (2016, p. 16) explain that after adolescence, a language user’s flexibility in linguistic behaviour drops, so much so that the linguistic behaviour of adults does not change by more than 10% after adolescence. As a result of this, the linguistic behaviour of an older language user is said to reflect their linguistic behaviour at the time of early adulthood (and by implication the linguistic behaviour of the speech community at the time of the language user’s early adulthood) (Baxter & Croft, 2016, p. 16). This inflexibility to take up or adapt to changes in language means that older language users, such as those in the samples of editors, are generally considered to be much more conservative in their linguistic behaviour.

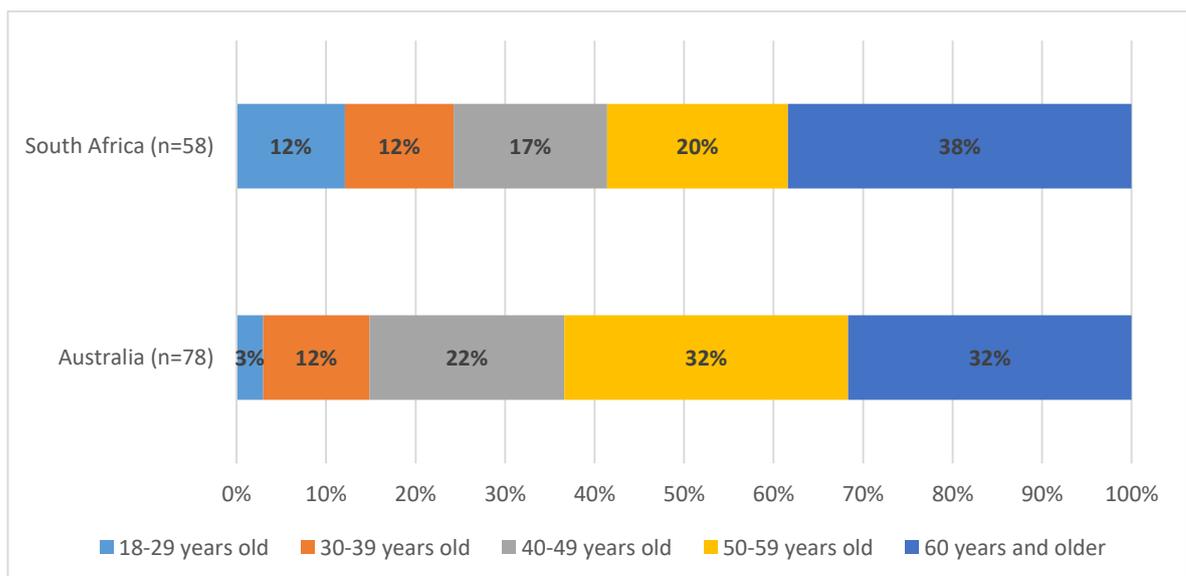


Figure 5.1: Respondents’ age by country

For the second question, the respondents were asked to indicate the gender to which they most identify. Three fixed-alternative options ('Prefer not to answer', 'Female', 'Male') and one 'Other' option in which respondents could specify a gender not covered in the selection options were provided. Two options were selected ('Female' and 'Male') and one was provided ('Mostly female'). The distributions for gender per country are presented in Table 5.1. The overwhelming majority of respondents in both samples are female. As with age, variationist sociolinguistic investigations into differences in language use across genders have shown a correlation between gender and a preference for standard or prestige forms (Labov, 2001; Tagliamonte, 2012). In fact, gender is considered one of the most telling social factors in linguistic behaviour, and it is widely accepted that females have a much stronger preference for standard language forms than males. However, a paradox arises in respect of gender. While females are more inclined to make use of standard or prestigious forms than their male counterparts, females have also been shown to take up creative and newer forms in language much more quickly (Labov, 2001, p. 293). Since the majority of the respondents in the two samples of editors are females, the investigation of editorial interventions might offer a unique perspective into how the tensions between the two preferences play out, and whether one wins out over another (or not).

	Percentage of respondents		
	Female	Mostly Female	Male
Australia (n = 78)	82	0	18
South Africa (n = 58)	83	2	15

Table 5.1: Respondents' gender by country

Next, the respondents were asked to indicate the number of years' editing experience they had by selecting one of five options: 1–10 years, 11–20 years, 21–30 years, 31–40 years and 41 or more years. Figure 5.2 reflects the percentage of respondents per option per country. For both countries, the greatest percentage of respondents had 10 or fewer years editing experience, with the percentage decreasing as the number of years of experience increased. Only two of the Australian respondents had more than 40 years' editing experience, while none of the South African respondents selected this option, although this seems to be counterbalanced by the higher percentage of South African respondents who have 31–40 years' experience, compared to the percentage of Australian respondents with the same amount of experience. Given that the majority of the respondents in the two countries are 50 years old or older, it would appear that in both countries editors enter into the editing profession later in life, possibly after another career.

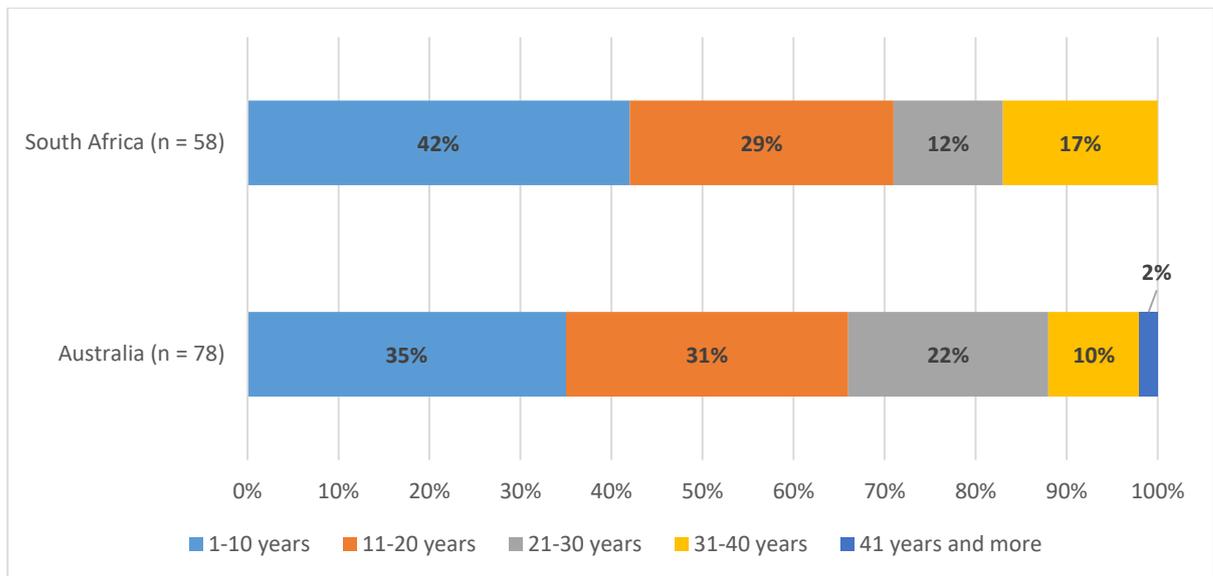


Figure 5.2: Number of years' editing experience

The respondents were next asked to indicate their highest level of qualification, the answers of which are reflected in Table 5.2. The data per qualification type for Honours degree and (Post-)graduate diploma or certificate are presented separately to distinguish the different qualification types; however, both are classified by the respective countries' qualifications authorities as level 8 qualifications. One Australian respondent listed accreditation from IPed as their highest qualification, which was categorised as 'Not specified' since this is not an officially recognised qualification by the Australian Qualifications Authority.

Qualification level	Percentage of respondents: Australia (n = 78)	Percentage of respondents: South Africa (n = 58)
Not specified	1	0
Year 12 / Matric	4	2
First degree or equivalent	26	22
Honours degree	18	34
(Post-)graduate certificate or diploma	14	2
Master's degree	22	34
Doctoral degree	15	5

Table 5.2: Respondents' highest qualifications

Most of the Australian respondents hold a first degree or equivalent, while a Master's degree is the second most frequent highest qualification. The distribution of qualification types among the sample of Australian respondents is similar to that given in IPed's survey of its members in 2016 where 93% held a first degree, 64% a qualification higher than an undergraduate degree and 12% a doctorate (IPed, 2016, p. 3). The greatest proportion of the South African respondents hold either an Honours or a Master's degree, a finding that is similar to previous studies that have surveyed editors in South Africa (Law & Kruger, 2008). The data therefore show that in Australia and South Africa, editors are highly educated individuals, with nearly all holding a post-school or tertiary qualification, and the majority having obtained a post-graduate qualification. The level of education of a person, as with age and gender, has been shown to correlate with

a preference for standard or prestigious forms of language (Labov, 2001; Romaine, 2003; Tagliamonte, 2012), and it may thus be concluded that because the editor respondents in the two samples are highly educated individuals, they will most likely be more conservative language users.

Broad categorisation of qualification	Subcategories of qualification	Percentage of respondents: Australia (n = 78)	Percentage of respondents: South Africa (n = 58)
Commerce	<ul style="list-style-type: none"> • Accounting science • Business and Governance or Management • Information management 	6	5
Education	<ul style="list-style-type: none"> • Education 	10	0
Humanities or Arts	<ul style="list-style-type: none"> • Ancient Language and Text Studies • Communication Science • Language, Linguistics or Literary Studies • Law • Not specified - Arts 	46	81
Science	<ul style="list-style-type: none"> • Agricultural Science • Animal Science • Biological Science • Engineering • Geo- and Spatial Sciences • Health Science • Industrial Design • Not specified - Science • Social Science 	33	12
Senior certificate	<ul style="list-style-type: none"> • Senior certificate 	4	2

Table 5.3: Categorisation of qualification areas

The questions thus far have largely dealt with the respondents' demographic details and provided some insight into the social factors relevant to editors' profiles in the two countries. From the data, it is clear that the respondents in the two samples represent the broader population of editors in the two countries, as evidenced by the comparability of the findings with other surveys of editors in Australia and South Africa. More importantly the data show a general pattern for editors in terms of the age, gender and education level. Given the correlation between age, gender and education level and linguistic behaviour, the fact that the respondents to the questionnaire are mostly older females who hold post-graduate qualifications suggests that editors in the two varieties fit the typical profile of conservative language users who orient their linguistic preferences towards standard or prestigious forms. However, the social factors of age, gender and education level for language users' linguistic behaviour and preferences do not exist in a vacuum, but rather interact with other factors such as working context and linguistic background. The next set of questions therefore sought to gauge information on the editors' working contexts.

The respondents were asked whether they are employed on a full-time or part-time basis and whether they work as in-house editors or as freelance editors. The respondents who indicated that they worked as editors on a part-time basis were asked to provide information on the other kinds of work that they do in addition to editing, while the respondents who worked as in-house editors were asked to describe the context of their in-house employment.

In terms of part-time versus full-time work, the majority of the Australian respondents work as part-time editors (56%), followed by 41% who work full-time. Two of the Australian respondents did not answer this question. For the South African sample, the percentage of respondents who work full-time or part-time is equal (48%), with one respondent indicating that they work as an editor both full-time and part-time (mostly likely through full-time employed in-house work and part-time as a freelancer outside of their full-time employment hours). There is therefore a relatively even distribution across the two countries in terms of whether editing work is done full-time or part-time.

Most of the respondents who work as part-time editors indicated that their other part-time work includes language- and text-based work (Australia, 43%; South Africa, 57%), such as writing, journalism, translation and proofreading, suggesting that most of the part-time editors engage in work related to writing and the creation of texts, in addition to the evaluative work associated with editing. Other kinds of part-time work for the Australian respondents included research-based work (20%), teaching, tutoring and lecturing (16%), administrative work (11%), project management (mostly within the context of publishing) (9%) and work in marketing, communications or media (7%). Other types of part-time work for the South African respondents included work in marketing, communications or media (14%), teaching, tutoring and lecturing (4%) and English marking (4%). A quarter of the respondents in each country did not specify what their other part-time work was.

The overwhelming majority of respondents in both Australia and South Africa work on a freelance or self-employed basis (see Figure 5.3), while just over a fifth in each country are employed in-house. The percentages reported here are in line with current industry employment trends for editors, as reported in previous studies of editors in the two countries (see for example IPEd (2016), Law (2011), and Law and Kruger (2008)). Unlike editors who are employed on an in-house basis, freelance editors work for different clients (and potentially with many different types of texts), each of whom most likely has a preferred norm-providing source to be used for editing work, such as style guides and dictionaries. This means that freelance editors are in all likelihood required to draw on a diverse range of norm-providing sources and are therefore exposed to a less homogenous range of norms and consequently work in a less cohesive normative environment, compared to their counterparts who work in-house.

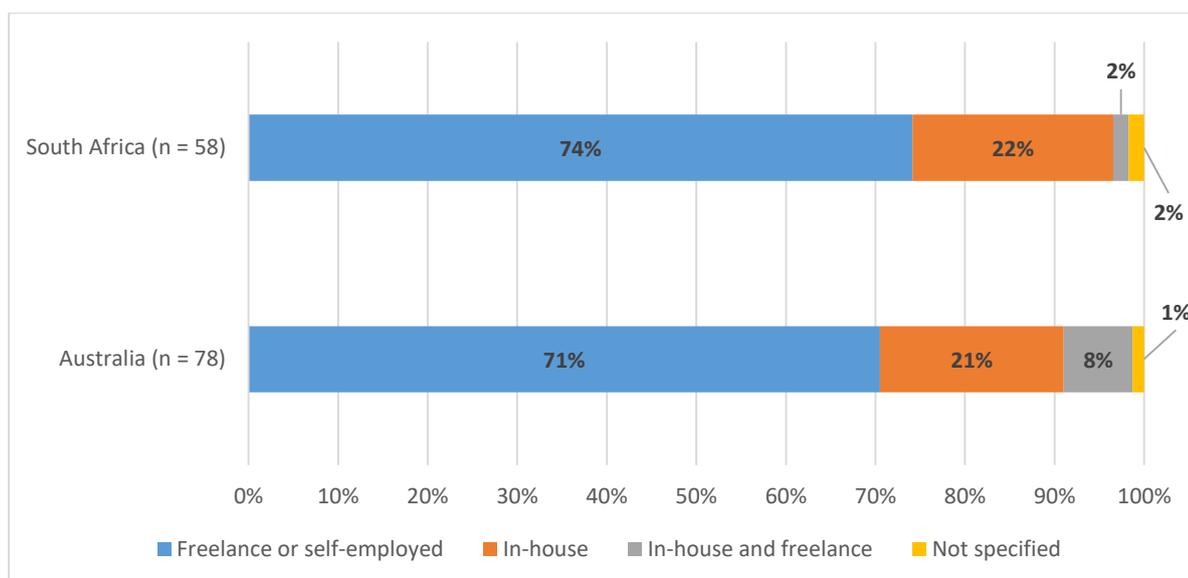


Figure 5.3: Freelance versus in-house employment

In both countries, most editors who work on an in-house basis are employed by government (see Table 5.4). Institutions of higher education are the second most common in-house employers of editors in Australia, while the second most common employer of in-house editors for the South African respondents is publishers. For both the Australian and South African respondents, the third most frequent employer of in-house editors is the corporate sector, which includes businesses, non-governmental organisations and not-for-profit organisations. The editors in the sample therefore represent the typical contexts in which editors would find in-house employment in the two countries. Noble (2014), for example, states that the traditional view that a large portion of editorial work comes from the publishing sector is no longer true and that the corporate sector and government(s) are two of the greatest sources of employment for editors.

	Percentage of respondents (with raw number in brackets)				
	Government	Higher education / Research	Language services agencies	Publishing	Corporate / Business / NPO / NGO
Australia (n = 22)	36 (8)	23 (5)	5 (1)	9 (2)	14 (3)
South Africa (n = 14)	36 (5)	7 (1)	7 (1)	21 (3)	14 (2)

Table 5.4: In-house employment details per country

Editors work on many different kinds of texts and while some (such as those working on an in-house basis) may specialise in certain types of texts, others may work on a wide variety of texts. In order to understand the types of texts typically edited by the respondents, they were asked to indicate all the texts in which they have editing experience. This question consisted of 19 fixed-answer options and one ‘Other’ option which could be used to list texts not included in options provided. All 19 options were selected by the respondents and an additional 128 texts were listed using the ‘Other’ option. Therefore, the data for this question required some pre-processing. After inspecting all of the responses, the decision was made to categorise the texts using the five register labels that were used in the corpus of this study (see Table 5.5) with the addition of four further categories. Although the data from the questionnaire cannot be correlated with the corpus data, it is possible to ensure some continuity across the two data types by using similar labels to

categorise the text types. The respondents' answers were therefore categorised into nine categories (see Table 5.5).

Category	Examples of text types
Academic	Academic articles appearing in scholarly journals, Academic books, Conference proceedings, Dissertations and theses, Research reports
Administrative	Administrative texts (including minutes and institutional communication), CVs, Executive summaries, Funding or grant applications, Reports: Annual, Economic, Productivity, Federal Government and Statistical
Creative	Children's books, Classics, Fiction, Film scripts, Novels, novellas and short stories, Plays, Poetry
Educational	Course materials, Digital learning objects and resources, Schoolbooks, Study guides, Teacher's guides, Training manuals, University textbooks
General informational	Brochures, Event programmes, Fact sheets, Leaflets, Medical and Pharmaceutical texts, Non-government treatises, Product disclosure statements
Instructional	Bills and Acts, Contracts and legal documents, Government guides and policy documents, Instruction manuals, HR manuals, Recruitment processes
Persuasive	Advertising brochures, Marketing materials, Promotional materials, Scripts for adverts, Speeches
Popular	Blog articles, Biographies, Cultural knowledge, Dictionaries, Magazine articles, Memoirs, Religious texts, Trade magazines, Trade non-fiction books, Travel books, Travel guides
Reportage	News reports: online, News reports: print, Newsletters, Press releases

Table 5.5: Categorisation of text types

The distribution of responses per category per country are reflected in Table 5.6, which shows that the samples from both countries include editors who have experience in all the text type categories. The majority of the Australian respondents have experience editing popular, administrative and academic texts. Persuasive and general informational texts are the two text types in which the fewest Australian respondents have experience, while more than half have edited reportage, instructional and educational texts. Academic and administrative texts are by far the two text types in which most of the South African editors have experience. Only one South African respondent has experience editing persuasive texts and just over a quarter have edited creative texts. Similar to the Australian respondents, more than half of the South African respondents have edited popular, reportage, educational and instructional texts.

The distributions across the two countries per text type category are, to a large extent, relatively similar, with the greatest difference in the creative and popular categories. In both of these instances, many more Australian respondents have editing experience than their South African counterparts. The most likely reason for this is that there is probably less work available for South African editors for these kinds of texts. Le Roux and Cassells (2018), for example, state that the South African trade publishing sector (which includes creative and popular texts) accounted for only 23,19% of the industry's turnover in the 2016/2017 financial year and is dominated by multinationals that import texts for the South African market.

	Percentage of respondents: Australia (n = 78)	Percentage of respondents: South Africa (n = 58)
Academic	78	90
Administrative	78	74
Creative	45	29
Educational	51	57
General informational	23	31
Instructional	59	53
Persuasive	9	2
Popular	81	62
Reportage	62	62

Table 5.6: Respondents' experience editing different text types

The last question in this section sought to gather information on how the respondents view their editorial function or role, particularly in terms of a normative versus a communicative function (see Section 2.6.2). The respondents were asked to select one of five fixed-alternative answer options that best described how they view their function as an editor. The five options were:

- an exclusively normative function
- a primarily normative function with communicative elements
- an equally normative and communicative function
- a primarily communicative function with normative elements
- an exclusively communicative function.

The respondents were provided with brief definitions that distinguished between the normative and communicative dimensions. Normative work was defined as work that aims to bring a text into conformance with predefined rules of usage as set out in dictionaries, house styles, style manuals, usage guides and other prescriptive sources. Communicative work was defined as work that aims to ensure that the text communicates optimally with its intended reader, and which may make use of language not (yet) set out in dictionaries, house styles, style manuals, usage guides and other prescriptive sources.

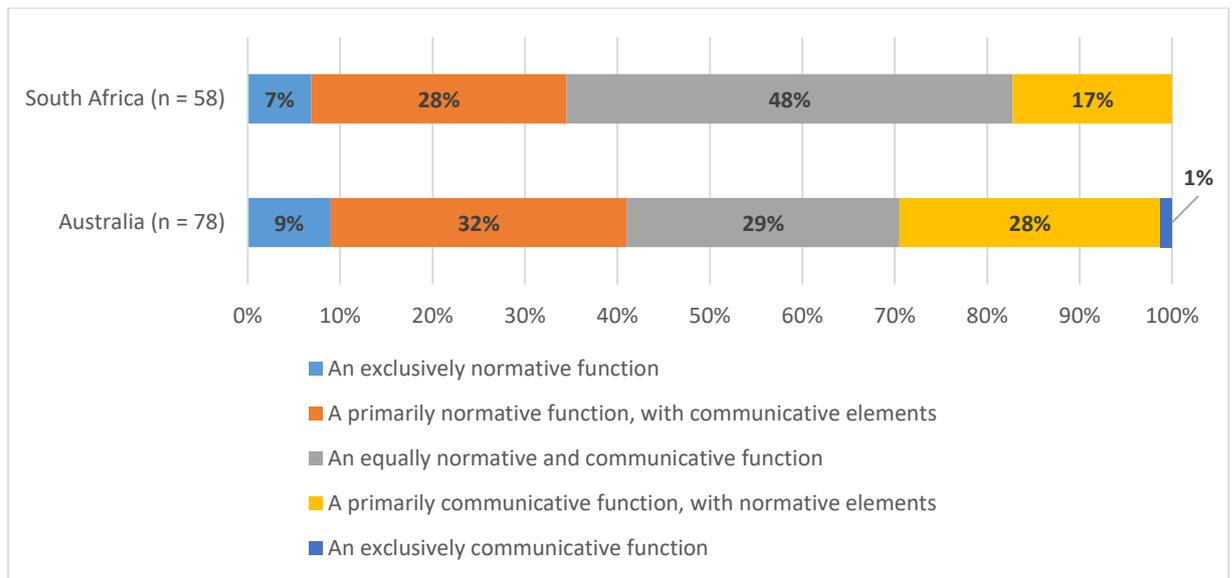


Figure 5.4: Respondents' conceptualisation of their editorial role

Figure 5.4 reflects the respondents' answers to this question and shows that a small percentage of the respondents in each of the two countries hold an exclusively normative view of their editorial function. On the other hand, only one Australian respondent and none of the South African respondents view their role as fulfilling an exclusively communicative function. Overall, it appears that the majority of respondents in both countries view the two functions as important to their work. However, many more South African respondents assign equal weight to the two functions than do their Australian counterparts.

The respondents were also asked to motivate why they held a particular view of their editorial role. This question was included in order to gain a deeper understanding of the factors informing particular editorial approaches. For the respondents who prioritised the communicative dimension, nearly all cited the importance of communication as central to a text's purpose. Many of these editors stated that while the normative dimension of editing was important, clear communication, engagement with the audience and the optimisation of content were far more important. Others cited readers' proficiency (which were in many cases linked to non-native readers) as the primary motivation for holding a communicative view of their role. Some of the more extreme views, as demonstrated in the quotes below, framed normative work as something that is 'part of the job' and which does not add much value:

AusE46: *I take complex documents that are poorly structured, difficult to read and do not comply with my style rules or their organisation's style rules, and extensively rewrite and restructure them to address various issues I encounter: what you describe as 'normative' comes late in the process and does not add a great deal of value, although of course it is essential.*

SAE12: *I match the style of academic articles to the target journal, including referencing. But this is a minor part of my job (tiresome donkey work). I edit in-depth to make the text suitable for the*

intended readership. I sometimes 'repurpose' documents, e.g. rewriting an academic text for a non-academic readership.

In contrast, editors who held a primarily normative view of their work, stated that most of their work was for clients who required that texts adhere to rules set out in specific preferred norm-providing sources (AusE33), while others drew a direct link between their own knowledge of language and their editorial focus (SAfE9), and found the suggestion that editors might introduce language not sanctioned in norm-providing sources unusual (AusE35):

AusE33: *We must ensure that all public communications are of a professional standard.*

SAfE9: *I have a firm background in English grammar and stick to conservative language principles, but reduce jargon and academic speak in favour of clarity and readability [sic] and flow.*

AusE35: *I find the definition of communicative work unusual: typically I would see language that does not yet appear in dictionaries etc. as suboptimally communicative (for example, academic writing that is full of neologisms, noun verbing, unclear expression, etc.). Taking jargon-laden text and making it use easily understood words and commonly used grammatical and typographical conventions is a key part of making sure that the text is meaningful and communicative.*

For the respondents who assigned equal weight to the two dimensions, a clear difference in the motivation for this can be found between the two samples' responses. The prevailing reason among Australian editors who hold this view is that the two dimensions cannot be separated if a text is to communicate optimally with the intended audience (AusE1 and AusE2); however, for the South African editors, their editorial approach is based on either using language that is more inclusive of different readers of English texts (such as first- and second-language users) or helping non-native English authors improve their writing by addressing both communicative and normative dimensions of the text (SAfE52 and SAfE56).

AusE1: *I believe both functions are equally important and interrelated. The normative function (such as removing errors, inconsistencies and other distractions and following a particular style) assists with effective communication. At the same time, using language that the intended audience understands and can relate to is also important and such language can have its own 'norms'.*

AusE2: *Not all communication is about perfect grammar and obeying rules. It depends on the client, the medium and the purpose of that communication.*

SAfE52: *We receive texts from various industries. First and foremost we must ensure it adheres to house style and grammar norms etc but the audiences are diverse so we must in equal parts ensure they understand the content and can relate to it. We must be aware of not excluding anyone in the way something is written.*

SAfE57: *Most of those who ask me to edit their work come from second language English backgrounds and I try to improve the content of their writing in ways that communicate their understanding of the work they are doing. To make it presentable for examination and publication at the institution, to try and communicate with the institution about ways in which they can assist a thesis writer to grasp the essentials of what they are doing (eg. Supplying a House Guide).*

Overall, the responses to the demographic questions suggest that the respondents to the survey broadly represent the population of editors in the two countries (as found in relevant industry surveys), and that the demographic profiles of the editors in the two samples are similar. The respondents are predominantly female, 50 years old or older and appear to have entered editing later in life. Nearly all of the respondents hold post-school or tertiary qualifications, and are therefore highly educated individuals. In both countries, roughly half of the respondents work as editors on a full-time basis while the other half work as editors part-time with most doing other language- and text-based work in addition to editing. Furthermore, the majority of the respondents in both countries work on a freelance or self-employed basis, with about a quarter working as in-house editors. Although the respondents in the two countries appear to have broadly similar views of their approach to their editing work, it does appear that the motivations for particular views differ across the two varieties. For Australian editors, the ability of a text to communicate effectively is the primary reason for holding an equally normative and communicative view of their editorial role. However, for South African editors, the same view seems to be motivated by either a greater need for inclusivity of the range of readers or by the sense that (mostly non-native) authors need assistance to communicate effectively.

5.2.2 Respondents' language profiles

The next set of questions sought to gather information on the respondents' language profiles and included three questions in which participants were asked to indicate their first or home language, their strongest languages (up to a maximum of five) and their editing languages (up to a maximum of five). These questions were included because the language profiles of the editor respondents serve to provide some background to the kinds of linguistic knowledge that editors bring to bear on their editorial interventions, which, together with the information on their age, gender and education level, provide on the one hand information on the kinds of grammatical representations that editors have (as informed by their social-linguistic backgrounds, which give rise to the covert norms) and a framework within which to investigate the interaction between the overt norm-providing sources that the editors in the two countries draw on and their similar/different grammatical representations.

The first question asked the respondents to indicate their first or home language. Four selection options and an 'Other' option were provided. The four selection options were based on the fact that the target population was editors of English texts, but included selection options for the indigenous languages in the two countries. The selection options were 'English', 'Afrikaans', 'an Indigenous Australian language', and 'an

Indigenous South African language’. No selection options were provided to the questions on the respondents’ strongest languages and editing languages, and they were instead asked to list their language(s) in order of strongest to weakest, up to a maximum of five.

In answering the question related to first or home language, only two options, ‘English’ and ‘Afrikaans’, were selected and one ‘Other’ option, Flemish, was provided. Figure 5.5 reflects the answers to this question as a percentage of respondents per language per country. All of the Australian respondents indicated English as their first or home language, while just over half of the South African respondents indicated that Afrikaans is their first or home language, and the remainder selecting English.

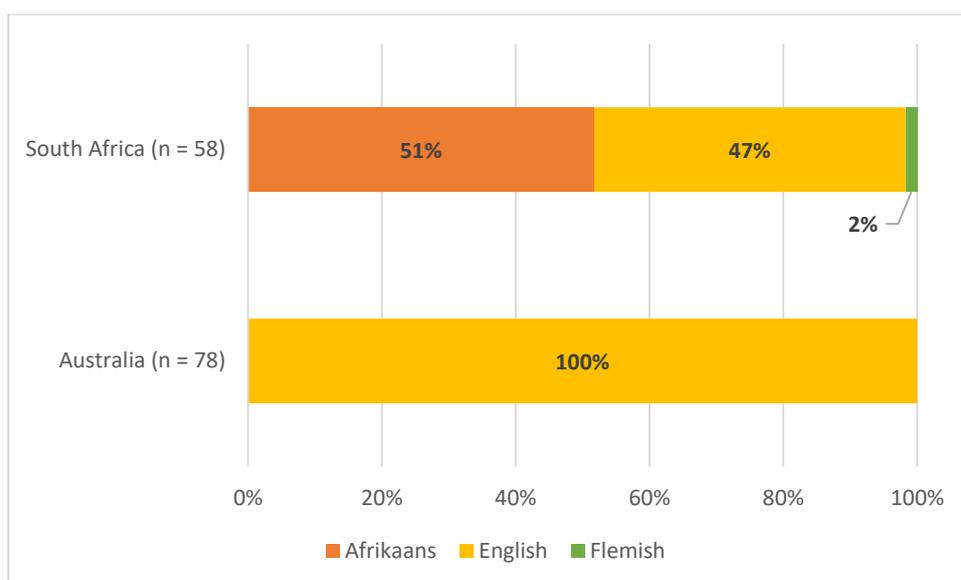


Figure 5.5: Respondents’ first or home language

Unlike their Australian counterparts, slightly more than half of the South African editors of English texts are not first- or home-language users of English. Since the target population of this study is editors of English texts, this finding may seem unusual. However, as discussed in Section 3.2.2, South Africa is a linguistically diverse country, with a constitution that affords official status to 11 languages, including English and Afrikaans. English and Afrikaans are by no means numerically dominant languages in South Africa – a 2011 census of the South African population puts the number of users of Afrikaans as a first or home language at 13,5%, while English is spoken as a first or home language by 9,6% of the population (Statistics South Africa, 2012, p. 24). However, English and Afrikaans are the two most widely used languages for written communication in high functions, including the media, publishing, education and government, and are therefore the two languages in which most editing (and other language- and text-based) work is available. Furthermore, since some of the forums in which the invitation to complete the questionnaire was distributed also represent the interests of translators, it is likely that many of the South African respondents are translators who utilise their proficiency in the languages in which they do translation work to also do work in editing. Approximately half of the South African respondents (48%) indicated that they work as editors on a part-time basis, and more than half of these (57%) indicated that

they do other kinds of language work, like translation, in addition to editing (see Section 5.2.1). Bi- and multilingualism is therefore wide spread in South Africa and is a particularly unique feature of the South African context, and an important feature in the development of the sub-varieties of English in the country.

The bilingualism of the South African respondents is an important factor not only in their selection of sources, which will be examined in more detail in the discussion on the respondents' strongest languages and editing languages, but is also an important consideration in the different kinds of linguistic representations of editors and the effects of these representations on the choices they make in the course of their work. It is well known that for bilinguals, the underlying representational systems of their languages are "constantly active and available during language use activities" (Bialystok, 2007, p. 211), opening up the possibility of the influence of these two active but competing linguistic systems on the linguistic choices of bilingual individuals. This means that high English–Afrikaans bilingualism among the South African editors creates the opportunity for Afrikaans to influence the choices of bilingual editors of English texts.

As a follow up to the question on first or home language, the respondents were asked to list their strongest languages (up to a maximum of five), in order of strongest to weakest. The data for the respondents' answers to this question are discussed in three stages:

- First, their strongest languages are discussed and compared across samples and with reference to the answers to the question on their first or home language.
- Second, the number of respondents that indicated second, third, fourth and fifth languages per sample are presented, and used to interpret the levels of monolingualism, bilingualism and multilingualism within and across the two samples.
- Third, the languages listed by the respondents as their second, third, fourth and fifth strongest languages are discussed and compared.

Only two languages were chosen as the respondents' strongest language: English and Afrikaans. All of the Australian respondents indicated that English is their strongest language, while 57% of the South African respondents indicated that English is their strongest language and the remaining 43% indicated that Afrikaans is their strongest language. As with the question on first or home language, the answers given by the respondents across the two samples, and within the South African sample, differ. Although the South African respondents named the same two languages in answer to this question as they did for their first or home language, the percentages for the two languages are reversed, suggesting that the first or home language of some South African respondents is not their strongest language.

To determine the number of respondents for which this is true, the respondents' answers regarding their first or home languages and strongest languages were categorised into groups based on how the first or home and strongest languages combined (see Table 5.7). These language profiles reveal that the majority of the South African respondents have the same first or home and strongest language, and that this is (almost evenly) distributed across English and Afrikaans. In other words, just less than half of the respondents have

English as a first or home language and regard it as their strongest language, and just less than half have Afrikaans as a first or home language and regard it as their strongest language. However, for 14% of the respondents their language profiles constitute a combination of two languages where their first or home language is not regarded as their strongest language. This is true mostly for the Afrikaans home-language respondents, where 6 of the 30 Afrikaans home-language users indicated that English is their strongest language. Although these findings might seem unusual, studies of the language profiles of Afrikaans home-language users in South Africa reflect a similar trend. For example, Coetzee-Van Rooy (2013) investigates the multilingual repertoires of urban South Africans and reports that in her study, 7,3% of Afrikaans home-language users perceive English as their strongest language. The South African editors in this study therefore represent a somewhat mixed bag of language profiles, in which the same two languages combine in different ways to represent the editors' first or home and strongest languages.

First or home language	Strongest language	Percentage of respondents (n = 58)
Afrikaans	Afrikaans	41 (24)
Afrikaans	English	10 (6)
Flemish	English	2 (1)
English	Afrikaans	2 (1)
English	English	45 (26)

Table 5.7: Comparison of first or home language and strongest language: South Africa

English and Afrikaans are not the only two languages available in the profiles of the respondents to the questionnaire, and many respondents indicated that they know more than two languages. Table 5.8 summarises the percentage of respondents per country who indicated that they have a second, third, fourth and fifth language and shows that while the majority of respondents in both countries have a second language, just less than half also have a third language. Most notably, the percentages reveal that additional languages are much more common among the South African respondents than they are among the Australian respondents.

	Percentage of Australian respondents (n = 78)	Percentage of South African respondents (n = 58)
Second language	60	97
Third language	40	49
Fourth language	14	43
Fifth language	12	26

Table 5.8: Percentage of respondents with second, third, fourth and fifth languages, per sample

To better understand the language profiles of the respondents in each sample, the percentage of respondents who are monolingual, bilingual and multilingual (three or more languages in the profile) were calculated and compared. From the percentages per sample in Figure 5.6, it is clear that many more of the Australian respondents know only one language, compared to the South African respondents where only a smattering indicated that they are monolingual. Only a fifth of the Australian respondents are bilingual but just less than half are multilingual. Bilingualism and multilingualism are much more common among the South

African respondents, compared to the Australian respondents, but given that functional bilingualism and multilingualism are much more common in South Africa than in other countries, such as Australia (Mesthrie, 2002), this difference is not unexpected. However, it does point to differences in the linguistic backgrounds of editors of English texts in the two countries, and in particular to the English–Afrikaans bilingualism of the South African editors.

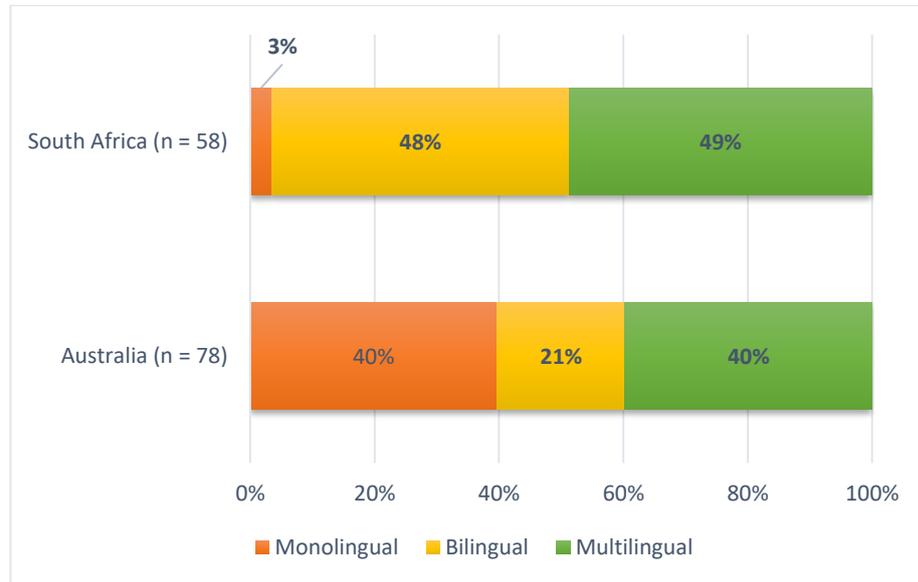


Figure 5.6: Mono-, bi- and multilingualism among the respondents, per sample

The languages that the respondents listed as their second, third, fourth and fifth strongest languages, together with the number of respondents that listed each language, per country, are given in Table 5.9 and Table 5.10.

In total, 15 different languages were listed by the respondents as a second strongest language, of which 12 were listed by the Australian respondents and 5 by the South African. There is therefore a less diverse range of second languages among the South African respondents, even if there are more users of a second language in the South African group. The two most widely cited second languages by the Australian editors are French and German, with French being by far the more popular second language. Four of the respondents listed Afrikaans (an indigenous South African language) as their second strongest language, which may seem unusual but is most likely a reflection of the number of South Africans who have immigrated to Australia. For the South African respondents, English and Afrikaans are the most widely reported second languages, which is to be expected given the prevalence of both of these languages as the respondents' first or home languages and strongest languages. Together with the data on the respondents' first or home and strongest languages, this is further evidence of the high degree of English–Afrikaans and Afrikaans–English bilingualism among the South African respondents.

The absence of the nine official African languages in the list of second strongest languages among the South African editors is striking, for two reasons. First, these languages are used as a first or home language by three quarters of the population (Statistics South Africa, 2012, p. 24). Second, the official status of these

languages is enshrined in the South African constitution and in the additive multilingual language policies of the country. However, despite the official status of the African languages and the sanctioning of their use in the additive multilingual language policies of the country, they are still not widely used in high status functions (Coetzee-Van Rooy, 2013, p. 180). Clearly this has an influence on the availability of work for editors in the different official South African languages, with most work being available in English and Afrikaans.

Second strongest language	Australia (n = 41)	South Africa (n = 56)
Aboriginal English	1	-
Afrikaans	4	25
Arabic	-	1
Auslan	1	-
Chinese	1	-
Dutch	1	-
English	-	25
Flemish	-	1
French	21	4
German	8	-
Greek	1	-
Indonesian/Malay	2	-
Italian	2	-
Latin	1	-
Spanish	4	-

Table 5.9: Respondents' second strongest languages

The data for the respondents' third, fourth and fifth languages are listed in Table 5.10. A total of 26 different languages contributed to the multilingual language profiles of the respondents. In order to facilitate the processing of the data, the languages were clustered, roughly, into four groups:

- Languages indigenous to South Africa: Afrikaans, Fanagalo, Swati, Xhosa and Zulu.
- European languages: Dutch, Esperanto, Finnish, French, German, Hungarian, Italian, Polish, Portuguese, Russian and Spanish.
- Asian languages: Bugis, Indonesian/Malay, Japanese, Mandarin, Thai, Turkish and Vietnamese.
- Classical languages: Ancient Greek and Latin.

European languages, and especially French and German, were by far the most common third strongest languages among both groups of participants. French (n = 7) is the more common third strongest language among the Australians followed by German (n = 4), while German (n = 14) is the more common third strongest language for the South African respondents, followed by French (n = 9). It is interesting, but not all that unexpected given the high levels of Afrikaans–English bilingualism among the South African respondents, to note that other indigenous African languages are only included in the editors' language

profiles as third, fourth and fifth strongest languages, and even then these are consistently less strongly represented than the European languages.

More than a quarter of the Australian respondents listed an Asian language as their third strongest language, which in all likelihood relates to two factors. First, the high number of people from Asian countries who immigrated to Australia after World War II and who continue to do so today might mean that for a number of the Australian respondents, an Asian language is a heritage language (Guy, 1991). Second, Australia's proximity to Asian nations and the realisation during World War II of the disproportional relationship between Australia and Britain gave "rise to a young, self-dependent nation, rooted in her territory, connected with her Asia-Pacific environment" (Schneider, 2007, p. 124). These two factors may also account for the fact that Asian languages make up the bulk of the Australian respondents' fourth strongest languages. Only 9 Australian respondents listed a fifth strongest language, which included 8 of the 11 languages listed overall. There is therefore a high degree of diversity in terms of the Australian respondents' fifth strongest languages, even if these eight languages are grouped into two of the language groups.

	Percentage		Percentage		Percentage	
	Third strongest language		Fourth strongest language		Fifth strongest language	
	Australia (n = 31)	South Africa (n = 38)	Australia (n = 11)	South Africa (n = 25)	Australia (n = 9)	South Africa (n = 15)
English	-	3 (1)	-	-	-	-
Asian	26 (8)	-	64 (7)	-	33 (3)	-
Classical	10 (3)	5 (2)	-	-	-	7 (1)
European	61 (19)	84 (32)	36 (4)	80 (20)	64 (6)	86 (13)
Indigenous	3 (1)	8 (3)	-	20 (5)	-	7 (1)

Table 5.10: Respondents' third, fourth and fifth strongest languages

The respondents' answers regarding their strongest languages paint a picture in which the linguistic profiles of the editors in the two countries are quite different. On the one hand, there are many more English monolinguals in Australia than in South Africa, and by implication many more native users of English in the Australian sample than in the South African. This is to be expected, particularly given the language policies of the two countries and the privileged position that English (and in the case of South Africa, Afrikaans) holds in both. With regard to bilingualism, the South African editors are quite different from their Australian counterparts, and unique for two possible reasons. The languages that make up the bilingual language profiles of the South African respondents include English and one indigenous South African language, Afrikaans, and it is this indigenous bilingualism (together with other socio-economic factors of language use in the country) that gives rise to a group of bilinguals with a high level of bilingual proficiency that is unique in the world (Coetzee-Van Rooy, 2013, p. 199). The bilingualism among these editors is therefore a 'home-grown' type, but it is striking that none of the South African editors are BSAfE users. This means that the work of BSAfE authors are not edited by users of BSAfE, and therefore BSAfE users do not constitute membership of this linguistic gatekeeper community. In contrast, there are only a smattering of bilingual Australian respondents, and where the editor respondents are bilingual, their profiles

are much more diffuse than their South African counterparts and rooted in the inclusion of languages that have been brought into the country, most likely through immigration.

To understand if the high levels of bilingualism and multilingualism among the respondents are also reflected in their editing work, the respondents were asked to indicate the languages in which they do editing in order of strongest to weakest and up to a maximum of five. This question was included to determine if the respondents do editing work in more than one language and how their working languages compared to the languages that they know. The data for the respondents' answers to this question are discussed in three sections:

- First, the respondents' strongest editing languages are discussed and compared across samples.
- Second, the percentage of respondents who indicated that they worked in more than one language are presented and compared across the two samples.
- Third, the languages listed by the respondents as their second, third, fourth and fifth strongest editing languages are discussed and compared.

All of the Australian respondents and most of the South African respondents indicated that English is their strongest editing language. Only a handful of South African respondents indicated that Afrikaans is their strongest editing language (see Figure 5.7). An analysis of the language profiles of the six South African respondents who regard Afrikaans as their strongest editing language revealed that all six also indicated that Afrikaans is their first or home language and their strongest language (see Table 5.11).

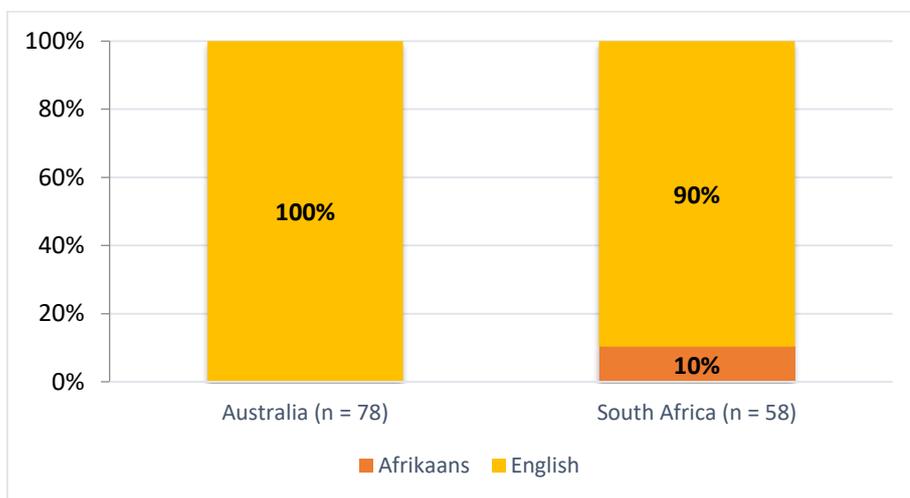


Figure 5.7: Respondents' strongest editing language

The 52 South African respondents who indicated that English is their strongest editing language could be grouped into five categories based on how their first or home language and strongest language combined (see Table 5.11). These combinations show that 18 of these respondents actually have Afrikaans as their first or home language and regard it as their strongest language. These 18 respondents represent 31% of the total South African sample, and therefore a third of English editors in South Africa do not view English as their strongest editing language. Only 26 respondents in the South African sample, which represents 45%

of the total South African sample, indicated that English is their first or home language, strongest language and strongest editing language. All of this means that the South Africa respondents to the questionnaire have diverse language profiles in which English and Afrikaans combine in several different ways not only to represent the respondents' first or home languages and strongest languages, but also their strongest editing languages.

First or home language	Strongest language	Strongest editing language (English)	Number of respondents (n = 52)
Afrikaans	Afrikaans	English	18
Afrikaans	English	English	6
English	Afrikaans	English	1
English	English	English	26
Flemish	English	English	1

Table 5.11: Combination of respondents' first or home, strongest and strongest editing language(s): South Africa

The high levels of Afrikaans–English bilingualism among South Africans alluded to earlier in the discussion are confirmed by these findings, but also the findings regarding the respondents' second, third, fourth and fifth strongest editing languages. Hardly any respondents in either sample indicated that they edit in a third, fourth and fifth language; however, the majority of the South African respondents indicated that they edit in a second language, while only a handful of the Australian respondents indicated that they did too (see Table 5.12).

Of the South African respondents (n = 33) who edit in a second language, 25 edit in Afrikaans, 6 edit in English and one each edit in Spanish and French. The 25 respondents whose second strongest editing language is Afrikaans, all indicated that English was their strongest editing language, while the 6 whose second strongest editing language is English, all indicated that Afrikaans was their strongest editing language.

	Percentage of respondents	
	Australia (n = 78)	South Africa (n = 58)
Second language	4	57
Third language	3	3
Fourth language	3	2
Fifth language	1	2

Table 5.12: Respondents who edit in second, third, fourth and fifth languages

The 25 South African respondents who indicated that they edit in only one language, all indicated that this language is English. Two of these respondents are monolingual, 10 are bilingual and 13 are multilingual. Five have Afrikaans as first or home language, one has Flemish as first or home language and the remaining 19 have English as first or home language. Of the five Afrikaans first or home language respondents, three are English–Afrikaans bilinguals while the other three are Afrikaans–English bilinguals. It is interesting that the three Afrikaans–English bilinguals, who all have Afrikaans as their first or home and strongest language, only edit in English.

The discussion of the respondents' demographic information in Section 5.2.1 reveals that the typical profile of the editors in this study (and also of those within the broader population) is very similar across the two countries. However, representing editors only in terms of particular social factors such as their age, gender or education level is an oversimplification and misrepresentation of their complete backgrounds. Instead, context and language profile need to be considered too. While one might be lead to believe that the editor respondents form a homogenous group, the data on the respondents' language profiles paint an entirely different picture and show important differences in the language profiles of the editor respondents from the two countries, all of which have implications for their linguistic knowledge and consequently, their editorial practice.

In terms of language profiles, the Australian editors form a homogenous group in which editors are overwhelmingly first language or native users of English. Even though many of the Australian respondents indicated that they know more than one language, their proficiency in languages other than English was not such that they edit in this additional language. The group of South African editors is, in comparison, more diffuse and represents a combination of profiles that include both first- and second-language users of English, most of whom edit in English and Afrikaans. The significance of this is not to be underestimated: as alluded to earlier in the discussion (and in Section 2.2.1), it is well known that the underlying representational systems of bilingual language users' languages are active and available during language use activities (Bialystok, 2007, p. 211), such as editing. Given that English and Afrikaans are very strongly represented in the minds of many of the South African editors, it might be that in terms of the editorial interventions of these editors, the representation of the two languages in the minds of these editors acts as a point of contact between the two languages, opening up opportunities for the two linguistic systems to influence the linguistic choices of bilingual editors. A further crucial finding with regard to the language profiles of the South African editors is that none of these editors are first- or second-language users of BSAfE. This means that the editors who work on the texts of BSAfE authors are users of different strands of English, which limits the opportunity for the same kind of cross-linguistic transfer effect, but also means that the representation of the different strands of English among South African editors is skewed in favour of the WSAfE and AfrE.

5.2.3 Norm-providing sources

As mentioned in Section 5.2, the second primary aim of the questionnaire was to gather information on the kinds of norm-providing instruments used by editors of English texts in Australia and South Africa as a way to determine what kinds of sources editors in the two varieties use and whether the sources used reflect local norms or not. Such sources serve as overt norms for editors (see Section 2.6.1). The respondents were therefore asked to list all of the sources that they use in the course of their editing work. Although the questionnaire initially divided these sources into four broad categories (online dictionaries, print dictionaries, usage guides and style manuals), the respondents' answers to these questions demonstrated that editors draw on many different kinds of sources for their work, well beyond the scope covered by the

four categories, and therefore the decision was made to extend the four categories to include the additional kinds of sources that emerged from the data. Nine broad categories of sources were thus identified: dictionaries, glossaries, thesauruses, style guides, usage guides, editing and writing sources, language reference works, people, and various online and print sources. (See Section 4.2.4 for an explanation and definition of the different source type categories.)

The discussion of the results regarding the respondents' use of norm-providing sources and their orientation is distributed over four sections. In the first section, an overview of the responses and their categorisation is given. Section 5.2.3.1 provides information on the number of responses per source type category and the number of distinct sources per category. In addition, summary data on the number of respondents in each country that make use of each source type are presented. In Sections 5.2.3.2 and 5.2.3.3, the responses by the Australian and South African editors are analysed in terms of the norm orientations of the source types used and the most frequently listed sources in each source type category. The last section, Section 5.2.3.4, concludes the discussion by comparing the results for the Australian and South African editors.

5.2.3.1 Overview of responses

Each source listed by the respondents was individually inspected to confirm its bibliographic information, to determine the source type category to which it belonged and to determine its linguistic norm orientation. The complete list of sources, categorised by source type and orientation, is reflected in Appendix C. In total, 720 individual responses were recorded for the Australian respondents, while 567 individual responses were recorded for the South African respondents. Although it may seem that many more responses were returned for the Australian sample, proportionally, the number of responses translates to 9,23 per respondent for the Australian sample and 9,77 per respondent for the South African sample, reflecting a similar number of responses per respondent per country. Table 5.13 reflects the number of responses per source type category per country as well as the number of distinct sources per source type category per country.

Source type category	Australia		South Africa	
	Number of responses	Number of distinct sources	Number of responses	Number of distinct sources
Dictionary	259	41	255	59
Thesaurus	7	5	7	7
Glossary	1	1	3	3
Style guide	255	48	132	44
Usage guide	71	12	37	8
Editing and writing sources	57	27	53	18
Language reference works	40	19	51	17
People	8	4	16	8
Various (online and print)	22	12	13	6
Total	720	169	567	170

Table 5.13: Number of responses and number of distinct sources per source type category

The same five source types returned the highest number of individual responses for both the Australian and South African editor respondents (dictionaries, style guides, usage guides, editing and writing sources and language reference works) and will therefore form the focus of the following discussion.

It comes as no surprise that dictionaries and style guides returned the highest number of individual responses for both groups of editors: dictionaries and style guides are the two source types most typically associated with editorial work and are generally viewed as absolutely essential within the context of copy- or sub-editing. For both respondent samples, dictionaries accounted for more than 250 of the total number of responses. A near similar number of responses was returned by the Australian respondents for style guides, while the number of responses for style guides by the South African sample is considerably lower, both in comparison to the Australian sample and in comparison to the number of responses for dictionaries by the same sample. Despite the lower number of responses for style guides by the South African sample, collectively, dictionaries and style guides accounted for just over two thirds of the total number of responses for both samples (Australia, 514/720; South Africa, 387/567).

Usage guides, editing and writing sources and language reference works respectively returned the third, fourth and fifth highest number of individual responses for the Australian respondents. Editing and writing sources, language reference works and usage guides respectively returned the third, fourth and fifth highest number of responses from the South African sample. Although there are slight differences across the two samples in terms of the number of individual responses for these three source type categories, collectively they account for a quarter of the total number of responses for each sample (Australia, 168/720; South Africa, 141/567).

The remaining source type categories – thesauruses, glossaries, various online and print sources, and people – all returned a considerably lower number of individual responses in comparison to the top five. Despite this, the fact that there were data for these categories clearly suggests that, at least for some editors, these sources are consulted in the course of their work, and points to the diversity of sources that editors, from both samples, draw on.

To further understand how the different source types are represented within the two samples, the number of distinct sources per source type category was calculated (see Table 5.13). While the respondents were asked to provide edition information on each source they listed or selected, the calculation of distinct sources in Table 5.13 does not take this into consideration and instead treats different editions of the same source as one distinct source. Information on the different editions will be discussed in Sections 5.2.3.2 and 5.2.3.3, which deal with the country-specific responses. The data on distinct sources show that the five source types that returned the highest number of individual responses also accounted for the widest range of distinct sources. Dictionaries offered the widest range of choice for the South African sample, who indicated that they used 59 different dictionaries, of which 44 were print dictionaries and 15 were online.

The Australian respondents draw on a slightly narrower range of dictionaries (41) compared to their South African counterparts, with more print dictionaries represented (24) than online (17).

Style guides offered the widest range of choice among the Australian respondents (48), and the second widest range among the South African respondents (44). In both cases, the respondents listed various kinds of style guides, such as in-house style guides as well as style guides published by book publishers and various scientific institutions that carry weight beyond their own organisation (Peters, 2014b). With regard to in-house style guides, the respondents varied in how much information was specified for the source(s) they listed. For example, some respondents simply listed “in-house style guide” while others named the organisation to which the in-house style guide belongs. As a result of this, the number of distinct sources listed here may not accurately reflect the full range of style guides used by the editors, since 44 individual responses from the Australian sample and 19 from the South African sample were for unspecified in-house style guides and are therefore counted as one distinct source. Nevertheless, dictionaries and style guides still accounted for the greatest range of distinct sources used by the editor respondents.

Editing and writing sources included the third highest number of distinct sources for both the Australian (27) and South African (18) respondents. Clearly, the Australian respondents draw on a wider range of editing and writing sources than their South African counterparts.

The data show that language reference works accounted for the fourth widest range of distinct sources for editors from both countries (Australia, 19; South Africa, 17). Only four of these sources were shared by respondents in both countries, suggesting that while a similar number of distinct sources are represented by the respondents from the two countries, the actual sources used are not the same. For both samples, usage guides accounted for the fifth widest range of distinct sources (Australia, 12; South Africa, 8), and five of those sources were shared by editors in both samples. Editors in both countries therefore draw on a narrower range of usage guides compared to the other source type categories.

	Australia (n = 78)		South Africa (n = 58)	
	Number	Percentage	Number	Percentage
Dictionary	78	100	55	95
Style guide	76	97	52	90
Usage guide	43	55	23	40
Editing and Writing sources	31	40	27	47
Language reference works	30	38	36	62
Various online and print sources	17	22	8	14
People	6	8	12	21
Thesaurus	5	6	5	6
Glossary	1	1	3	5

Table 5.14: Number and percentage of respondents per source type category

Although the number of responses per source type category and the number of distinct sources per source type provide some descriptive information on the prevalence of the different source types among the respondents, they do not provide information on the number of editors in each of the two samples that actually use each of the source types. Table 5.14, which reflects the number of respondents per source type category per sample, shows that the five source type categories that returned the highest number of individual responses were also used by the greatest number of respondents, which is further evidence of their importance for editorial work.

All of the Australian respondents (78) indicated that they use a dictionary; 72 make use of online dictionaries and 48 make use of print dictionaries. Three South African respondents indicated that they do not make use of dictionaries at all, and of the 55 who indicated that they do use dictionaries, 50 make use of online dictionaries and 38 make use of print dictionaries. It should be noted that two of the South African respondents who do not make use of dictionaries indicated that they make use of *Pharos Online*, an online language reference tool (categorised under language reference works) that includes a number of SAfE, Afrikaans and other South African language dictionaries, an editing and writing handbook and a style guide. Online dictionaries are therefore the more popular choice of dictionary type among both the Australian and South African respondents. The source type used by the second highest number of respondents is style guides, and the overwhelming majority of respondents in each of the two samples indicated that they make use of style guides

Just less than half of Australian and South African respondents indicated that they use editing and writing sources (Australia, 40%; South Africa, 47%), which seems to reiterate the shared importance of these sources for editors in the two countries. With regard to usage guides and language reference works, there is less comparability between the two countries in terms of the number of respondents who use these sources. For example, more than half of the Australian respondents make use of usage guides, while just less than 40% of the South African respondents do. Usage guides are therefore more widely used among the Australian respondents than the South African. There may be several reasons for this, the most likely of which is that there are more locally-orientated usage guides for AusE than there are for South African (see Sections 5.2.3.2 and 5.2.3.3). To compensate for this, it seems that South African editors may turn to language reference sources, which are used by more than 60% of the South African respondents, compared to the 38% of Australian respondents who indicated that they make use of these types of sources. Nevertheless, despite these differences it is clear that dictionaries, style guides, editing and writing sources, language reference works and usage guides are the most commonly used sources among both sets of editors.

Only a handful of Australian and South African respondents make use of glossaries and thesauruses, and only a few of the Australian respondents indicated that they consult people to find solutions to the problems they encounter in their editing work. Just over 20% of the South African respondents indicated that they consult people in the course of their work, such as colleagues or other editors via PEG's chat group. Nearly 22% of the Australian respondents make use of various online and print sources, such as Google, Wikipedia

and various organisations' websites, while only 14% of the South African respondents make use of these kinds of sources.

The overview of the data presented here shows a fairly consistent distribution of responses across the two samples. However, there are some noteworthy findings. Editors appear to draw on a large range of sources in the course of their editorial work that extend beyond the codification instruments most typically associated with editing (dictionaries and style manuals). The responses also show that for both the Australian and South African respondents, the same five source types returned the highest number of responses and were the five source types that offered the greatest range of distinct sources. This is particularly the case for dictionaries and style guides, which are used by the majority of respondents in both samples, but are also the two source types with the greatest number of distinct sources. Since dictionaries are in all likelihood used to obtain guidance on spelling (and probably also word meaning) and because style guides are highly prescriptive, it seems that for most editors, the most overtly prescriptive of the sources are the most widely used, suggesting that editors are purveyors of prescriptivism. It is striking that these are also the two source types that had the highest number of distinct sources, suggesting that editors draw on a very diffuse set of these prescriptive sources, compared to the other sources. Editing and writing sources, language reference works and usage guides are widely used among the respondents; however, the number of respondents who make use of these sources is considerably lower than those for dictionaries and style guides, as is the number of distinct sources listed for each. A fifth of the South African respondents indicated that they consult people in the course of their work, while only a handful of the Australian respondents indicated that they do this. This might speak to the lack of codification instruments available for SAfE (see Section 5.2.3.3), compared to those available for AusE. Perhaps South African editors compensate for this by consulting fellow editors or colleagues when they encounter an editing problem unique to SAfE.

5.2.3.2 Australian respondents

The codification of AusE in local dictionaries, style guides and usage guides is well known, as is the recognition and acceptance of AusE as a national variety and as one of the major reference varieties of English, alongside AmE and BrE (Peters, 2014a, 2014b; Schneider, 2007). In comparison to other varieties of English represented by the sources listed, AusE dictionaries, style guides and usage guides returned the highest number of individual responses for each of these source type categories by the Australian respondents. In fact, nearly half of the responses for dictionaries (121 out of 259), style guides (100 out of 255) and usage guides (34 out of 71) were Australian. Of the 682 individual responses for the five source types most used by the Australian editors, 44% are oriented towards the Australian norm, compared to 23% which represent AmE and 18% which represent BrE.

At first glance, it appears as if, in terms of dictionaries, style guides, usage guides, language reference works and editing and writing sources, Australian editors prefer drawing on endonormative sources. However,

when the individual responses for each of the source types are tallied along the lines of whether they reflect a local norm (AusE) or an external norm (all other norm orientations represented in the sources), it becomes clear that there is a roughly equal distribution between exonormative and endonormative sources, with only slightly more exonormative sources than endonormative. Table 5.15 shows that 46% of the individual responses are for sources that reflect a norm external to AusE, 44% reflect the local norm and 10% are of indeterminate norm orientation. This data may be interpreted as demonstrating a tendency amongst the Australian respondents to draw equally on sources reflecting either a local or an external norm, but this may not actually be the case. Because the respondents had the opportunity to (and did) list all of the distinct sources that they use in the course of their work, the number of responses for the different norm orientations in Table 5.15 reflect the range of sources the editors use, and not the sources used most often by the Australian editors. To determine which sources are most widely used among the Australian respondents, all distinct sources that were listed by at least 10% of the respondents were identified.

Orientation	Dictionaries	Style guides	Usage guides	Editing and writing sources	Language reference works	Total
American English	39	96	1	10	9	155
American and British English	2	-	-	-	2	4
Multinational	21	-	-	-	5	26
Australian English	121	100	34	33	9	297
British English	59	10	34	11	12	126
English and a foreign language	2	-	-	-	-	2
Indeterminate	15	49	1	1	3	69
South African English	-	-	1	2	-	3
Total	259	255	71	57	40	682
External	123	106	36	23	28	316
Internal	121	100	34	33	9	297
Indeterminate	15	49	1	1	3	69

Table 5.15: Number of responses per norm orientation for the top five source types: Australia

As indicated in Section 5.2.3.1, all of the Australian respondents ($n = 78$) indicated that they make use of at least one dictionary. Table 5.16 lists the twelve dictionaries, in order of popularity, that are used by at least 10% of the Australian respondents. Two of the six dictionaries listed represent dictionaries produced in Australia since the 1980s (Laugesen, 2014; Peters, 2014b) and are thus clearly endonormative. Two reflect a BrE orientation, one an AmE orientation, while one is multinational (representing different varieties of English used in the world without according primary status to any one).

	Dictionary	Print or online	Orientation	Percentage of respondents (n = 78)
1.	<i>The Macquarie Dictionary</i> https://www.macquariedictionary.com.au/	Online	Australian English	79
2.	<i>The Macquarie Dictionary</i> Sydney: Macquarie Dictionary Publishers Pty. Ltd.	Print	Australian English	54
3.	<i>Merriam-Webster Dictionary and Thesaurus</i> http://www.merriam-webster.com/	Online	American English	35
4.	<i>Oxford English Dictionary (OED)</i> http://www.oed.com/	Online	Multinational	30
5.	<i>Oxford Dictionaries</i> https://en.oxforddictionaries.com/definition/	Online	British English	24
6.	<i>Collins English Dictionary</i> http://www.collinsdictionary.com/dictionary/english	Online	British English	13

Table 5.16: Most frequently selected dictionaries by the Australian respondents

The online and print versions of the *Macquarie Dictionary* are the two most frequently listed dictionaries by the Australian respondents, with the majority of respondents indicating that they use the online version and just over half indicating that they use the print version. The popularity of the *Macquarie Dictionary* among the Australian editor respondents is not surprising, particularly given that it is the most frequently updated dictionary of the AusE lexicon (Peters, 2014b, p. 585), having established itself as “a hallmark of Australia’s national identity” (Schneider, 2007, p. 52) and been adopted by Australian governmental, educational, broadcasting and journalism organisations (Delbridge, 1999). The *Merriam-Webster.com* online dictionary, originally based on the eleventh edition of the *Merriam-Webster’s Collegiate® Dictionary*, is the third most popular dictionary among the Australian editor respondents, with just over a third of respondents making use of the dictionary. It is noteworthy that a much smaller number of editors make use of *Merriam-Webster.com* compared to the online and print versions of the *Macquarie Dictionary*, highlighting the dominant status of the *Macquarie Dictionary* among AusE editors.

Given that AusE achieved endonormative stabilisation in the late twentieth century and that it has been codified in a number of AusE dictionaries, it is surprising that an overtly AmE dictionary would be used among a cohort of AusE editors. The most likely reasons for this are probably the fact that access to *Merriam-Webster.com* online is free and that Webster dictionaries are often the preferred dictionaries to be used alongside AmE style guides (see discussion of style guides). However, it might also be evidence of the recently noted influence of AmE words and expressions on AusE as a result of the rapid development in communication technologies and the “increasing influence of American culture in Australia ... as Australian English is brought closer than ever before to the ‘now’ of American English” (Collins, 2014, p. 452). The extent of the influence of AmE on AusE is a topic on which opinions differ, but the very fact that an AmE dictionary is used by just over a third of the Australian editors might point to (at the very least) an awareness of and (to more than negligible degree) a tacit acceptance of AmE norms among some AusE editors.

The online version of the *Oxford English Dictionary* (OED) was the fourth most popular dictionary among the Australian editor respondents, and was selected by just less than a third of the respondents. The version

of OED available online reflects the current revision process for the third edition, which seeks to widen the geographical coverage of the dictionary by including entries that reflect the world's Englishes and remove the BrE bias of earlier editions (Salazar, 2014). The OED online is therefore a multinational dictionary of the English language as it is used in different regions and does not model its entries on a particular variety. Since the OED is widely regarded as the most authoritative and comprehensive dictionary of the English language, it is not surprising that it is one of the top four most widely used dictionaries among the Australian editor respondents. Its lower position in comparison to *Merriam-Webster.com* might be due to the fact that access to the OED online is via paid subscription, whereas access to *Merriam-Webster.com* is free.

BrE dictionaries came in fifth and sixth place, and are far less frequently mentioned, and therefore used, by the Australian editors. The lower percentage of Australian editors who make use of BrE dictionaries compared to those who make use of AusE and AmE dictionaries demonstrates a very clear preference for endonormative sources among the Australian editors, and provides some support for the increasing influence of AmE in Australia, which might be bolstered by the fact that access to the dictionaries mentioned are free. Whatever the motivation, it is noteworthy that there is a decline in the influence of the historically important British norm. This preference for AusE and AmE sources is also demonstrated in the number of AusE and AmE style guides that were listed by the respondents.

Many organisations in Australia provide their employees with style guides (Delbridge, 1999). Peters (2014b) points out that the codification of key elements of Australian orthography and editorial style started in the 1960s with the publication of the *Australian Government Style Manual*, which has since been revised in six editions and is now the preferred style guide of many Australian book and newspaper publishers. The popularity of the *Australian Government Style Manual* and in-house guides among publishers/producers of English texts in Australia is reiterated in the most frequently listed style guides by the Australian editor respondents (see Table 5.17).

	Style guide	Type	Orientation	Percentage of respondents (n = 76)
1.	<i>Australian Government Style Manual for Authors, Editors and Printers</i> Milton Qld: John Wiley.	Association	Australian English	92
2.	<i>In-house Style Guide</i> (44, unspecified; 13, specified)	In-house	Indeterminate	75
3.	<i>The Chicago Manual of Style</i> Chicago: University of Chicago Press.	Book publisher	American English	49
4.	<i>Publication Manual of the American Psychological Association</i> New York: American Psychological Association.	Association	American English	29
5.	<i>The Elements of Style</i> New York: Harcourt/Macmillian/Pearson Longman.	Popular	American English	21

Table 5.17: Most frequently selected style guides by the Australian respondents

Most of the Australian respondents (76) indicated that they make use of a style guide. The overwhelming majority indicated that they use the *Australian Government Style Manual for Authors, Editors and Printers* (82% of whom use the sixth edition, 7% the fifth edition, and 3% the first edition). In-house style guides

are used by three quarters of the respondents, which is the second most frequently listed style guide among the Australian editors. Only three other style guides were listed by at least 10% of the respondents, and all of these were American English style guides. By far the most popular of these was the print version of the *Chicago Manual of Style*, which is used by half of the respondents.

There is obviously a clear preference among the Australian editor respondents for endonormative style guides, as evidenced by the popularity of the *Australian Government Style Manual*, but given the availability of style guides for AusE, it is curious that three of the top five guides are for AmE, and none are for BrE. It might be that AmE style guides such as *The Chicago Manual of Style* and the *Publication Manual of the American Psychological Association* are used simply because either is often the preferred style guide of publishers of academic texts (a text type in which the majority of the Australian editors indicated that they have experience). This might also explain the number of Australian editors who make use of AmE dictionaries.

The findings for the Australian editors' use of style guides, together with those for dictionaries, point to the widespread acceptance and use of local norms among Australian editors; however, the higher frequency of use of AmE dictionaries and style guides in comparison to BrE dictionaries and style guides, suggest a movement away from BrE 'input' in Australia, and possibly towards a greater acceptance of AmE norms.

Usage guides on AusE first appeared in the 1980s and initially drew on the structure, content and style of Fowler's model (Delbridge, 1999), providing personal impressions of usage rather than providing systematic evidence of current usage (Delbridge, 1999; Peters, 2014b). In the 1990s, the *Cambridge Australian English Style Guide* (which, in the late 2000s, reappeared in a second edition under the title *Cambridge Guide to Australian English Usage*) broke away from the tradition, and instead adopted a descriptive approach to AusE usage that drew on corpus-linguistic methods (Delbridge, 1999). In keeping with the popularity of endonormative sources among the Australian editor respondents, two AusE usage guides are among the Australian editors' top four most popular choices (see Table 5.18).

	Usage guide	Orientation	Percentage of respondents (n = 43)
1.	<i>Cambridge Guide to Australian English Usage</i> Melbourne: Cambridge University Press	Australian English	49
2.	<i>Fowler's Modern English Usage</i> Oxford: Oxford University Press	British English	42
3.	<i>Cambridge Guide to English Usage</i> Cambridge: Cambridge University Press	British English	26
4.	<i>Modern Australian Usage: A Practical Guide for Writers and Editors</i> Melbourne: Oxford University Press	Australian English	16

Table 5.18: Most frequently selected usage guides by the Australian respondents

The *Cambridge Guide to Australian English Usage*, which was listed by half of the respondents, was the most popular choice, with the majority of these respondents (61%) drawing on the most recent edition. *Modern Australian Usage* was the fourth most popular choice of usage guide, and of the respondents who

indicated that they use this guide, 86% made use of the most recent 2015 edition. BrE usage guides accounted for the remaining two top spots: *Fowler's Modern English Usage* was the second most popular choice, with most of the respondents (9 of the 18), most of whom were 50 years old or older, indicating that they make use of Gower's now rather dated 1965 edition. Burchfield's third and third revised editions are used by two and six respondents respectively, while the most recent edition, Butterfield's fourth edition, is used by only one respondent. The archetypal status of *Fowler*, among linguists and non-linguists alike, is the most likely reason why nearly as many respondents who use the *Cambridge Guide to Australian English Usage* use this usage guide. The absence of AmE usage guides in the list of most used usage guides is striking, particularly given the popularity of AmE dictionaries and style guides.

As pointed out in Section 5.2.3.1, dictionaries and style guides are widely considered to be absolutely essential for editing work, setting out rules for issues related to lexis and style. Such issues are usually addressed during copyediting, the heart of the editorial process (Mackenzie, 2011), and it might be that because most editors work on texts published by particular institutions, and because most institutions typically inform the editor of the preferred dictionary and style guide to be used, and less often advise on the use of other norm-providing sources, editors use usage guides most relevant to the variety they are editing or the most well-known usage guides.

In addition to the sources already discussed, the Australian editor respondents also indicated that they make use of various editing and writing sources to guide their editorial choices. Just less than half of the total number of Australian respondents indicated that they make use of editing and writing sources, of which six were listed by at least 10% of the respondents. All except one of these are locally produced editing and writing handbooks, suggesting a clear preference for Australian editing and writing sources among the editor respondents (see Table 5.19). *Butcher's Copy-editing*, the third most popular editing and writing source is a British copy-editing handbook, which, according to the Society of Editors and Proofreaders (SfEP) (2015), is widely considered as the 'copy-editor's bible', and is therefore unsurprising in its inclusion in the list.

	Editing and writing source	Type	Orientation	Percentage of respondents (n = 31)
1.	<i>Working Words</i> Manuka, ACT: Canberra Society of Editors.	Handbook	Australian English	26
2.	<i>Effective Writing: Plain English at Work</i> Melbourne: Pitman.	Handbook	Australian English	23
3.	<i>Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors, and Proofreaders</i> Cambridge: Cambridge University Press.	Handbook	British English	19
4.	<i>The Australian Editing Handbook</i> Milton, Qld: John Wiley & Sons.	Handbook	Australian English	16
5.	<i>The Editor's Companion</i> Port Melbourne: Cambridge University Press.	Handbook	Australian English	16
6.	<i>The Little Red Writing Book</i> Sydney: UNSW Press.	Handbook	Australian English	10

Table 5.19: Most frequently selected editing and writing sources by the Australian respondents

Language reference works, as reported in Section 5.2.3.1, are the fifth most common source type used by the Australian editors. A total of 30 of the 78 respondents indicated that they make use of this type of source and only 3 sources in this category are used by at least 10% of the respondents. The three most popular sources represent an eclectic range of language reference works published by two individuals and a well-known university press.

Language reference work	Type	Orientation	Percentage of respondents (n = 30)
<i>The Little Green Grammar Book</i> Sydney: UNSW Press.	Print book	Australian English	30
<i>Grammar Girl: Quick and Dirty Tips</i> https://www.quickanddirtytips.com/grammar-girl	Language website	American English	27
<i>Oxford Reference Online</i> www.oxfordreference.com	Online language reference	Multinational	17

Table 5.20: Most frequently selected language reference works by the Australian respondents

Most respondents indicated that they use Mark Tredinnick's *The Little Green Grammar Book*. Tredinnick is a popular Australian poet and essayist, who, although admitting to not being a grammarian, does teach grammar. *The Little Green Grammar Book* is described as "a book of grammar for the writer in everyone" (Tredinnick, 2008, p. 3) and is therefore a popular language reference work.

The US-based grammar blog, *Grammar Girl: Quick and Dirty Tips*, which has become an increasingly popular source of grammar advice, is the second most popular language reference work cited by the Australian respondents. Over a quarter of the respondents indicated that they consult this blog, and while this number might be quite low, it represents just over 10% of all the Australian respondents. The high ranking of this blog might seem unusual among a cohort of professional editors, but Lukač (2017) argues that grammar blogs authored by individuals, such as *Grammar Girl*, have established themselves as modern day, online, language authorities, and it is very likely that editors use this blog as a quick way to find answers to any grammar questions they may have. It may therefore be that ease of access and popularity are two reasons motivating the editors' use of particular sources.

The *Oxford Reference Online* was the third most popular language reference work. *Oxford Reference Online* offers access to a range of Oxford University Press publications via annual subscription, including general and subject-specific dictionaries, usage guides, companions and encyclopaedias, and represents a range of English-speaking regions, making it a multinational language reference source.

Very few of the Australian editor respondents indicated that they make use of glossaries (n = 1), thesauruses (n = 5) and people (n = 6), which most likely speaks to the perceived unimportance of these sources for the editing work of the majority of the Australian editor respondents, and suggests that the respondents are able to find sufficient advice in the top five source types. The last source type category, various online and print sources, was listed by 17 respondents and includes a diverse range of sources used. The only specific source in this category that was mentioned by more than 10% of respondents is Google's search engine, which the

respondents indicated is used to quickly look up and confirm the spelling of subject-specific terms or to determine the frequency of usages (if not covered in a dictionary or style guide).

The Australian editor respondents clearly prefer drawing on local sources, as evidenced by the fact that for each of the top five most used sources, an AusE source was used by most respondents. This is especially true for dictionaries, style guides and usage guides. Given that the codification of AusE was attained in the second half of the twentieth century through the publication of AusE dictionaries, style guides and usage guides, the data on the respondents' use of norm-providing sources provides evidence of a clear endonormative orientation in the sources used by editors of AusE. Some exonormative sources are still used, especially AmE and BrE sources, but there appears to be a pattern linked to their use. AmE sources feature strongly among the source type categories associated with lexis and style (such as dictionaries and style guides), while BrE sources feature more strongly among sources associated with grammar and usage. Two distinct reasons most likely account for this pattern. Since most of the Australian editors have experience in editing academic texts, and since academic editing is typically limited to interventions of lexis and style only, the prevalence of AmE dictionaries and style guides is in all likelihood linked to the publishers' or clients' preference for AmE norms or the fact that these dictionaries are accessible for free, rather than an explicit decision on the part of the editor to draw on these sources. The prevalence of BrE sources that deal with issues of grammar and usage seems also to be motivated not by an explicit attempt among Australian editors to draw on sources reflecting an external norm, but rather to be motivated on the one hand by ease of access to the source and on the other by the popularity of the sources. This suggests an important point: Particular norm orientations may gain influence as a result of other forces, such as client preference or cost, conspiring in ways that have little to do with reflecting a preferred norm orientation.

5.2.3.3 South African respondents

In contrast to the extensive codification of AusE, the codification of SAfE is largely limited to the publication of dictionaries. The first dictionary of SAfE, *Africanderisms* (Pettman, 1913), appeared in the early twentieth century; however, the most decisive step towards the codification of SAfE only took place towards the end of the century with the publication of two substantial scholarly dictionaries (Schneider, 2007). Since then, the codification of SAfE has not moved beyond the lexical plane, and while some style guides and usage guides have appeared for SAfE, they are by no means substantial or widely recognised and used by speakers of English in South Africa. It is therefore to be expected that most of the norm-providing sources that editors of SAfE draw on are exonormative. The data in Table 5.21 show this to be the case: of the 528 responses, 70% were for externally-oriented sources and only 22% of responses were for sources that reflect a norm local to the South African context.

BrE sources returned the highest number of responses, accounting for nearly half of all responses (41%) (see Table 5.21). While there are clearly many more responses for sources that reflect a norm external to South Africa, an analysis of the individual orientations shows that SAfE (15%) and bilingual English–

Afrikaans (5%) sources together accounted for the second highest number of responses (20%), while an additional 13 (2%) responses were returned for Afrikaans. AmE sources come in third place, followed by multinational sources and sources of indeterminate orientation. Although there is evidence of emerging endonormativity within SAfE (Kruger & Van Rooy, 2017; Schneider, 2007; Spencer, 2011b; Van Rooy & Terblanche, 2010), there is still a wide-spread perception that BrE is the standard to aspire to, as captured in a chapter titled, *No 'Afringlish', please, we're British* in Linnegar (2009, p. 3).

Orientation	Dictionaries	Style guides	Usage guides	Editing and writing sources	Language reference works	Total
American English	39	41	5	2	5	92
American and British English	4	-	-	-	3	7
Multinational	35	-	-	-	9	44
Australian English	-	3	-	6	-	9
British English	117	36	18	26	17	214
English and a foreign language	4	-	-	-	-	4
Indeterminate	14	21	-	2	3	40
Afrikaans	5	8	-	-	-	13
South African English	23	23	14	17	-	77
South African English and Afrikaans	14	-	-	-	14	28
Total	255	132	37	53	51	528
External	199	80	23	34	34	370
Internal (incl. Afrikaans)	42	31	14	17	14	118
Indeterminate	14	21	0	2	3	40

Table 5.21: Number of responses per norm orientation for the top five source types: South Africa

Dictionaries were by far the most popular norm-providing sources among the South African editor respondents, with 95% of the respondents indicating that they make use of dictionaries. As pointed out in Section 5.2.3.1, two of the three editors who do not make use of dictionaries do make use of *Pharos Online*, an online language reference that includes a number of general and specialised dictionaries. It is therefore highly likely that these two editors do make use of dictionaries, but that they access these via an online language reference site.

	Dictionary	Print or online	Orientation	Percentage of respondents (n = 55)
1.	<i>Oxford English Dictionary (OED)</i> http://www.oed.com/	Online	Multinational	64
2.	<i>Oxford Dictionaries</i> https://en.oxforddictionaries.com/definition/	Online	British English	51
3.	<i>Merriam-Webster Dictionary and Thesaurus</i> http://www.merriam-webster.com/	Online	American English	40
4.	<i>Collins English Dictionary</i> http://www.collinsdictionary.com/dictionary/english	Online	British English	36
5.	<i>New Oxford Dictionary for Writers and Editors</i> Oxford: Oxford University Press.	Print	British English	29

6.	<i>Cambridge Free English Dictionary and Thesaurus</i> http://dictionary.cambridge.org/dictionary/english/	Online	British English	22
7.	<i>South African Concise Oxford Dictionary/ Concise South African Oxford Dictionary</i> Cape Town: Oxford University Press.	Print	South African English	18
8.	<i>Wiktionary</i> https://en.wiktionary.org/wiki/Wiktionary:Main_Page	Online	Indeterminate	16
9.	<i>Macmillan Dictionary</i> http://www.macmillandictionary.com/	Online	British English	15
10.	<i>The Free Dictionary by Farlex</i> www.thefreedictionary.com	Online	American English	13
11.	<i>Concise Oxford English Dictionary</i> Oxford: Oxford University Press.	Print	British English	13
12.	<i>Pharos Afrikaans–English English–Afrikaans Dictionary</i> Cape Town: Pharos.	Print	South African English – Afrikaans	11

Table 5.22: Most frequently selected dictionaries by the South African respondents

Table 5.22 lists the 12 dictionaries that were selected by at least 10% of the respondents. The *OED Online* is the most popular dictionary and was selected by the majority of the respondents. As mentioned previously, the online version of the OED is a multinational dictionary of the English language that reflects the different regions in which English is used, including South Africa, without modelling its entries on a specific reference variety. The popularity of the *OED Online* among the South African editors is probably due to its status as the most authoritative and comprehensive dictionary of the English language. Editors in South Africa also seem to place a higher premium on Oxford University Press publications, as demonstrated by the fact that 5 of the top 12 dictionaries are published by this university press.

The popularity of Oxford dictionaries is further demonstrated in the fact that the online version of *Oxford Dictionaries* was the second most popular dictionary for this sample. The content in the online version of *Oxford Dictionaries* differs from that of the *OED Online* because it focuses on current English usage and practice, and includes a short grammar section that addresses spelling, punctuation and usage. While the dictionary is oriented towards the British norm, it does include access to a US dictionary, but explicitly marks this dictionary as US, clearly setting the BrE version as the standard reference from which the US version deviates. The preference for *Oxford Dictionaries* is likely bolstered by the fact that the members of the Professional Editors' Guild (PEG) (one of the sampling frames from which the sample of South African editors was drawn) can access the premium sources on *Oxford Dictionaries* for free via the organisation's membership portal. This is the only language source that PEG makes available to its members. The endorsement of this dictionary by PEG as well as the fact that access to it is free to PEG's members in all likelihood adds to the status of the dictionary among editors of English texts in South Africa. As with the case of dictionaries for Australian editors, considerations of cost might also be at play here, which provides further evidence to suggest that factors beyond conscious normative choices might conspire to entrench certain norm orientations.

BrE dictionaries also took fourth, fifth and sixth place while only 2 AmE dictionaries, *Merriam-Webster's Dictionary and Thesaurus* and *The Free Dictionary by Farlex*, both of which are free online dictionaries,

were among the 12 most used dictionaries, coming in at third and tenth place respectively. Two South African dictionaries, the two editions of the *South African Concise Oxford Dictionary* and a bilingual SAfE–Afrikaans dictionary published by Pharos, a well-known publisher of language reference works in English and Afrikaans in South Africa, were listed by more than 10% of the respondents. The most comprehensive and recent SAfE dictionary, the *South African Concise Oxford Dictionary*, was used by only 18% of respondents, of which half make use of the first edition published in 2002 and the other half the second edition published in 2010.

The inclusion of a bilingual English–Afrikaans dictionary, *Pharos Afrikaans–English English–Afrikaans Dictionary*, in the top 12 dictionaries, and the fact that 10 different bilingual dictionaries were listed by the South African editor respondents, is probably due to the recommendation by Janse van Rensburg (2016) to PEG members that a good bilingual dictionary is indispensable for editing and proofreading work, and that the *Pharos Afrikaans–English English–Afrikaans* dictionary is the best source available. The high level of English–Afrikaans bilingualism amongst the South African editors (see Section 5.2.2) is an important reason for the wide-spread use of bilingual English–Afrikaans dictionaries and points to the complex interaction between English and Afrikaans within the South African linguistic landscape. Not only are South African editors bilinguals (influencing their covert norms), but they draw on bilingual sources, which opens up opportunities for the influence of overt Afrikaans norms on the editorial choices of editors of English texts in South Africa, not only in terms of the potential for cross-linguistic influence in the editors’ bilingual representations of English and Afrikaans (Kruger & Van Rooy, 2016), but also in terms of the influence of Afrikaans norms as set out in Afrikaans and bilingual norm-providing sources. Overall it seems that most of the South African editor respondents draw on dictionaries reflecting a norm external to SAfE, and favour BrE input far more than AmE.

Style guides were the second most popular source among the South African editor respondents, and were selected by 52 of the 58 respondents in this sample. There are far fewer individual responses for style guides than there are for dictionaries, suggesting that South African editors draw on a narrower range of style guides than they do dictionaries. Unlike Australia, where one style guide is widely recognised as the standard for AusE publications, no single style guide enjoys such status in South Africa. Six different style guides were selected by at least 10% of the South African editors, of which most were BrE and AmE style guides and only one was specifically South African, but was for Afrikaans and not English.

	Style guide	Type	Orientation	Percentage of respondents (n = 52)
1.	<i>In-house Style Guide</i> (19, unspecified: 23, specified)	In-house	Indeterminate	73
2.	<i>Hart's Rules: The Oxford Style Guide/The New Oxford Style Manual</i> Oxford: Oxford University Press	Book publisher	British English	37

3.	<i>The Chicago Manual of Style</i> Chicago: University of Chicago Press.	Book publisher	American English	25
4.	<i>The Elements of Style</i> New York: Harcourt/Macmillian/Pearson Longman.	Popular	American English	17
5.	<i>Publication Manual of the American Psychological Association</i> New York: American Psychological Association.	Association	American English	13
6.	<i>Afrikaanse Woordelys en Spelreëls</i> Cape Town: Pharos.	Association	Afrikaans	10

Table 5.23: Most frequently selected style guides by the South African respondents

The majority of the South African editors indicated that they make use of in-house style guides (of which 19 were unspecified and 23 were specified). Since many of the in-house style guides were unspecified, it is not possible to determine if these they reflect an endonormative or an exonormative orientation, and their orientation is therefore indeterminate. Since these style guides emanate from the bilingual South African publishing environment, it is likely that they reflect the kinds of norms that are set for English and Afrikaans publications in South Africa, and therefore in all likelihood reflect an internal, South African, norm (that would also, to variable extents, show traces of external norms which may have informed their development).

The second most popular style guide among the South African editor respondents is the BrE *Hart's Rules* series of style guides, which have been published under different titles since 1904. The popularity of this style guide is not surprising, particularly given its long publication history, status as an authoritative guide, the recommendation for its use by PEG (Linnegar, 2016) and the South African Department of Government Communication and Information System's Editorial Style Guide (GCIS, 2013), and the high status of Oxford University Press publications among South African editors. Three AmE style guides come in in third, fourth and fifth place: *The Chicago Manual of Style*, William Strunk's popular *Elements of Style* and the *Publication Manual of the American Psychological Association*.

The sixth most popular style guide used by the respondents is an Afrikaans text that provides rules for the spelling of Afrikaans words, but includes information on orthographic aspects of language use, as well as guidelines for writing in Afrikaans. Thus the content of the source extends beyond spelling and contains discussion of other areas of language that might be transportable to the editing of English texts. It is surprising that such a text would be used by so many editors of English texts, but given the high level of bilingualism amongst the South African editor respondents, and the fact that many work in both English and Afrikaans, it is likely that these editors use the sources from one language to help guide the editorial choices they make when editing another (as is also demonstrated in the popularity of English–Afrikaans bilingual dictionaries among the respondents). A further contributing factor may be the fact that in comparison to SAFE, many more codification instruments for Afrikaans exist, and it could be argued that coupled with the English–Afrikaans bilingualism of the respondents, editors of English in South Africa do make use of sources internal to the South African context, even if they are in a language other than English. Furthermore, the *Afrikaanse Woordelys en Spelreëls* (The Afrikaans Wordlist and Spelling Rules) is a

unique publication that holds a special place in the codification of and the development of Afrikaans in South Africa, and is widely considered to be the most authoritative source for the rules of Afrikaans spelling. This, together with the findings on the dictionaries used by South African editors provides further evidence of the potential influence of Afrikaans (and especially Afrikaans norm-providing sources) on the editing of English texts in South Africa.

Language reference works were the third most popular type of source used by the South African editors, 37 of whom indicated that they make use of this type of source in the course of their work. Three sources were used by at least 10% of the respondents (see Table 5.24).

	Language reference work	Type	Orientation	Percentage of respondents (n = 37)
1.	<i>Pharos Online</i> https://www.pharosonline.co.za/home	Online language reference	South African English and Afrikaans	38
2.	<i>The Oxford A-Z of Grammar & Punctuation</i> Oxford: Oxford University Press.	Print book	British English	30
3.	<i>Oxford Reference Online</i> www.oxfordreference.com	Online language reference	Multinational	24

Table 5.24: Most frequently selected language reference works by the South African respondents

The most popular language reference work is *Pharos Online*, which is accessible via subscription and includes a number of monolingual and bilingual general and specialised dictionaries, word lists and thesauruses. The site also includes access to the electronic version of *The (Essential) Companion for Writers and Editors*, a “ready reference to information commonly required, errors commonly committed, conventions commonly accepted, and problems and language idiosyncrasies commonly encountered by writers and editors” (Hendry, 2011, p. 7), as well as access to the electronic version of the *Afrikaanse Woordelys en Spelreëls*. As with dictionaries and style guides, it again seems that South African editors of English texts consult English–Afrikaans bilingual or Afrikaans monolingual sources in the course of their work, reaffirming the finding that these sources are popular among the respondents and providing further evidence of the potential influence of Afrikaans and Afrikaans norm-providing sources on the editing of English texts in South Africa.

The second and third most popular sources are both Oxford University Press publications. *The Oxford A–Z of Grammar & Punctuation* is a print book oriented towards the British norm that sets out a discussion of grammar, various aspects of usage and punctuation. The fourth most popular source is the *Oxford Reference Online*, a multinational repository of Oxford publications including general and subject-specific dictionaries, usage guides, companions and encyclopaedias.

Editing and writing sources were the fourth most popular source type among the respondents, and 27 of the South African editors indicated that they make use of these kinds of sources (see Table 5.25). Of the 27 respondents, the majority indicated that they make use of *Butcher’s Copy-editing*. The *Oxford Guide to*

Plain English and *Pharos The Write Stuff: The Style Guide with a Difference* were both used by 26% of the respondents each, while *Working Words* was used by 11% of the respondents. The popularity of the plain English guide, which is an Oxford University Press publication with a BrE orientation, is most likely due to the fact that the use of plain language is sanctioned in a number of acts in South African legislation (Cornelius, 2015). Because of this, many organisations outsource the plain language editing of their documents to South African language practitioners, including editors.

Only one South African editing and writing source was used by more than 10% of the respondents and one Australian source was listed. Despite its title, the *Pharos The Write Stuff: The Style Guide with a Difference* is not a style guide but rather, as the authors claim, a collection of items one would find “in a dictionary, grammar book and style guide – all in one book” (Dykman, Geldenhuys, & Viljoen-Smook, 2008, p. 5). The guide, whose content is structured alphabetically, covers a range of topics deemed important by the authors, including subject verb agreement, easily confused words and a one-page discussion on AmE and BrE (with hardly any mention of SAfE) that positions BrE as Standard English and the standard to which South African editors should aim.

	Editing and writing source	Type	Orientation	Percentage of respondents (n = 27)
1.	<i>Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors, and Proofreaders</i> Cambridge: Cambridge University Press.	Handbook	British English	59
2.	<i>Oxford Guide to Plain English</i> Oxford: Oxford University Press.	Handbook	British English	30
3.	<i>Pharos The Write Stuff: The Style Guide with a Difference</i> Cape Town: Pharos.	Handbook	South African English	30
4.	<i>Working words</i> Manuka, ACT: Canberra Society of Editors	Handbook	Australian English	11

Table 5.25: Most frequently selected editing and writing sources by the South African respondents

Usage guides are the last of the top five source types listed by the South African respondents, and only 23 of the total number of 58 respondents indicated that they make use of usage guides (see Table 5.26). There may be several reasons for this: firstly, very few usage guides for SAfE exist, and those that do exist mostly rely on the authors' personal impressions of language usage, drawing on anecdotal examples to discuss usage issues. As a result of the limited number of SAfE usage guides, it may be that South African editors are turning to other sources for usage advice, such as online language reference works, in order to access such information, which may explain why online language reference works were the third most popular type of source among the South African editor respondents.

	Usage guide	Orientation	Percentage of respondents (n = 23)
1.	<i>Engleish, our Engleish: Common Errors in South African English and How to Solve Them</i> Cape Town: Pharos.	South African English	61
2.	<i>Fowler's Modern English Usage</i> Oxford: Oxford University Press	British English	43
3.	<i>Garner's Modern English Usage</i> New York: Oxford University Press	American English	22

Table 5.26: Most frequently selected usage guides by the South African respondents

Nevertheless, the data from the South African respondents show that three usage guides are used by more than five respondents. The most widely used of these is a South African usage guide, *Engleish, our Engleish* (Linnegar, 2009), that is used by 61% of the respondents. The guide, which the author describes as a “slim vade mecum”, provides a list of “incorrect usages . . . [compiled] over several decades in manuscripts, newspapers and magazines” (Linnegar, 2009, p. xiii). The author of the guide is a well-known editor and editing trainer in South African freelance editing circles. He has served as chairperson of PEG, has co-authored a textbook on editing for South African editors and is a regular contributor to the organisation’s chat group (a platform where members can raise and answer editing- and language-related questions). The popularity of this source among the editors is therefore most likely due to the author’s status as an authority on English in the South African editing community.

The second most popular usage guide is *Fowler’s*, which is used by 43% of the respondents, all of whom indicated that they make use of a print edition. None of the respondents use the most recent Butterfield’s edition of the guide; 33% make use of Burchfield’s 2004 edition and 20% of the 1996 edition. A total of 10% of the respondents use a reprinted first edition from 1954, while 20% make use of Gower’s second edition. The inclusion of *Fowler’s* among the most used usage guides is not all that surprising, particularly given its status as the most influential usage guide of the twentieth century (Lukač, 2017).

Garner’s Modern English Usage, which like *Fowler’s* is an Oxford University Press publication, is the third most popular usage guide and is used by 22% of the respondents. All of these respondents indicated that they make use of the most recent 2016 edition, which despite the inclusion of BrE data, is still very much modelled on the AmE norm. One interesting point of difference that arises regarding the second and third most popular usage guides is that the latest edition of *Fowler’s* is not represented in the South African sample at all, while only the latest edition of the AmE *Garner’s* is. This may point to a changing preference for the AmE usage guide among the South African editor respondents.

Overall, it seems that the few South African editors who do make use of usage guides in the course of their work draw on a range of guides that reflect South African, British, and to a lesser extent AmE norms. The South African editors therefore draw on a much more diffuse range of sources in terms of normative orientation, compared to their Australian counterparts who show a much more focused preference for AusE.

With regard to the remaining four source type categories, the data show that very few South African editors make use of these sources (see Table 5.27). Glossaries are used by only 5% of the South African respondents, while 8% make use of a thesaurus and 14% consult various online and print sources. Just over a fifth of the South African respondents (21%) indicated that they make use of people, which included consulting colleagues, subject experts and other language practitioners (via four language practitioner organisations in South Africa). The most popular of these is the PEG chat group, an email chat group in which the members of PEG raise and answer various questions related to editing. The use of networks of people as a tool to seek out advice on editing decisions might be linked to the absence of clear codification instruments and more diffuse norms for South African editors.

South Africa					
Orientation	Glossaries	Thesauruses	People	Various online and print sources	Total
American English	1	1	-	-	2
American and British English	-	1	-	-	1
British English	-	4	-	2	6
Indeterminate	-	-	7	11	18
Afrikaans	-	1		-	1
South African English – Afrikaans	2	-	9	-	11
Total	3	7	16	13	39

Table 5.27: Responses for glossaries, thesauruses, people and various online and print sources

The South Africa editor respondents clearly draw on a diffuse set of overt norm-providing sources: with a strong preference for sources reflecting the external BrE norm, coupled with a handful of SAfE and some monolingual AfrE and bilingual Afrikaans–English sources. Further information regarding the findings for the norm-providing sources used by editors of SAfE is provided in Section 5.2.3.4, which compares the findings for the two countries.

5.2.3.4 Comparison of norm-providing sources

The data on the sources used by the Australian and South African editors need to be interpreted against the information on their demographic and linguistic profiles because the two sets of data provide information on editors' covert and overt norms. The typical demographic profile of the editors shows that across both samples editors are usually older women who are highly educated. Given what is known about the role of these factors in influencing language behaviour, it may be concluded that editors are most likely conservative language users with a preference for standard language usage. However, the fact that most of the respondents are women also raises the possibility that even though they might orient their linguistic choices towards standard forms, they are also quicker to take up features associated with ongoing language change. Because no single social factor is entirely deterministic of language behaviour, this gender paradox

most likely interacts with the other factors, which together with the normative nature of editorial work might pull the linguistic choices of the editors towards more conservative linguistic choices.

In addition to near similar profiles in terms of age, gender and education, the respondents are also similar in their working backgrounds and how they view their editorial roles. Most of the editor respondents work on a freelance basis and consequently have experience in a large range of text types. The effect of this is a more diffuse normative environment in which the editors work on many different kinds of texts and for multiple clients and organisations. This diverse range of text types and the accompanying diversity of stylistic and textual conventions associated with each, might account for the respondents' views of the editorial role as comprising both normative and communicative dimensions with each assigned somewhat different weight.

Among the Australian respondents, the normative dimension of editorial work appears to be prioritised slightly more, compared to the South African respondents who expressed a much more balanced view in terms of the importance of each dimension. The respondents' motivations for these views are telling, and speak to the different contexts in which their editorial work is carried out. For all respondents, target audience expectations and text type conventions are crucial, but within the particular context in which they work, the weighting of their views of the editorial role is cast somewhat differently. For the South African editors, the fact that many readers of the texts they edit are second-language users of English seems to result in a view in which both dimensions of the editorial role are considered equally important, yet for the Australian editors the same consideration seems to bolster the normative dimension. One reason for this might be the fact that within the South African context, linguistic norms are much more diffuse (as evidenced by the kinds of norm-providing sources these editors draw on). In the absence of clear norms for the South African context, South African editors might turn to the communicative dimension to ensure that texts are optimised for their audiences, opening up the opportunity for their own usage preferences (covert norms) to have an influence. In contrast, the norms for AusE are clearer and easily accessible in the range of codification instruments available for the variety. As a result, there is a clear set of norms to aim for and against which a text can be assessed, providing Australian editors with the opportunity to take a more focused normative approach to their editorial work but also providing less space for personal preferences (covert norms) to play a role.

This contrast between a more focused versus a more diffuse approach is also evident in the language backgrounds of the editors from the two samples, where (unlike age, gender, education and work experience profiles) the profiles of the respondents are quite different. The Australian editors are all first or home language users of English, and although a fair number are multilingual, these additional languages do not extend to their editing work, and nearly all edit in English only. This homogeneity in language profiles together with a much more clearly codified variety of English (in other words, a larger degree of convergence between overt and covert norms) means that for the Australian editors, their linguistic

representations are most likely more reflective of the Australian norm, and in all likelihood results in linguistic editorial choices that converge on the Australian norm.

The South African respondents' language profiles are much more complex, with a roughly equal distribution of English and Afrikaans first or home language users, most of whom edit in both languages. This high degree of bilingualism and the accompanying bilingual linguistic representations, together with the much more diffuse normative environment is characteristic of the earlier stage of the variety along the Dynamic Model (in comparison to AusE) (Spencer, 2011b), and most likely results in linguistic editorial choices that are more divergent. In other words, far more diffuse usage contexts in South Africa feed in to editors' covert norms. Furthermore, it is striking that only two strands of English are represented in the language profiles of the South African editors, and that none of the editors represent the other IDG strand, BSAfE. The linguistic gatekeeping role in SAfE, also for texts produced by BSAfE users, is thus fulfilled by users from other strands (which is not the case for AfrE and WSAfE).

This contrast between a more homogenous and a more diffuse normative environment is also evident in the kinds of norm-providing sources that the respondents from the two samples draw on as well as the orientation of these sources. In other words, the same distinction between more focused and more diffuse norms evident at the level of covert norms is also evident at the level of overt norms. For the Australian respondents, there is a strong preference for endonormative sources, as evidenced by the most frequently used sources across the top five categories. Dictionaries and style guides, the two types of sources most strongly associated with editing work (Cameron, 2012), are the two source types used by most of the Australian editors and offer the widest range of choice of AusE texts. However, dictionaries and style guides are also the two source type categories in which AmE sources are extremely popular, which suggests that even though most editors of AusE texts draw on key endonormative dictionaries and style guides, many supplement these with AmE dictionaries and style guides.

The number of Australian respondents who indicated that they use usage guides, editing and writing sources and language reference works demonstrates that these sources are somewhat less important for editing work, compared to dictionaries and style guides. Australian editors seem to be aware of and draw on the usage guides developed for their own variety of English, but these are used by less than half of the respondents, and compete with BrE usage guides. Editing and writing sources are used by still fewer Australian editors; however, when they are used the preference is clearly for sources written for the Australian editing and writing industry. Lastly, language reference works are the least popular of the top five sources, with less than half of the respondents indicating that they make use of these sources. As with the dictionaries and style guides, the language reference works used by the Australian editors are oriented towards AusE and AmE norms.

As is to be expected, the data from the South African respondents are quite different from the data from the Australian respondents, and there is clearly a preference for exonormative sources among the South African

editors. This is to a considerable extent due to the unavailability of SAfE sources for most of the source type categories, and the fact that in the absence of codification instruments for SAfE, the appeal of BrE sources is strong (de Klerk, 1996, p. 10). Across the top five source type categories, the South African editors consistently demonstrate a preference for BrE sources, especially those published by Oxford University Press. This is especially true for dictionaries and style guides, arguably the most important sources in the ‘editorial armoury’ (Linnegar, 2016). There is therefore a clear discrepancy between the overt and covert norms of editors of SAfE, with the overt norm orientation towards BrE even though these editors are not users of this variety.

In addition to a preference for BrE sources, the South African editors also have a preference for bilingual English–Afrikaans and monolingual Afrikaans sources, many of which are published by the same South African publisher. Therefore, an unexpected local influence arises in the overt norms that editors of English in South Africa invoke. As discussed, the use of these sources is most likely due to the high level of English–Afrikaans bilingualism among the South African editor respondents, as reflected in the data on their language profiles, but it may also be an indication that in the absence of explicit SAfE norm-providing sources, these bilingual editors utilise their high levels of bilingual proficiency and draw on sources written within the South African context (even if those sources are in or include another language). This opens up their editorial decisions on the English texts they edit to potential influence from Afrikaans, not only in terms of the influence of overt Afrikaans norms, but also in terms of the bilingual linguistic representations of these editors.

The South African editors also differ from their Australian counterparts in their preference for usage guides, with far more preferring to make use of language reference works. Lukač (2017) argues that the ease with which people can access online language reference works might have an influence on the use of print sources. Elsewhere, she finds evidence that highly educated language professionals, such as editors, prefer consulting online rather than print sources (Lukač, 2016). She argues that online language blogs and websites might be becoming the preferred point of guidance on usage issues (Lukač, 2016, 2017). It may therefore be that in the absence of substantial, scholarly usage guides for SAfE, editors turn to these online language reference sites to access information on usage issues.

The points that Lukač (2017) raise also tie in with an further finding regarding the use of norm-providing sources, and especially the use of dictionaries. In both the Australian and the South African data, dictionaries that can be accessed for free were some of the most popular sources among the respondents to the questionnaire. This raises that possibility that the selection of norm-providing sources might be less guided by explicit consideration of appropriate norms, than they are by questions of cost. It is noteworthy that traces of preferred norm-orientation may still be found across the two samples: Australian editors make use of free AmE dictionaries, while South African editors make use of BrE dictionaries.

As discussed in Chapter 2, the kinds of norm-providing sources that editors draw on provides information on the kinds of overt norms that they invoke in their editing work, while the information on editors' linguistic backgrounds provides some information on the kinds of usage events they are exposed to and their grammatical representations. But this does not provide any information on the extent to which the overt norms or covert norms find their way into the actual editorial choices that editors make, nor how these overt norms might interact with covert emergent norms. To assess this, it is necessary to investigate the actual changes that editors make, which is the focus of the corpus-based analysis in the next section.

5.3 Presentation and analysis of corpus findings

Research question 2.3 of this study seeks to investigate how editors of English texts in Australia and South Africa respond to the presence of features associated with ongoing language change in the texts that they edit as a way to gauge what role these linguistic gatekeepers play in the dissemination and conventionalisation of these features in the varieties within which they work. As already mentioned, the aim of this question is to determine how overt norms (sanctioned in norm providing sources) interact with covert norms in editorial choices, and how this might help to move a variety towards endonormativity.

This section discusses the findings of the cross-varietal and cross-register analysis of the alternation between the two major adnominal genitive constructions in unedited and edited writing. Sections 5.3.1 and 5.3.2 present the analysis of the alternation between these two constructions in AusE and the different sub-varieties of SAfE, as well as editors' treatment of these genitive constructions across five registers in the two varieties. Since very little is known about genitive alternation within the context of AusE and SAfE and across different published written registers, Section 5.3.1 sets out a distributional analysis of the proportional and normalised frequencies of the two constructions in the edited (and therefore published) written texts for each register in each of the sub-varieties. The analysis in Section 5.3.1 lays the foundation for a contrastive analysis of variation between the two constructions in the two varieties and across the different registers, while Section 5.3.2 adds the role of editorial intervention as an additional factor in influencing genitive alternation. Thereafter, in Section 5.3.3, the focus shifts to include the linguistic factors known to influence the alternation between the two constructions in order to consider how different external/stylistic and linguistic factors and editorial intervention interact in the use of the two variants. The language-external factors considered include the different sub-varieties, the five registers, and the edited and unedited status of the texts, while the linguistic factors considered include the semantic relation between the possessor and possessum, animacy of the possessor, syntactic weight of the possessor and possessum, definiteness of the possessor and final sibilance.

5.3.1 Distributional analysis of genitive constructions in AusE and the sub-varieties of SAfE

The first step in the distributional analysis of the two genitive constructions is to investigate the proportions and normalised frequencies of the two constructions across the sub-varieties represented in the corpus, for

the edited texts. This analysis serves as an important indication of how the genitive constructions are used within the context of written published English in the two varieties. The analysis of the proportional frequencies is presented in Section 5.3.1.1, while the analysis of the normalised frequencies is presented in Section 5.3.1.2.

5.3.1.1 Distributional analysis: Proportional frequency

Table 5.28 presents the overall estimated proportions (see Appendix D for an exposition on the calculation of these proportions and normalised frequencies) of the two constructions in the edited texts of each sub-variety, alongside proportions for AmE and BrE. The proportions for AmE are derived from Heller (2018b), who investigates changes in the proportions of the two constructions across the various written registers in the Brown and Frown corpora. The proportions reported here are for the Frown data in Heller (2018b). The only comparable data for written published registers in BrE is Leech et al. (2009, p. 308).¹⁶ All studies that investigate genitive alternation in contemporary BrE either focus exclusively on written data from reportage, by for example, focusing on the A and B portions of the F-LOB corpus, or focus on BrE more broadly, lumping together the frequencies of the two constructions for both spoken and written language. Leech et al. (2009, p. 308) do report on the frequencies of the constructions in F-LOB; however, their data show that the *s*-genitive is far more frequent in BrE than it is in AmE, which contradicts the findings by other scholars, such as Hinrichs and Szmrecsanyi (2007), Rosenbach (2002) and Szmrecsanyi and Hinrichs (2008), who show that the increase in the *s*-genitive is more advanced in AmE than in BrE. The data for BrE in this table should therefore be interpreted with caution.

	AfrE		BSAfE		WSAfE		AusE		AmE		BrE	
	Count	% of total	Count	% of total	Count	% of total	Count	% of total	Count	% of total	Count	% of total
<i>s</i> -genitive	4 353	18	1 014	12	3 121	27	2 498	30	1 109	43	6 148	57
<i>of</i> -genitive	19 889	82	7 799	88	8 296	73	5 770	70	1 490	57	4 723	43
	24 242		8 813		11 417		8 268		2 599			10 871

Table 5.28: Overall proportional frequencies of the *s*- and *of*-genitives in the sub-varieties, in edited texts.

The two native sub-varieties in the corpus, AusE and WSAfE, have the highest overall proportion of *s*-genitives, followed by AfrE and BSAfE, where the overall proportion of *s*-genitives is somewhat lower. AusE is only slightly more advanced than WSAfE in the higher proportion of *s*-genitives, with AfrE lagging in third place and BSAfE considerably behind the two native sub-varieties in fourth place. This mirrors the findings by Heller et al. (2017b) who suggest that the proportion of *s*-genitives is lower in some non-native

¹⁶ The data from Leech et al. (2009) need to be interpreted with caution since their data are approximations based on a 2% sample of *of*-phrases taken from the LOB and F-LOB corpora and that were analysed manually. The authors note that a similar procedure was not carried out for the *s*-genitives, and instead base their approximation of the number and proportion of *s*-genitives on the assertion by Hinrichs and Szmrecsanyi (2007) that 80% of *s*-genitives are paraphrasable by *of*-genitives. Furthermore, it is not clear what criteria were used for the identification of interchangeable instances, and therefore the figures cited here are not fully comparable with the data on the Australian and South African (sub-)varieties.

varieties, compared to native varieties. Furthermore, AusE, WSAfE, BSAfE and AfrE all have lower overall proportions of the *s*-genitive, compared to AmE, suggesting that the four sub-varieties investigated in this study are not as advanced in taking up the *s*-genitive as AmE; however, of the four varieties investigated, AusE is more advanced than all three of the SAfE sub-varieties and most closely aligned with AmE. To investigate how register might influence the proportional frequencies of the two constructions in the different the sub-varieties, the proportions per register per sub-variety were calculated and are presented in the figures that follow.

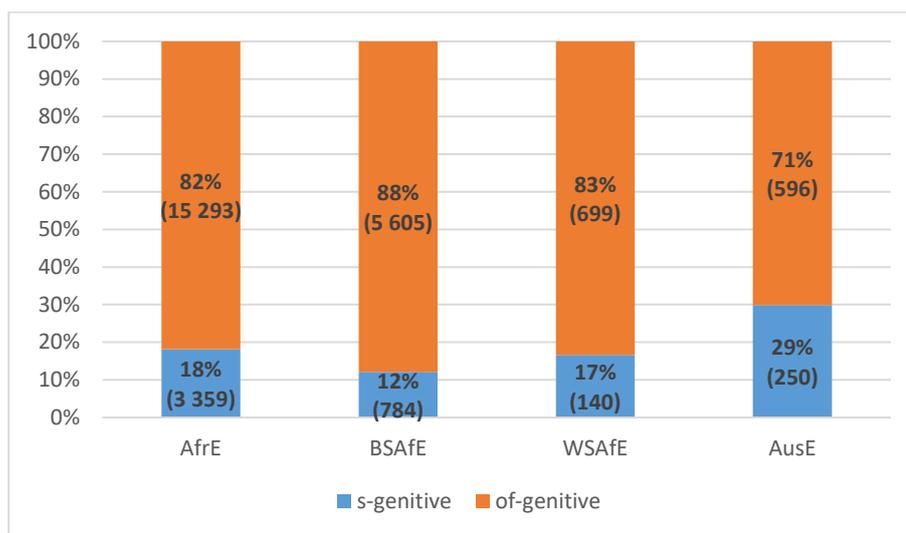


Figure 5.8: Proportional frequencies of *s*- and *of*-genitives in published academic writing across sub-varieties, in edited texts

Figure 5.8 shows that the proportional frequency of the *s*-genitive in written academic AusE is much higher than in AfrE and WSAfE, which have near similar proportional frequencies, and BSAfE, which has the lowest proportional frequency. Written academic AusE therefore appears to be much more advanced in its receptivity to the incoming *s*-genitive than the other three sub-varieties, and particularly BSAfE. Heller (2018a), who investigates genitive alternation in nine first- and second-language varieties of English, reports an overall proportional frequency of *s*-genitives in his academic data of 13,3%.¹⁷ The three South African sub-varieties are therefore (broadly) in line with the varieties investigated by Heller (2018a). However, clearly academic AusE is much more receptive to the less formal construction in a register that is widely considered to be extremely formal (and regarded as increasing in formality (Grafmiller, 2014)). Given that AusE has been shown to be particularly receptive to features associated with colloquialisation (Collins, 2014), the higher proportion of *s*-genitives in this variety suggests that this receptivity extends even to the most formal of the registers.

¹⁷ The varieties investigated by Heller (2018a) are BrE, CanE, HkE, InE, IrE, JamE, NZE, PhiE and SingE, and the proportional frequencies cited here are overall proportions for all varieties. Given the assertion by Heller et al. (2017b) that *s*-genitives have a higher proportional frequency in first- and second-language varieties, it is likely that the proportions for the *s*-genitive in their first-language varieties are higher than the overall proportions reported here.

As mentioned in Section 4.3.2, AfrE and BSAfE are not represented in the creative register due to difficulties in obtaining unedited texts for these two sub-varieties. Figure 5.9, which reflects the proportional frequencies for the two genitive constructions for WSAfE and AusE in the creative register, shows that while the *of*-genitive has the higher proportional frequency of the two constructions in both varieties, the proportional frequency of the *s*-genitive is much higher in creative writing than it is in academic writing. Furthermore, the proportional frequencies for both constructions are identical in the two sub-varieties, suggesting that the creative writing of the two native sub-varieties investigated in this study is similarly receptive to the incoming *s*-genitive. The proportions for the *s*-genitive in the AusE and WSAfE data are somewhat lower than the proportion reported by Heller (2018a), 48,1%, suggesting that AusE and WSAfE are slightly behind the varieties represented in Heller's data.

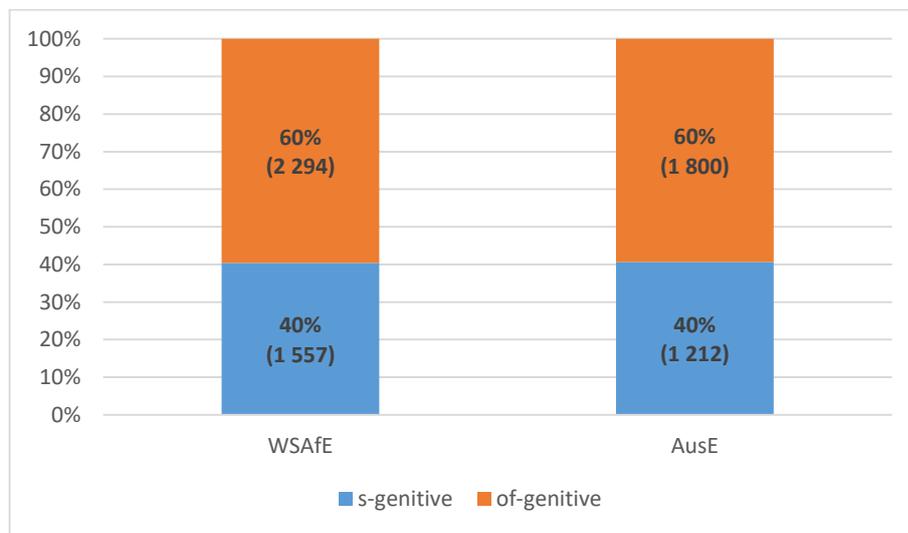


Figure 5.9: Proportional frequencies of *s*- and *of*-genitives in published creative writing in WSAfE and AusE, in edited texts

Instructional writing appears to be the most conservative register in terms of its openness to the *s*-genitive for all four sub-varieties. Near similar proportional frequencies are found for AfrE, WSAfE and AusE (see Figure 5.10) – all of which are very close to the proportions reported by Heller (2018a) for the same register (*of*-genitives, 79,5%; *s*-genitives, 20,5%). With regard to AusE, the trend toward colloquialisation noted for the proportional frequencies of academic editing, does not appear to have as strong an impact on AusE instructional writing. Furthermore, the extremely low proportional frequency for the *s*-genitive in BSAfE instructional writing suggests that BSAfE authors have a near categorical preference for the analytical *of*-genitive in instructional writing.

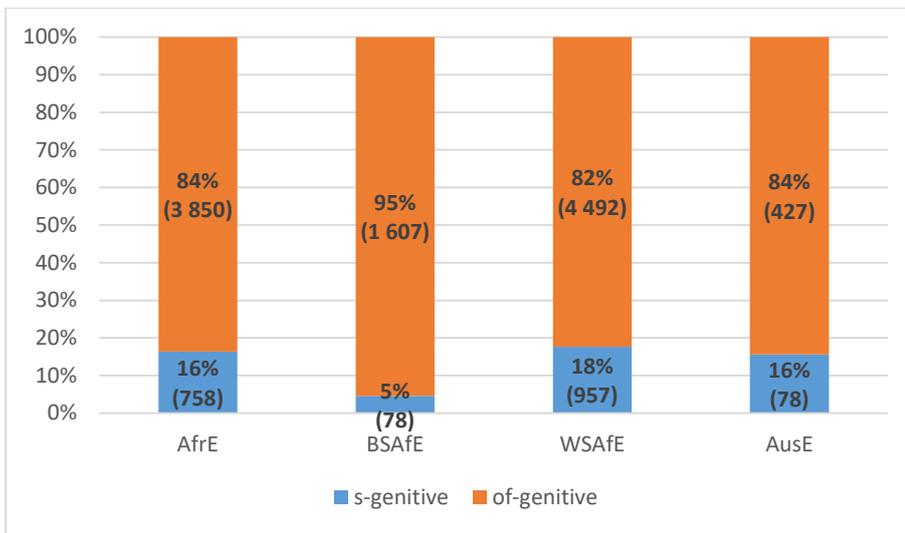


Figure 5.10: Proportional frequencies of *s-* and *of-*genitives in published instructional writing across sub-varieties, in edited texts

In the popular register, WSAfE has the highest proportion of *s-*genitives of the four sub-varieties, with an *s-*genitive proportion that is only slightly lower than the same construction’s proportion for the same sub-variety in the creative register. AfrE and AusE have near identical proportional frequencies, though the proportion of *s-*genitives for these two sub-varieties is considerably lower than for WSAfE. BSAfE is again the sub-variety with the lowest proportional frequency of *s-*genitives, providing more evidence that this sub-variety lags behind the others in terms of its openness to this particular construction. Heller (2018a) reports an overall proportional frequency of the *s-*genitive in his popular data as 31,5%, suggesting that popular WSAfE is quite advanced in its openness to the increasing frequency of the *s-*genitive, and that AfrE and AusE, and particularly BSAfE, are, by comparison, less advanced.

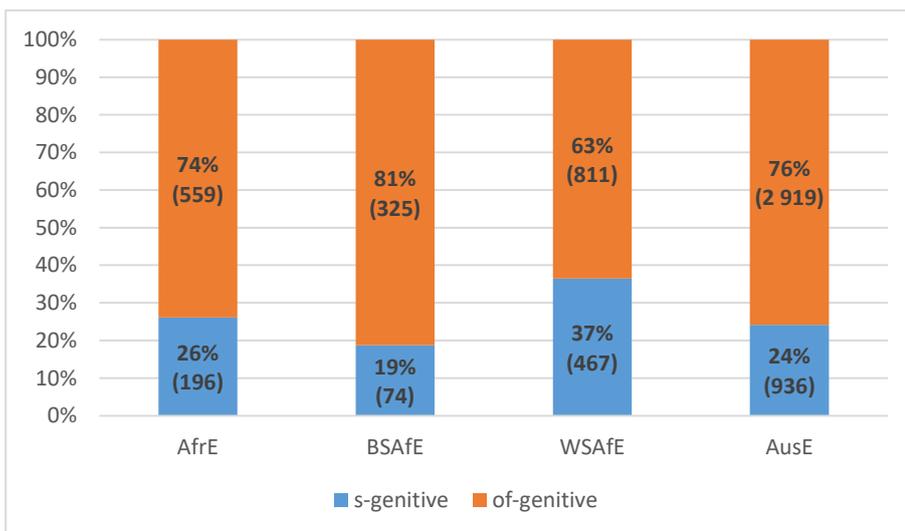


Figure 5.11: Proportional frequencies of *s-* and *of-*genitives in published popular writing across sub-varieties, in edited texts

As mentioned in Section 4.3.2, no texts authored by WSAfE users were collected for the reportage register. In addition, the reportage register in the AusE corpus is composed of one text of 4 481 words. The data for the reportage register therefore do not provide a representative picture of the sub-varieties under investigation, and should be interpreted with caution. In terms of the South African sub-varieties represented, it is interesting to note the very low proportional frequency of the *s*-genitive in the AfrE and BSAfE data, particularly given the generally held view that reportage is the most advanced of the written published registers in taking up the *s*-genitive. Heller (2018a) reports an overall proportional frequency of 46,7% for the *s*-genitive in his data, which is very similar to AusE: clearly the two non-native South African sub-varieties are much less advanced than AusE and the varieties in Heller's data.

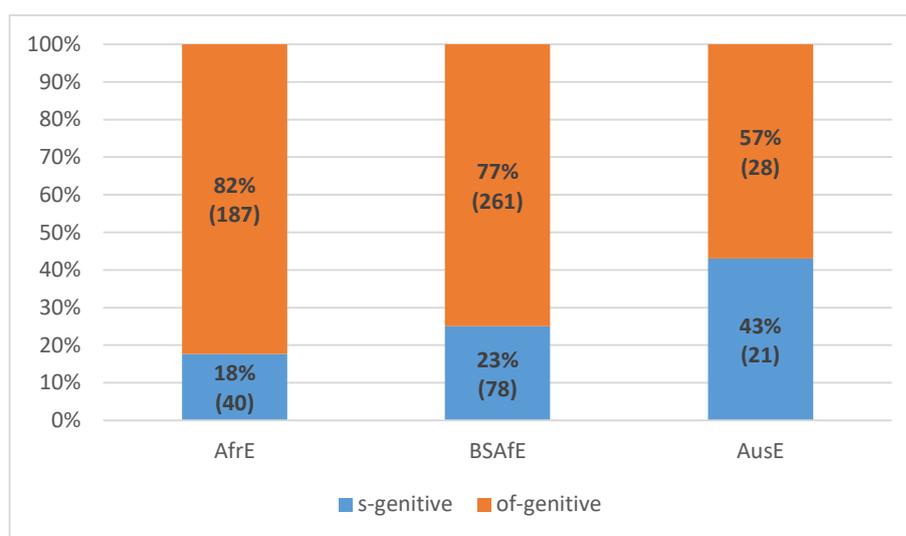


Figure 5.12: Proportional frequencies of *s*- and *of*-genitives in reportage across sub-varieties, in edited texts

The general pattern that emerges from the analysis of the proportional frequencies of the two constructions across sub-varieties and the different registers, and bearing in mind that there are no data for AfrE and BSAfE in the creative register and no data for WSAfE in the reportage register, provide an indication of differences in the preference for the *s*-genitive across sub-varieties and registers. These differences appear to pattern along native and non-native variety lines, and the formality of the registers, and largely support the findings of existing studies (Grafmiller, 2014; Heller, 2018b; Heller et al., 2017b).

In terms of register preferences, the *of*-genitive is clearly preferred in the instructional and academic registers. Although the proportional frequency of the *of*-genitive is higher in all five registers, there is a stronger preference for the *s*-genitive in the creative register, and for the two native sub-varieties in the popular register, compared to the other three registers. Somewhat unexpectedly, the proportion of *s*-genitives in reportage is much lower in the SAfE varieties represented in the corpus, compared to AusE.

One of the two native varieties (WSAfE and AusE) consistently has the highest proportion of *s*-genitives in all five registers, with the other native variety and AfrE clustering together and BSAfE lagging behind. This suggests that the two native varieties are the more advanced sub-varieties in taking up the incoming *s*-

genitive. Furthermore, while WSAfE and AusE have near similar proportions in the creative and instructional registers, this clustering of the two native varieties does not continue through the academic and popular registers, where AusE has a much higher proportion of *s*-genitives in academic writing and WSAfE has a much higher proportion of *s*-genitives in popular writing.

AfrE appears to cluster with one of the two native varieties in the academic, instructional and popular registers, but it does so with the native variety that has the second highest proportion of *s*-genitives in these three registers. The fact that with the exception of the reportage register, AfrE clusters with one of the two native varieties, suggests that of the two non-native varieties, AfrE is more advanced in taking up the incoming form. A contrastive study of the choice of genitive constructions between WSAfE and AfrE users by Rosenbach (2017) provides an explanation for why this may be. Rosenbach (2017) finds evidence that AfrE users use the *s*-genitive in contexts where WSAfE users are more likely to use the *of*-genitive, and attributes this to a transfer effect of the structural parallelism between the prenominal and postnominal realised genitive constructions in English and Afrikaans, and the influence of a weaker animacy constraint in Afrikaans that is carried over to the genitive choices of AfrE users (see Section 3.5 for a more detailed discussion).

BSAfE consistently has the lowest proportion of *s*-genitives, except for in the reportage register, where the proportion is slightly higher than the proportion for AfrE. BSAfE is therefore the least advanced of the four sub-varieties in taking up the incoming *s*-genitive. One reason for this may be due to an effect of the substrate on the choices of these non-native English users: the South-East Bantu languages do not have available a genitive construction in which the possessor is realised prenominal (see Section 3.5 for a more detailed discussion), and so it may be that in BSAfE writing, there is a substrate effect of the author's first language on their non-native English writing that pulls the alternation between the two constructions closer to the more conservative variant. Furthermore, BSAfE is least advanced of the four varieties along the stages of the Dynamic Model (see Section 3.2.2), and this might also account for the higher preference for the *of*-genitive in this sub-variety. The higher proportion of *of*-genitives in the BSAfE writing might thus also be due to the fact that BSAfE is a non-native variety in which language users prefer the more analytic and transparent *of*-genitive construction (Heller et al., 2017b); this does not appear to be the case for the other non-native variety, AfrE. AfrE is much more advanced along the stages of the Dynamic Model, compared to BSAfE, and this, coupled with the possible transfer effect of a lower animacy constraint in Afrikaans, in all likelihood boosts the proportion of *s*-genitives in this second language sub-variety.

5.3.1.2 Distributional analysis: Normalised frequency

To further explore the differences across the sub-varieties and registers, the normalised frequencies (per 1 000 words) of each interchangeable construction in each sub-variety and across registers were calculated (see Table 5.29). As explained in Section 4.3.2, data were not obtained for AfrE and BSAfE in the creative register and WSAfE in the reportage register, which is why some data are missing from the table.

	AfrE		BSAfE		WSAfE		AusE	
	<i>s-gen</i>	<i>of-gen</i>	<i>s-gen</i>	<i>of-gen</i>	<i>s-gen</i>	<i>of-gen</i>	<i>s-gen</i>	<i>of-gen</i>
Academic	3,67	16,72	2,57	18,37	3,01	15,03	6,53	15,50
Creative					4,25	6,27	3,64	5,40
Instructional	2,69	13,66	1,23	25,31	2,69	12,62	1,70	9,28
Popular	3,43	9,81	3,47	15,22	2,66	4,62	3,33	10,38
Reportage	2,88	13,29	2,78	9,34			4,69	6,25

Table 5.29: Normalised frequencies (per 1 000 words) of *s-* and *of-*genitives per sub-variety and register, in edited texts

The table shows that across all registers and for all sub-varieties, the normalised frequency of the *of*-genitive is higher than that of the *s*-genitive (as expected from the proportional analysis). The strength of this preference is lower in the WSAfE creative and popular registers, and the AusE creative and reportage registers where there is less divergence in the normalised frequencies of the *s-* and *of*-genitives.

In terms of the differences across the sub-varieties, with the exception of reportage, BSAfE has the highest normalised frequency of *of*-genitives in all registers, indicating that the *of*-genitive is used more often in BSAfE. The pattern for the remaining three sub-varieties is less straight-forward: with the exception of the popular register, AfrE has the second highest normalised frequency of *of*-genitives. AusE has the third highest normalised frequency of *of*-genitives in the academic and popular registers, but the lowest in the creative and reportage registers, while WSAfE has the third highest normalised frequency of *of*-genitives in the instructional register, and the lowest in the academic and popular registers. The pattern that therefore emerges is that (broadly) *of*-genitives are used more often in BSAfE texts, followed by AfrE and then WSAfE and AusE (which vary across registers).

There is a less stable pattern regarding the normalised frequency of *s*-genitives. In the academic and instructional registers, BSAfE has the lowest normalised frequency of *s*-genitives (together with the highest normalised frequency of *of*-genitives). BSAfE academic and instructional writing is therefore very conservative in its use of the genitive, with a much stronger preference for the more formal and older of the two constructions, than the incoming form. Furthermore, in the academic and instructional registers, AfrE and WSAfE have similar normalised frequencies for the *s*-genitive, suggesting some convergence between the two sub-varieties in the use of the *s*-genitive in these two registers. (Further evidence for this convergence can be found in the similar normalised frequencies of the *of*-genitive for these two sub-varieties in the same two registers.) AusE has the highest normalised frequency of *s*-genitives in the academic register and the lowest in the instructional, which together with the normalised frequencies for the *of*-genitives, suggest that AusE has a particularly strong preference for the incoming form in the academic register and a low preference for the same form in the instructional register. In the popular register, BSAfE, AfrE and AusE have similar normalised frequencies for the *s*-genitive (but divergent frequencies for the *of*-genitive). WSAfE, on the other hand, has the lowest normalised frequency of *s*-genitives. Framed against the normalised frequency of the *of*-genitive for WSAfE in the popular register,

which is also very low, it may be concluded that genitives are used less frequently in WSAfE popular writing (compared to the other sub-varieties, but also the other WSAfE registers). Lastly, AfrE and BSAfE use the *s*-genitive less frequently in reportage, compared to AusE, where the normalised frequency is the second highest for that register. Furthermore, interpreted against the normalised frequencies for the *of*-genitives in reportage, it appears that AfrE and BSAfE have a stronger preference for the *of*-genitive in this register, and use it more frequently than it is used in AusE.

5.3.1.3 Statistical evaluation of the interaction between variety and register

So far, the overall analysis of the genitive alternation has focused on a discussion of the differences in the proportions and normalised frequencies of the two genitive constructions across the four sub-varieties and five registers. To evaluate these differences statistically, a conditional inference tree in which the proportion of *s*-genitives is set as the outcome variable, and REGISTER and SUB-VARIETY are set as predictor variables, is used (see discussion of conditional inference trees in Section 4.3.4).

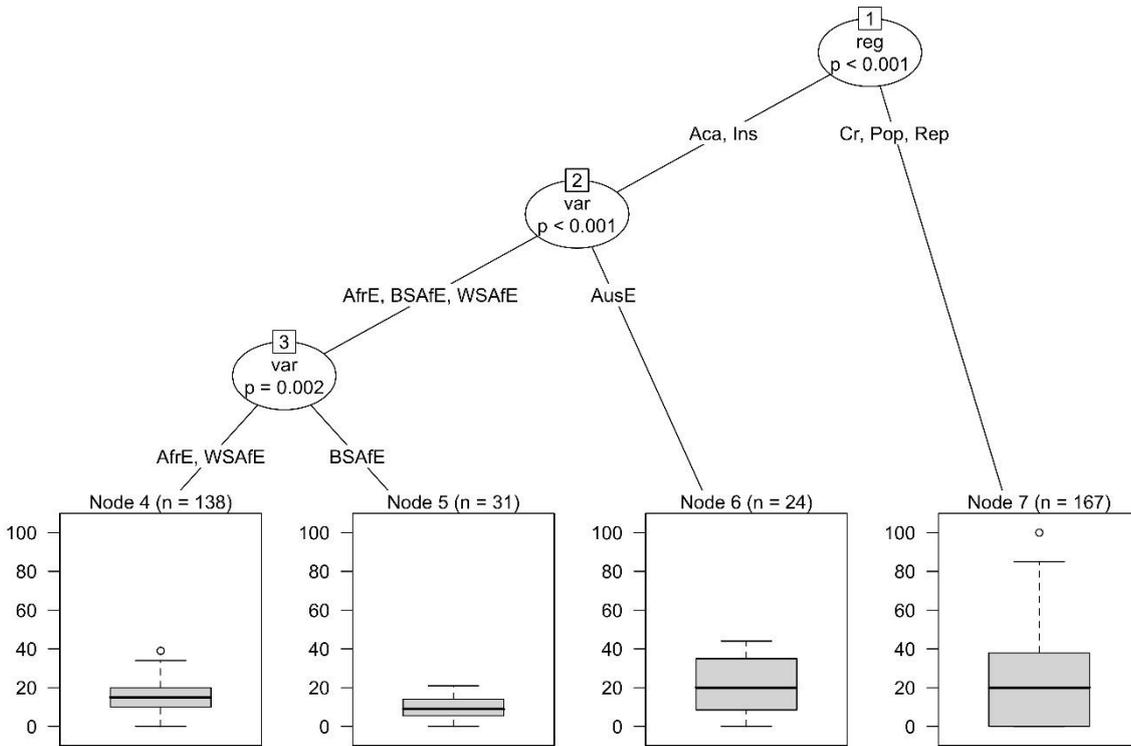


Figure 5.13: Conditional inference tree for the proportion of *s*-genitives, with register and sub-variety as predictors

The conditional inference tree in Figure 5.13 shows that the highest level predictor for the proportion of *s*-genitives is REGISTER (node 1), with the formal academic and instructional registers clustering in the left main branch of node 1 (with overall lower proportions of *s*-genitives), and the creative, popular and reportage registers clustering in the right main branch of node 1 (with overall higher proportions of *s*-genitives). This provides evidence that in the current dataset, the proportion of *s*-genitives is strongly

influenced by the formality of registers. The creative, popular and reportage registers in the right main branch of node 1 are not further partitioned into sub-varieties; however, since these are the three registers in which data are missing for some (sub-)varieties (there are no data for AfrE and BSafE in the creative register, and no data for WSAfE in reportage), it is not possible to determine if there is an effect for sub-variety in the proportional preference of the *s*-genitive in these registers.

The left main branch of node 1, does point to differences in the proportion of the *s*-genitives for SUB-VARIETY. Node 2 splits the proportional differences in the academic and instructional registers along the lines of variety, and partitions the sub-varieties along country lines, so that in the academic and instructional registers there is a clear and significant difference for the two countries: AusE, which is represented in the right branch of node 2, has a higher proportion of *s*-genitives in both registers, compared to the three South African sub-varieties, AfrE, BSafE and WSAfE, which are further partitioned in node 3. The difference between AfrE and WSAfE, and BSafE in terms of the proportional preference of the *s*-genitive is clear: AfrE and WSAfE have higher proportions of the *s*-genitive in academic and instructional writing, compared to BSafE where the proportion of *s*-genitives is significantly lower.

The statistical evaluation of the effect of register and variety on *s*-genitive proportions reveals a pattern that confirms the findings of the general analysis of the proportions and normalised frequencies of the two genitive constructions across registers and sub-varieties: in the current model, the main predictor in determining differences in the proportion of *s*-genitives is register, with the academic and instructional registers showing a lower proportion of *s*-genitives than the creative, popular and instructional registers. For the academic and instructional registers, where there are data for all four sub-varieties, there are clear differences in the proportion of *s*-genitives across the two countries, with AusE showing a stronger preference for more informal *s*-genitives in these two registers, reflecting a stronger tendency towards colloquialisation in AusE than in the other sub-varieties. Within the South African sub-varieties, AfrE and WSAfE cluster together in respect of the proportional frequency of the *s*-genitive, providing evidence that the two sub-varieties function like a dialect cluster, at least in academic and instructional writing, with BSafE distinguishing itself as being separate and less advanced in taking up the *s*-genitive, in comparison.

5.3.2 The influence of editorial intervention on genitive alternation across varieties and registers

Having established that there are differences in the proportions and normalised frequencies of *s*- and *of*-genitives across registers and sub-varieties in edited texts, the analysis now shifts the focus to the influence of editorial intervention on the proportions and frequencies of the two constructions across the registers and sub-varieties. As outlined in Section 4.3.4, the analysis of the influence of editorial intervention is presented in three stages. First, differences in the proportional preferences of the two constructions are analysed across sub-varieties, registers, and unedited and edited texts. These differences are further subjected to statistical evaluation carried out using conditional inference trees. Second, the normalised frequencies of the two

constructions are analysed across sub-varieties, registers, and unedited and edited texts in order to determine if shifts in the proportional preferences of the two constructions are due to changes in the frequency of one construction or both. Third, a more fine-grained analysis of the shifts in proportional and normalised frequency is conducted in order to explore what types of editorial changes account for the changing proportional and normalised frequencies. This analysis discusses the frequency shifts across each register and sub-variety using a derived calculation of the percentage of *s-* and *of-*genitives that are added, removed, replaced or retained in each register in each sub-variety.

5.3.2.1 Distributional and statistical analysis of the influence of editorial intervention on the proportional frequencies of genitive variants across variety and register

Table 5.30 shows the proportional changes in the frequency of the *s-* and *of-*genitives across the unedited and edited texts in each sub-variety and for each register. From the table it appears that editorial intervention leads to substantial shifts in proportional frequencies in AfrE academic writing and reportage, WSAfE creative writing, WSAfE instructional writing, and BSAfE reportage.

	AfrE		BSAfE		WSAfE		AusE	
	Unedited (s% : of%)	Edited (s% : of%)						
Academic	14 : 86	18 : 82	12 : 88	12 : 88	17 : 83	17 : 83	28 : 72	29 : 71
Creative					61 : 39	40 : 60	44 : 56	40 : 60
Instructional	15 : 85	16 : 84	5 : 95	5 : 95	50 : 50	18 : 82	14 : 86	16 : 86
Popular	27 : 73	26 : 74	18 : 82	19 : 81	39 : 61	37 : 63	24 : 76	24 : 76
Reportage	21 : 79	18 : 82	15 : 85	23 : 77			43 : 56	43 : 57

Table 5.30: Proportional frequencies of *s-* and *of-*genitives across variety, register, and unedited and edited texts

With regard to the academic register, editorial intervention increases the proportion of *s-*genitives by 4% in AfrE and by 2% in AusE, while leaving the proportions unchanged in WSAfE and BSAfE. As noted in Section 5.3.1.1, the proportion of *s-*genitives in the edited AusE academic register is unusually high, and it seems that this has its origins in the choices of AusE academic authors. The fact that the Australian editors accept (and even slightly increase) this proportion suggests that for these linguistic gatekeepers, the use of the less formal variant in this extremely formal register is acceptable. Furthermore, the shift in the proportion of *s-*genitives in the AfrE academic register brings the proportion closer to that of WSAfE, and might point to a tendency among editors to pull the two sub-varieties closer together (even if the proportion of *s-*genitives in WSAfE is decreased slightly).

Editorial intervention is much more extensive in the creative register than it is in the other registers. For this register, it appears that SAfE and AusE editors make similar kinds of changes, just to differing degrees. In both sub-varieties, there is a decrease in the proportion of *s-*genitives in the edited creative register, so that in the edited texts, the proportion of genitives is identical for the two sub-varieties. The extremely high proportion of *s-*genitives in unedited WSAfE suggests that WSAfE authors are particularly receptive to the use of the *s-*genitive, but that this use is constrained by editors in this register.

Editorial intervention in BSAfE instructional writing does not alter the proportions of the two genitive constructions, but it is noteworthy that editors of BSAfE maintain the extremely low proportion of *s*-genitives for this sub-variety in this register, suggesting that the near categorical use of the *of*-genitive among BSAfE authors is reinforced by editors in this register. Interpreted against the slight increase in the proportion of *s*-genitives in the edited texts of the AfrE instructional register, and the drastic reduction of the *s*-genitive in the WSAfE instructional register, it appears that editorial intervention in SAfE has two effects in this register: first, editors reinforce the near categorical use of *of*-genitives by BSAfE authors, and second, editors appear to be pulling AfrE and WSAfE closer together, by increasing the proportion of *s*-genitives in AfrE writing and drastically reducing the proportion in WSAfE. In so doing, they move AfrE further away from BSAfE, but pull WSAfE in a conservative direction. Editors of AusE also influence the proportions of the two constructions in the instructional register by slightly decreasing the proportion of *s*-genitives.

In the popular register, editorial intervention in AusE does not result in a shift towards a higher proportion of the *s*-genitive. As noted in Section 5.3.1.1, WSAfE is in the lead of the four sub-varieties in terms of its preference for the *s*-genitive in this register. This preference clearly originates in the writing of the WSAfE authors; however, editors of these texts appear to constrain this by decreasing the proportion of *s*-genitives in edited WSAfE writing, a tendency that mirrors that of the creative and instructional registers, although the effect is less strong in the popular register. In BSAfE, editorial intervention results in a slight increase in the proportion of *s*-genitives, while in AfrE it results in a slight decrease.

Lastly, in reportage, in AfrE and BSAfE editorial intervention influences the proportions of the two constructions, but in opposite directions. Whereas editorial intervention results in a decrease in the proportion of *s*-genitives by 3% in AfrE, it results in an 8% increase in the proportion of *s*-genitives in BSAfE, suggesting that editorial intervention actually facilitates (and even bolsters) the openness of this particular register to the incoming variant; however, this increase in the proportion of *s*-genitives might also be due to editors simply trying to adjust the proportions of the two genitive constructions to be more in line with the typical conventions of this register – but this only seems to be the case for editors of BSAfE and not editors of AfrE newswriting. As discussed in Section 5.3.1.1, BSAfE consistently has the lowest proportion of *s*-genitives across all registers, except in reportage. The analysis in Section 5.3.1.1 is based on the edited texts only (which are the kinds of texts typically included in the written published components of different corpora), and it is therefore striking to note that the proportion of *s*-genitives in unedited BSAfE reportage is much lower than in the unedited writing of the other sub-varieties represented. It therefore appears that BSAfE authors have a much lower preference for the *s*-genitive in all registers, but editorial intervention bolsters this in the (popular and reportage) edited registers, so that in reportage BSAfE appears to be more receptive to the incoming form than AfrE.

Editors do therefore play a role in shaping the proportional frequencies of the *s*- and *of*-genitive constructions, but that the effect differs across the varieties, and within each variety, across the different

registers. To evaluate these differences statistically, a conditional inference tree in which the proportion of *s*-genitives is set as the outcome variable, and REGISTER, SUB-VARIETY and STATUS (with two levels, edited and unedited) are set as predictor variables, was generated (see discussion of conditional inference trees in Section 4.3.4). This conditional inference tree is similar to the one presented in Section 5.3.1.3, with the exception that EDITED STATUS is added as predictor variable. A striking difference that emerges when editorial intervention is added as a predictor to the model is that the most important (first-level) predictor is SUB-VARIETY – unlike in the model for the edited writing only, in which the most important predictor is REGISTER. This suggests, in the very first instance, that editing removes some of the strong varietal distinctions in favour of a more homogenous norm in which register is a stronger predictor.

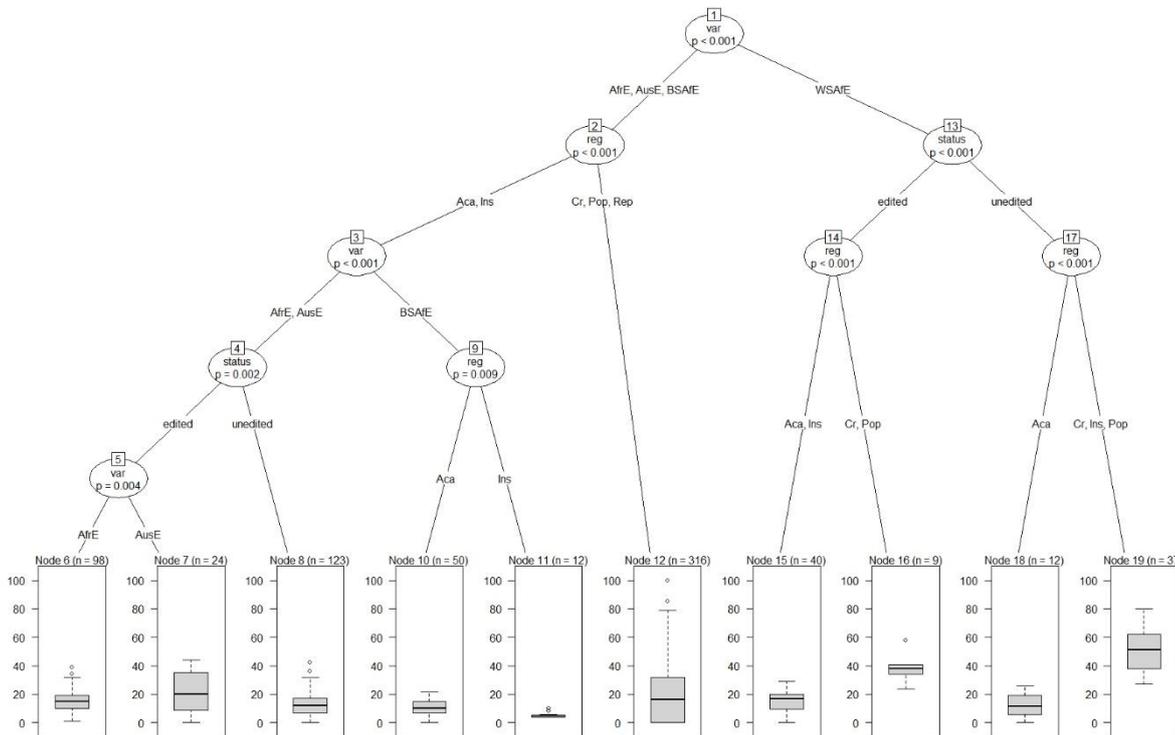


Figure 5.14: Conditional inference tree for the proportion of *s*-genitives, with register, sub-variety and edited status as predictors.

Figure 5.14 confirms the findings of the analysis of the effects of SUB-VARIETY, REGISTER and EDITED STATUS on the proportional frequencies of the *s*-genitive. Sub-variety, the first split in the tree at node 1 is the most important conditioning variable for the proportion of *s*-genitives, with AusE, AfrE and BSafE grouped together in the left branch, and WSAfE represented in the right branch, demonstrating that the proportion of *s*-genitives is conditioned differently in WSAfE than it is in AusE, AfrE and BSafE (when these three predictors are included in the model).

The right branch of node 1 splits WSAfE in terms of EDITED STATUS at node 13, demonstrating that in WSAfE editorial intervention is an important conditioning variable in the proportion of *s*-genitives. Beginning with unedited writing: node 17 splits the unedited WSAfE registers into two branches, in which

the academic register is represented in the left branch (node 18, median = 11.5) and the creative, instructional and popular registers in the right branch (node 19, median = 51). This shows that WSAfE authors of academic writing differ from the authors of the other three registers in their proportional use of *s*-genitives, which is much lower in the academic register than in the other three registers. The edited branch of node 13 as well as the unedited branch of node 13 are then further split on the basis of register at nodes 14 and 17 respectively. Node 14 splits the edited WSAfE registers into two branches: the left branch represents the academic and instructional registers (node 15, median = 16.6) and shows that the edited versions of these two registers cluster together in terms of the proportion of *s*-genitives, while the right branch groups the edited creative and popular registers together (node 16, median = 38). The conditional inference tree therefore demonstrates that editorial intervention plays an important role in conditioning the proportion of *s*-genitives in the five registers. Specifically, it shows that the greatest effect of editorial intervention is in the instructional register and that this effect pulls the proportion of *s*-genitives in this register closer to the academic register, by reducing the proportion of these *s*-genitives significantly.

The left branch of node 1 groups AusE, AfrE and BSAfE together, but splits these varieties in terms of register at node 2. The right branch of node 2 groups the creative, popular and reportage registers and does not split these further, suggesting that status is not an important conditioning variable for the proportional frequency of the *s*-genitive in these registers for these varieties. However, this is most likely due to the fact that there are no data for AfrE and BSAfE in the creative register, which skews the representation of the varieties somewhat. The left branch of node 2 splits the academic and instructional registers first by variety at node 3 (so that AfrE and AusE are represented in the left branch of this node, and BSAfE is represented in the right branch). With regard to BSAfE, the conditional inference tree splits this sub-variety in terms of register at node 9, so that the BSAfE academic and instructional registers are partitioned into separate nodes (BSAfE academic, node 10, median = 10; BSAfE instructional, node 11, median = 5). This confirms two important findings: first, status (unedited vs edited) does not appear to condition the proportional frequency of *s*-genitives in the BSAfE academic and instructional registers, but there are significant differences in the median proportional preferences of the *s*-genitive in these two registers in this sub-variety.

Node 4 splits AfrE and AusE academic and instructional writing first by status, and then the edited branch by variety at node 5. This demonstrates that in AfrE and AusE, there is no significant difference in the proportional preference of the *s*-genitive in the unedited academic and instructional registers (node 8, median = 12), but there is a difference in the edited academic and instructional registers of these two varieties, where AfrE has a significantly lower proportional frequency of *s*-genitives in edited academic and instructional writing (node 6, median 15), compared to AusE (node 7, median = 20).

The conditional inference tree confirms the effect of editorial intervention on the proportions of the two constructions across the sub-varieties and registers. . For WSAfE, the effect of editorial intervention is greatest for the instructional register, where editors significantly reduce the median proportions for *s*-genitives, moving the very high proportional frequency of this construction in this register away from the

similar frequencies in the creative and popular registers, and bringing it in line with academic writing. Editors of WSAfE therefore constrain the increase of the *s*-genitive in unedited instructional WSAfE writing and bring this register in line with conventional expectations of instructional writing. For BSAfE, the unedited texts have a very low overall proportional frequency of the *s*-genitive, which is in all likelihood due to the slower progress of this sub-variety along the stages of the Dynamic Model, but also due to a possible effect of substrate influence (where the *s*-genitive construction is unavailable in the Bantu languages). EDITED STATUS does not play any significant effect in shifting the proportion of *s*-genitives in the registers of this sub-variety: in other words, editors of BSAfE accept the very low proportional frequency of *s*-genitives in BSAfE, which provides evidence for endonormativity for this sub-variety.

In AfrE, editorial intervention only plays a significant role in the academic and instructional registers, where editors increase the proportional frequencies of the *s*-genitive in these two registers, compared to the popular and reportage registers, where the decrease in the proportion of *s*-genitives in edited writing is not shown to be significant. This demonstrates that editors of AfrE boost the proportion of *s*-genitives in these two most formal registers, which brings the proportions of the two constructions much more in line with WSAfE. This suggests convergence between the two sub-varieties, aided by editorial intervention.

In AusE, editorial intervention does not appear to play all that big a role in shifting the proportions of the two constructions in either direction: editors mostly accept the proportions of the two constructions across the registers, which reflects the endonormative stabilisation of this variety. One striking finding in this sub-variety, but which is not interpretable from the conditional inference tree, is AusE editors' acceptance of the very high proportion of *s*-genitives in unedited academic writing. This suggests that the use of this more informal construction in the very formal register is accepted in AusE, and forms part of the writing of AusE authors and the acceptability judgements of AusE editors.

To further probe the differences in the frequencies across the two constructions, the normalised frequencies of each construction for each variety and across the unedited and edited registers were also analysed.

5.3.2.2 Distributional analysis of the influence of editorial intervention on the normalised frequencies of genitive variants across variety and register

The normalised frequencies (per 1 000 words) of each interchangeable construction in each sub-variety and across registers were calculated for the unedited texts and their aligned edited segments (see Table 5.31). As explained in Section 4.3.2, data were not obtained for AfrE and BSAfE in the creative register and WSAfE in the reportage register, which is why some data are missing from the table.

	AfrE				BSAfE				WSAfE				AusE			
	Unedited		Edited													
	<i>s-</i>	<i>of-</i>														
Aca	2,68	17,11	3,67	16,72	2,44	17,56	2,57	18,37	3,00	14,95	3,01	15,03	7,65	19,90	6,53	15,50
Cr									3,44	2,18	4,25	6,27	5,33	6,73	3,64	5,40
Ins	2,53	14,42	2,69	13,66	1,00	17,96	1,23	25,31	2,84	2,87	2,69	12,62	2,20	13,01	1,70	9,28
Pop	3,66	10,10	3,43	9,81	3,54	15,84	3,47	15,22	2,78	4,40	2,66	4,62	3,98	12,32	3,33	10,38
Rep	3,74	14,06	2,88	13,29	1,90	10,57	2,78	9,34					4,84	6,45	4,69	6,25

Table 5.31: Normalised frequencies (per 1 000 words) of *s-* and *of-* genitives per sub-variety and register, in unedited and edited texts

Table 5.31 shows that overall, editorial intervention in AusE decreases the frequency of both *s-* and *of-* genitives across all five registers. The greatest (relative) decrease in the normalised frequencies of *s-* genitives during editorial intervention occurs in creative writing and reportage, while the greatest decrease in the normalised frequencies of the *of-* genitive during editorial intervention occurs in the academic and creative registers. The analysis of the proportional frequencies for AusE in Table 5.28 suggested that in the academic register, editors are increasing the occurrence of the *s-* genitive at the expense of the *of-* genitive. But the normalised frequencies suggest otherwise: editors are reducing the number of both constructions in this register; however, fewer *s-* genitives are removed per 1 000 words than are *of-* genitives, and therefore the proportion of *s-* genitives is higher in the edited texts than in the unedited. A similar pattern emerges in the instructional register: Editors are also decreasing the frequency of the two constructions, with more *of-* genitives being removed than *s-* genitives. Given the already low frequency of the *s-* genitive in this register, this has a greater impact on the proportions of the two constructions, which results in a proportional shift in favour of the *of-* genitive in edited instructional texts. The decrease in the normalised frequencies of the two constructions in the popular and reportage register is relatively even for the two constructions, and therefore do not result in proportional shifts in favour of either construction. An analysis of the total word counts for the different registers in AusE (see Table 4.3), shows that in all five registers, the word count of AusE increases by less than 10% in the edited texts, and so it may therefore be that the overall decrease in the normalised frequencies of the two constructions is due to an increasing overall word count, rather than the targeted removal of the two constructions by editors.

With regard to AfrE, editorial intervention in this sub-variety decreases the frequency of *of-* genitives in all registers, and decreases the frequency of the *s-* genitive in the popular and reportage registers, while increasing this construction's frequency in the academic and instructional registers. It therefore appears that the increase in the frequency of the *s-* genitive comes at the expense of the *of-* genitive in edited academic and instructional AfrE writing – a pattern that is mirrored in the proportional frequencies of the sub-variety for these registers (see Table 5.28). Furthermore, while both constructions decrease in normalised frequency in the popular and reportage registers, this decrease is (relatively) greater for the *s-* genitives than it is for the *of-* genitives, which shifts the proportional preferences of the two constructions in favour of the *of-* genitive. The very slight increase in the total word count of the edited popular and reportage registers (see Table 4.3), might suggest that the decrease in the normalised frequencies of the two constructions in these two registers is the by-product of an increasing word count. However, the decrease in the lower normalised

frequencies of the *s*-genitive, compared to the *of*-genitive, suggests that more *s*-genitives are removed than *of*-genitives. Whether this is a targeted removal on the part of the editors is unclear, but the effect shifts the proportional frequencies in favour of the *of*-genitive. Editorial intervention therefore clearly has an effect on the proportions and frequencies of the two constructions in AfrE, but the motivation appears to be different across registers: in the two more formal registers, academic and instructional, editorial intervention seems to target the perceived overuse of the *of*-genitive in these two registers by reducing the frequency of these constructions while at the same time increasing the frequency of the *s*-genitive. In the popular and reportage registers, editorial intervention does not appear to reduce the frequency of the one construction with a concomitant increase in the other. However, the *s*-genitive decreases far more than the *of*-genitive, which shifts the proportions in favour of the *of*-genitive.

In BSAfE, editorial intervention increases the normalised frequency of both constructions in the academic and instructional registers. In both registers, the increase in the normalised frequency of the *of*-genitive is greater, compared to the *s*-genitive. This does not appear to influence the proportional frequencies of the two constructions in these two registers, which is most likely due to the increase in the word count of the edited versions of these registers (see Table 4.3). With regard to the popular and reportage registers, there is a decrease in the normalised frequency of the *of*-genitive in the edited versions of these two registers. In the popular register, the *s*-genitive also decreases in frequency, which is most likely due to a slight decrease in the word counts across the unedited and edited texts, but this still shifts the proportions in favour of the *s*-genitive. In reportage, the decrease in the frequency of the *of*-genitive appears to come to the advantage of the *s*-genitive, which increases in frequency in the edited texts (quite substantially, by comparison). This results in a shift in the proportional frequencies of the two constructions in this register, so that the very low frequency of *s*-genitives in unedited BSAfE reportage is increased by editors in the edited texts.

The normalised frequencies in Table 5.31 reveal that editorial intervention in WSAfE increases the frequency of *of*-genitives in all registers, while the frequency of the *s*-genitive is increased in the academic and creative registers, and decreased in the instructional and popular registers. The increase in the frequency of the two constructions in the academic register is minor and does not appear to influence the proportions of the two constructions. Since there is a slight increase in the word count of the edited academic register, it is probable that editors are adding the two constructions in equal measure in the academic register as a by-product of increasing the overall word count (see Table 4.3). In the creative register, the increase in the normalised frequency of the *of*-genitive far outstrips that of the *s*-genitive (where the increase in the normalised frequency is much lower), so that the overall effect on the proportions moves the proportional preference in this register towards the *of*-genitive. Given that there is a 10% decrease in the word count of the edited creative register, and given that editorial intervention in creative texts is typically quite substantive, it seems likely that the drastic shift is due to some other substantive revision that reduces the total word count of the register but increases the frequency of both constructions. The instructional register of WSAfE stands out as the one register in which editorial intervention appears to have the greatest influence. The normalised frequencies of the two constructions are quite low and very similar in the

unedited instructional texts (*s*-genitive, 2,84; *of*-genitive, 2,87); however, there is a minor decrease in the frequency of the *s*-genitive (2,69) and a substantial increase in the *of*-genitive (12,62) in the edited texts (with an accompanying small decrease in the total word count of the edited texts). Editors therefore appear to be adding a great deal of *of*-genitives and only removing a few *s*-genitives, which has the effect of shifting the proportional preferences of the two constructions considerably. An inspection of the unedited texts in this register reveals that many of these texts are either learner books or teacher's guides on topics for school-going learners. The editing of these types of texts is typically characterised by a high level of intervention on the part of the editor, who does a substantive edit of the content and scope of this kind of text. This substantive editing usually entails the addition of material or a major rewriting of the text to optimise it for the reader in terms of their level of education and the purpose for which the text is used. It is therefore likely that the substantial shift towards the higher proportion and frequency of *of*-genitives is due to the editor rewriting substantial portions of these texts, and in so doing, adding many *of*-genitives, either because this is the preferred construction or because the editor views the less compact construction as more appropriate for the reader or text type.

5.3.2.3 The types of editorial changes that influence the shifting frequencies of the genitive constructions across variety and register

The analyses in Sections 5.3.2.1 and 5.3.2.2 demonstrate that editing introduces shifts in the proportional and normalised frequencies of the *s*- and *of*-genitives in the different sub-varieties and across registers within the sub-varieties. The analysis of the proportional and normalised frequencies shows that in all but the WSAfE academic, BSAfE academic and instructional, and AusE popular and reportage registers, editorial intervention results in a shift in the proportional preference of the two constructions, but that the effect of this intervention differs firstly at the level of variety and secondly, within each register. While these analyses provide evidence that editorial intervention does play a role in the ongoing variation in genitive alternation, they do not allow for a fine-grained interpretation of what exactly it is that editors are doing, and therefore mask *how* editorial intervention moves the unedited texts in each variety and register in either a conservative direction that pulls back the ongoing change by favouring more *of*-genitives, or in an innovative direction that accelerates the ongoing shift towards the increased use of *s*-genitives.

To explore this, this section focuses on an analysis of the types of editorial changes that account for the changing proportional and normalised frequencies across the unedited and edited texts. This analysis discusses the types of editorial changes using a derived calculation of the percentage of *s*- and *of*-genitives that are added, removed, replaced or retained in each register in each sub-variety. In the case of each construction, the derived percentage is based on the total estimated number of each type of change per construction, relative to the total estimated count of that construction in each register and sub-variety. For example, in Table 5.32, the two percentages given as removed for each construction in the academic register are not comparable with *each other*, but rather represent what the editor did with that type of construction in that register, measured against the other interventions for that same construction in that register.

Therefore, the editorial interventions for the *s*-genitive in the academic register of AusE in Table 5.32 are as follows: editors removed 4% of the *s*-genitives from the unedited texts, so that 96% of these genitives are retained in the edited version. The editorial interventions for the *of*-genitive in the same register read as follows: editors removed 8%, so that only 92% are retained from the unedited texts, but they add 1% to this. In this regard, editorial changes that add and retain *s*-genitives and remove and replace *of*-genitives are considered to move the edited texts in an innovating direction, while changes that add and retain *of*-genitives and remove and replace *s*-genitives are considered to move the edited texts in a conservative direction (see Table 4.6 for a detailed explanation and example of each type of change). To facilitate the interpretation in this section, changes that move the register in an innovating direction are shaded in grey.

Register	Percentage added <i>In edited, but not unedited</i>		Percentage removed <i>In unedited, but not edited</i>		Percentage replaced <i>The construction in the unedited is replaced with the variant form in the edited</i>		Percentage retained <i>In unedited and edited</i>	
	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive
Academic	0	1	4	8	0	0	96	92
Creative	1	13	20	16	0	0	80	84
Instructional	0	1	8	5	0	0	92	95
Popular	0	0	1	4	0	0	99	96
Reportage	0	0	15	10	0	0	85	90

Table 5.32: Percentages of the types of editorial changes influencing the frequencies of the *s*- and *of*-genitive constructions across registers in AusE

The analysis of the AusE data in Sections 5.3.2.1 and 0 showed that in this sub-variety, editorial intervention reduces the normalised frequency of both constructions in all registers. Table 5.32 shows that this is mostly due to removal of the constructions in the different registers, but in certain registers editors add and replace constructions.

In the academic register, more *of*-genitives are removed than *s*-genitives, which results in a marginally higher proportion of *s*-genitives in the edited version of this register (see examples (5.1a and 5.1b) and (5.2a and 5.2b)), while a small percentage of *of*-genitives are added (see example (5.3a and 5.3b)). The changes therefore favour the *s*-genitive, even if they do not specifically target the alternation between the two constructions. Furthermore, no *s*-genitives are replaced by the editors in this register, suggesting that editors accept the use of these forms in this register, which, as noted in Sections 5.3.2.1 and 5.3.2.2, is unusually high, compared to other varieties.

(5.1a) ...on saltwater and the role of salt, on **the value of sago**, on coral madrepores and limestone, and the polar night ... [Aus_A-015-O]

(5.1b) ...on saltwater and the role of salt, on coral madrepores and limestone and the polar night... [Aus_A-015-O]

(5.2a) ...On the Spiritual in Art by representing **Kandinsky's subjective impressions**... [Aus_A-012-O]

(5.2b) ...On the Spiritual in Art by representing his subjective impressions... [Aus_A-012-E]

(5.3a) ...is threatened by Vanikorans on his ship prior to departure. [Aus_A-015-O]

(5.3b) ...is threatened by **the presence of Vanikorans** on his ship in leaving Vanikoro... [Aus_A-015-E]

In the creative register, a large percentage of *s*- and *of*-genitives are removed. While the percentage removed for each appears to be similar, because there are many more *of*-genitives than *s*-genitives in the unedited texts (as evidenced by the normalised frequencies in Table 5.31), a greater number *of*-genitives are removed by editors than *s*-genitives, but this is counteracted by the percentage of *of*-genitives that are added, which results in a shift in this register in favour of the *of*-genitive. The removal of *s*-genitives and *of*-genitives, as exemplified in (5.4a and 5.4b) and (5.5a and 5.5b), is mostly due to the removal of the possessor from the construction or through the removal of the sentence in which the construction occurs.

(5.4a) ...and saw the tumbling brown house; ~~<character name deleted>~~'s house. [Aus_C-003-O]

(5.4b) ...they glanced up through the window and saw the tumbling brown house. [Aus_C-003-E]

(5.5a) ... had called him to retrace **the steps of his forgotten past**. [Aus_C-006-O]

(5.5b) ... to retrace his steps to his forgotten past. [Aus_C-006-E]

In the instructional register, editorial changes remove 5% of *of*-genitives, as exemplified in (5.6a and 5.6b) and 8% of *s*-genitives (see (5.7a and 5.7b)) while adding a small percentage of *of*-genitives. The editorial changes therefore favour the *of*-genitive, which serves to boost the proportion of *of*-genitives and moves the register in a constraining direction.

(5.6a) ...but **the eyes of males** are virtually touching dorsally... [Aus_I-009-O]

(5.6b) ...but in males, the eyes are virtually touching dorsally... [Aus_I-009-E]

(5.7a) For example, **Central America's Bullhorn Acacia tree** has massively bloated... [Aus_I-003-O]

(5.7b) Take the case of **the bullhorn acacia tree from Central America**. [Aus_I-003-E]

A similar pattern is found for the popular and reportage registers, where editors only remove the two constructions, without adding or replacing either. As exemplified in (5.8a and 5.8b), (5.9a and 5.9b), (5.10a and 5.10b) and (5.11a and 5.11b), nearly all of these removals are due to a complete removal of a sentence or paragraph in which the genitive occurs, suggesting that this removal is the by-product of substantive editing in these two registers. Since the frequency of the *of*-genitive is higher in the popular register than the *s*-genitive, the effect of the higher percentage of *of*-genitives that are removed, is balanced out against the lower percentage of *s*-genitives removed, so that the normalised frequencies of the two constructions decrease slightly, but the proportions stay the same.

(5.8a) ...is very effective in reducing **the seriousness of human-bear encounters**... [Aus_P-001-O]

(5.8b) ----- [Aus_P-001-E]

(5.9a) ... **Australia's first antivenom**... [Aus_P-001-O]

(5.9b) ----- [Aus_P-001-E]

(5.10a) ...key to this new framework is **the implementation of robust biosecurity practices**... [Aus_R-001-O]

(5.10b) ----- [Aus_R-001-E]

(5.11a) From **<organisation name removed>'s Biosecurity and Product Integrity Services team**... [Aus_R-001-O]

(5.11b) ----- [Aus_R-001-E]

Table 5.33 reflects the percentage of each change in the AfrE registers, and shows that editors make many kinds of interventions that add and remove the two constructions, with only a handful of changes targeting the replacement of the two constructions. As noted in Sections 5.3.2.1 and 5.3.2.2, shifts in the proportions and normalised frequencies of the two constructions in AfrE suggest that in the academic and instructional registers, editors increase the frequency of the *s*-genitive at the expense of the *of*-genitive, but decrease the frequencies of both constructions in the popular and reportage registers in a way that favours the *of*-genitive.

Register	Percentage added <i>In edited, but not unedited</i>		Percentage removed <i>In unedited, but not edited</i>		Percentage replaced <i>The construction in the unedited is replaced with the variant form in the edited</i>		Percentage retained <i>In unedited and edited</i>	
	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive
Academic	3	1	3	5	1	0	96	95
Creative								
Instructional	10	3	3	4	0	0	97	96
Popular	2	3	3	3	1	0	96	97
Reportage	8	1	12	9	0	0	88	91

Table 5.33: Percentages of the types of editorial changes influencing the frequencies of the *s*- and *of*-genitive constructions across registers in AfrE

With regard to academic editing, editors make many more changes that favour the *s*-genitive: more *of*-genitives are removed than *s*-genitives, and more *s*-genitives are added than *of*-genitives. Although the editors do replace 1% of *s*-genitives with *of*-genitives (see examples 5.12a and 5.12 b), because of the lower frequency of the *s*-genitive, these replacements make a minimal contribution to the *of*-genitives. These changes therefore increase the normalised frequency and proportion of *s*-genitives, while reducing the proportion and normalised frequency of the *of*-genitive, which brings the frequencies and proportions of the two constructions much closer to WSAfE. The motivation behind these changes is in all likelihood due to the perception of editors that AfrE authors overuse constructions containing the *of*-genitive, who address this by removing cases where the construction is not deemed necessary, as demonstrated in (5.13a and 5.13b).

(5.12a) **These churches' worship services** are statically drawn from prayer books. [SA_A-118-O]

(5.12b) **The worship services of these churches** are statically drawn from prayer books. [SA_A-011-E]

(5.13a) **For the purpose of this study**, the quantitative measuring instrument used... [SA_A-075-O]

(5.13b) The quantitative measuring instrument used ... [SA_A-062-E]

A slightly different picture emerges in the AfrE instructional register: editors add many more *s*-genitives than they remove, but remove slightly more *of*-genitives than they add. The slight increase in the proportion and normalised frequency of the *s*-genitive, with an accompanying decrease in the *of*-genitive, is therefore not due to editorial intervention that replaces one construction with another; rather, it is simply that editors are removing and adding the constructions at different rates, which boost the frequencies of the *s*-genitives. With regard to the addition of *s*-genitives, many of these interventions target formulation and clarity or seek to make the text more concise (see examples 5.14a and 5.14b), while others are due to the (substantive) addition of text by the editor (see example 5.15a and 5.15b). In both cases, editors invoke their own writing preferences and it is therefore noteworthy that the preference for these types of additions is for the *s*-genitive. The removal of *of*-genitives in this register appears to be associated with a tendency among editors to improve the plain writing of the texts: for example, editors remove framing constructions that contain the *of*-genitive as a sub-part (as exemplified in (5.16)).

(5.14a) These elements, when combined, create the context in which work is carried out, and a business is effectively, and efficiently run. [SA_I-059-O]

(5.14b) These elements, when combined, create the context in which **the business's activities** are carried out. [SA_I-059-E]

(5.15a) ----- [SA_I-059-O]

(5.15b) The virus attacks **the body's immune system**, so that it can not function properly to defend the body against illness. [SA_I-059-E]

(5.16a) ... the effectiveness **of this programme**... [SA_I-062-O]

(5.16b) ...its effectiveness...[SA_I-062-E]

The popular and reportage registers are the two AfrE registers in which editorial intervention reduces the proportion of *s*-genitives. In terms of the kinds of changes made in the popular register reflected in Table 5.33, editors remove a similar amount of *s*- and *of*-genitives (see examples (5.17a) and (5.17b)), but add fewer *s*-genitives than they add *of*-genitives (see examples (5.18a) and (5.18b)), and replace a small percentage of *s*-genitives (see examples (5.19a) and (5.19b)). The overall effect of this is to constrain the *s*-genitive and reduce both its proportional and normalised frequency. This points to a preference in the opposite direction as noted for instructional texts, but it may also be that in this register, language-internal factors (such as syntactic weight) play a role, such as in the replacement in examples (5.19a) and (5.19b).

(5.17a) I think **everyone's favourite wine spoilage factor** is the plaintive cry of "it's corked!" [SA_P-015-O]

(5.17b) I think the most common experience of wine spoilage is when a bottle is corked. [SA_P-015-E]

(5.18a) ----- [SA_P-017-O]

(5.18b) ...**the trial of Oscar Pistorius**... [SA_P-017-E]

(5.19a) She (Rust 2007:208) writes that these days it is all about “spiritual intelligence”, and in order to handle **the twenty first century’s demands**, this intelligence is a necessity. [SA+P-010-O]

(5.19b) She writes that these days it is all about “spiritual intelligence”, and in order to handle **the demands of the 21st century**, this type of intelligence is a necessity. [SA_P-010-E]

In the reportage register, although there appears to be a greater percentage of interventions that favour the *s*-genitive (8% are added, and 9% of *of*-genitives are removed), the frequency of the construction is so low in this register that such interventions have very little effect, compared to the interventions that favour the *of*-genitive (a 12% removal of *s*-genitives and 1% addition of *of*-genitives). Examples of changes that favour the *of*-genitive are presented in examples (5.20) and (5.21) respectively. Although these changes result in a decrease in the normalised frequency of both constructions, this decrease has a greater effect on the proportion of *s*-genitive, which decreases in the edited texts.

(5.20a) **The Blyde and Treur river’s names** were given to them by a group of Voortrekkers who had returned from Delagoa Bay. [SA_R-021-O]

(5.20b) The Blyde and Treur rivers were named by a group of Voortrekkers who had returned from Delagoa Bay. [SA_R-021-E]

(5.21a) ...to serve as a constant reminder on our desks and this was filled with very nice goodies that are helped warm up the cold winter days. [SA_R-016-O]

(5.21b) ...to serve as a constant reminder of **the importance of record management**. [SA_R-016-E]

The percentage of change per type of editorial intervention in BSAfE is reported in Table 5.34, and shows that in BSAfE the percentage of editorial changes is much higher, compared to AfrE. Furthermore, while most of this intervention targets the addition and removal of the two constructions, there are a number of registers in which this intervention replaces (mostly) the *s*-genitive. This suggests a more diverse interventionist approach by editors of this sub-variety, compared to the other sub-varieties.

Register	Percentage added <i>In edited, but not unedited</i>		Percentage removed <i>In unedited, but not edited</i>		Percentage replaced <i>The construction in the unedited is replaced with the variant form in the edited</i>		Percentage retained <i>In unedited and edited</i>	
	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive
Academic	9	10	3	1	16	1	82	98
Creative								
Instructional	48	41	30	9	0	0	70	91
Popular	9	1	1	4	7	0	91	96
Reportage	33	9	9	29	7	0	84	71

Table 5.34: Percentages of the types of editorial changes influencing the frequencies of the *s*- and *of*-genitive constructions across registers in BSAfE

In the BSAfE academic register, the normalised frequencies increase in the edited texts for both constructions, but this does not affect the proportions, which show a clear preference for the *of*-genitive in this register. Although a greater percentage of the types of editorial interventions in this register favours the *of*-genitive, the higher frequency of the *of*-genitive means that there is less overall difference across the unedited and edited texts. The kinds of interventions that add the two constructions are not due to the substantive addition of text (as reflected in the similar word counts across the unedited and edited texts), but seem to be triggered by problematic formulation, questionable grammar or unclear expression or references in the sentence in which the constructions occur (see examples (5.22) and (5.23)). Furthermore, the replacements made by the editors seem to address stylistic preferences for the particular register (as exemplified in (5.24)) or the reduction of wordy formulations (as exemplified in (5.25)). Therefore, the editorial intervention in the BSAfE academic register specifically targets the language choices of the BSAfE authors, rather than amending or adding content.

- (5.22a) Leadership continuously invested in education to be able to skilfully progress towards the economic needs of the time. [SA_A-009-O]
- (5.22b) **The country's leadership** continuously invested in education to be able to progress skilfully towards meeting the economic needs of the time. [SA_A-009-O]
- (5.23a) Regarding (LE) importance related to engagement, respondents viewed a direct correlation with engagement... [SA_A-009-O]
- (5.23b) Regarding **the relation of LE importance** to engagement, the respondents felt that... [SA_A-009-E]
- (5.24a) The study is exploratory in nature and as a result there is less concern with **the sample's representativeness**. [SA_A-022-O]
- (5.24b) ...there was less concern with **the representativeness of the sample**. [SA_A-022-E]
- (5.25a) ...**the state of the environment at home**, these schools do nothing to support these learners... [SA_A-090-O]
- (5.205b) ...**the learner's home situation**, some schools do nothing to support these learners. [SA_A-090-E]

As discussed in Sections 5.3.2.1 and 5.3.2.2, although there is no shift in the proportions of the two constructions in the BSAfE instructional register, there is a substantial (relative) increase in the normalised frequencies of the two constructions. In this register, the BSAfE authors show a near categorical preference for the *of*-genitive, which is maintained (and therefore endorsed) by the editors. The data on the percentage of editorial changes indicate that editors do make many changes in this register, but these changes relate only to the addition and removal of the constructions, rather than their replacement. Table 5.34 shows that more *s*-genitives are added than *of*-genitives, but that many more *s*-genitives are removed than *of*-genitives, the net effect of which does not alter the proportions of the two constructions. An inspection of the concordances extracted for this register reveals that all of the additions of the *s*-genitive are as a result of the correction of what is quite clearly an error in the formulation of the construction (see examples (5.26)

and (5.27)). The inspection of the *of*-genitives reveals that all additions are actually due to substantive changes made by the editors, either through the addition of text, or the substantive rewriting of text (see examples (5.28) and (5.29)). Therefore, when *s*-genitives are added by editors, this is due to an error in the formulation of the construction; however, when editors make substantive changes that require the inclusion of a genitive construction, they more often than not draw on the *of*-genitive construction, reinforcing the extremely high proportion and frequency of this construction in this register.

(5.26a) This is not withstanding **the Universities rules and procedures**. [SA_I-051-O]

(5.26b) This is notwithstanding **the university's rules and procedures**. [SA_I-051-E]

(5.27a) Disputed and rejected ballot papers will not be counted but will be kept separate for **Electoral Commissions decision**. [SA_I-051-O]

(5.27b) Disputed and rejected ballot papers will not be counted, but will be kept separate for **the Electoral Commission's decision**. [SA_I-051-E]

(5.28a) - To align disability and gender frameworks to the EPWP; [SA_I-071-O]

(5.28b) - **the alignment of disability and gender frameworks** to the EPWP; [SA_I-071-E]

(5.29a) ----- [SA_I-70-O]

(5.29b) What are **the demographics of your region?** [SA_I-70-E]

In terms of the percentage of types of editorial interventions in popular BSAfE writing, the data in Table 5.34 show that editors make fewer interventions compared to the other registers in the sub-variety. The analyses in Sections 5.3.2.1 and 5.3.2.2 showed that editors decrease the normalised frequencies of both constructions, but that this decrease is greater for the *of*-genitives, which results in a proportional shift in favour of the *s*-genitive. The percentage of intervention types supports this finding and show that interventions that favour the *s*-genitive (adding *s*-genitives and removing *of*-genitives) outstrip the types of interventions that favour the *of*-genitive (adding *of*-genitives, removing *s*-genitives and replacing *s*-genitives). It is noteworthy that of the interventions that favour the *of*-genitive, the greatest percentage deals with the replacement of *s*-genitives. An analysis of these replacements reveals that most are replacements in which the possessor ends in a final sibilant, suggesting that a linguistic factor influences which construction editors select (see examples (5.30) and (5.31)).

(5.30a) **CHURCH'S SLOGAN** [SA_P-007-O]

(5.30b) **SLOGAN OF THE CHURCH** [SA_P-007-E]

(5.31a) **Our Head office's capacity** has increased so much thus... [SA_P-007-O]

(5.31b) **The capacity of our head office** has increased so much... [SA_P-007-E]

Reportage is singled out as the register with the greatest shift in proportions in the BSAfE data: editorial intervention increases the normalised frequency of the *s*-genitive at the expense of the *of*-genitive, so that

the proportion of *s*-genitive increases substantially. The data in Table 5.34 bear this out, and demonstrate that the most common types of editorial interventions in this register favour the *s*-genitive: 33% of *s*-genitives are added and 29% of *of*-genitives are deleted, while only a handful of *of*-genitives are added (9%), *s*-genitives removed (9%) and a small percentage of *s*-genitives (7%) are replaced. The analysis of the *s*-genitives that are added reveals that the constructions are added either as a result of substantive rewriting of the unedited text or to include more information (see examples (5.32) and (5.33)), while the *of*-genitives that are removed appear to be motivated by factors of economy (which make the expressions more compact) and clarity (see examples (5.34) and (5.35)). An analysis of the *s*-genitives that are replaced reveals that as with the popular register, the motivation behind the replacement relates to a linguistic factor, except that in the case of reportage, the linguistic factor informing this replacement is the syntactic weight of the possessor and possessum, as exemplified in (5.36) and (5.37).

(5.32a) Hlengwa was dragged out of the house by the friend's mother to their house. [SA_R-085-O]

(5.32b) She went to **Hlengwa's house** and dragged him all the way to her house. [SA_R-085-E]

(5.33a) He said the councillors at the top of the list to loose their positions and be relegated are Emfuleni Executive mayor Gretha Hlongwane, Speaker Msebenzi Gqoloshe, MMC for infrastructure VM Godoli, MMC for finance Mabandla Ronyuza and MMC for public safety Velly Mcera. [SA_R-056-O]

(5.33b) The most influential position amongst those mentioned is that of ELM Mayor Gretha Hlongwane who, it is said, will be deployed to **a MMC's (Member of Mayoral Committee) position** in Sedibeng District Municipality and be replaced by SDM MMC for Strategic Planning and Economic Development, Maipato Tsokolibane. [SA_R-056-E]

(5.34a) **The primary aim of these visits** was to... [SA_R-029-O]

(5.34b) **The primary aim** was to see the progress made on registered docketts [SA_R-029-E]

(5.35a) ...**the houses of the poor** by a certain white minority. [SA_R-040-O]

(5.35b) ...houses belonging to the poor by a certain white minority. [SA_R-040-E]

(5.36a) It's **our fore-father's land** and it's not for sell. [SA_R-091-O]

(5.36b) It's **the land of our forefathers** and it's not for sale. [SA_R-091-O]

(5.37a) **These children's safety** lean on us as the landlords... [SA_R-121-o]

(5.37b) In my business **the safety of these children** is paramount. They depend on us... [SA_R-121-E]

The percentage of changes per type of editorial intervention for WSAfE in Table 5.35 suggests that editors of this sub-variety make many more changes and many more types of changes than in the AusE and AfrE sub-varieties. The analysis of the proportions of the two constructions in Section 5.3.2.1 showed that the effect of this was to move the creative, instructional and popular registers in a conservative direction, by reducing the high proportion of *s*-genitives in the unedited texts of these registers. This suggests that WSAfE authors show a stronger preference for the *s*-genitive in these registers, compared to the other sub-

varieties, but that the editors reign this in by increasing the frequency of *of*-genitives and in some instances also reducing the frequency of *s*-genitives.

Register	Percentage added <i>In edited, but not unedited</i>		Percentage removed <i>In unedited, but not edited</i>		Percentage replaced <i>The construction in the unedited is replaced with the variant form in the edited</i>		Percentage retained <i>In unedited and edited</i>	
	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive	<i>s</i> -genitive	<i>of</i> -genitive
Academic	9	3	7	0	0	0	93	100
Creative	4	5	22	14	0	0	78	86
Instructional	6	8	12	8	1	1	87	91
Popular	20	13	7	10	1	0	92	90
Reportage								

Table 5.35: Percentages of the types of editorial changes influencing the frequencies of the *s*- and *of*-genitive constructions across registers in WSAfE

In the WSAfE academic register, editors target their interventions at the *s*-genitive, either adding or removing them, with very little intervention that specifically targets the *of*-genitive. While the proportional frequencies, which show a clear preference for the *of*-genitive, remain unchanged across the unedited and edited versions, there is a slight increase in the normalised frequencies of both constructions. With regard to the *s*-genitives, this increase is due to the addition of more *s*-genitives than are removed, while the increase in *of*-genitives is due to the addition of a small number of *of*-genitives. The addition of *s*-genitives and *of*-genitives is, in nearly all cases, due either to the addition of a possessor to clarify the referent (as exemplified in (5.38) and (5.39)) or through the correction of an error in the construction (such as a missing 's to mark possession), as exemplified in (5.40). The removal of the *s*-genitive is mostly triggered by some other issue in the sentence in which it occurs, as exemplified in (5.41).

- (5.38a) For instance, different patterns of verbal and non-verbal communication may create serious errors and lead to misinterpretations, misunderstanding and confusion and affect the perceptions of others. [SA_A-052-O]
- (5.38b) Such differences also affect **people's perceptions** of one another. [SA_A-052-E]
- (5.39a) Tourists are classified as temporary visitors staying at least 24 hours, whose purpose could be categorized as leisure... [SA_A-051-O]
- (5.39b) Tourists are classified as temporary visitors staying at least 24 hours in the place they are visiting. **The purpose of their visit** may be categorised as leisure... [SA_A-051-E]
- (5.40a) The importance of culture studies is therefore imperative to the understanding of tourism socio-cultural impacts. [SA_A-049-O]
- (5.40b) Cultural studies are therefore imperative to the understanding of **tourism's socio-cultural impacts**. [SA_A-049-E]
- (5.41a) It is **the adolescent's** task to develop positive relationships with peers, members of the opposite sex and authority figures. [SA_A-011-O]

- (5.41b) The adolescent is tasked with developing positive relationships with peers, members of the opposite sex and authority figures. [SA_A-011-E]

Editors of texts in the WSAfE creative register make substantial changes during the editing process, as evidenced by a swopping of the proportions of the two constructions in the edited texts in favour of the *of*-genitive and the increase in the frequency of both constructions. These changes might seem contradictory but they need to be viewed against the background of a substantial decrease (of more than 10%) in the word count of the edited sub-corpus for this register. Table 5.35 shows that in this register a high percentage of the two constructions are removed, with many more *s*-genitives being removed than *of*-genitives. Despite this decrease, the substantial drop in the total word count of the edited texts means that the normalised frequencies of the two constructions increase. An analysis of the individual observations that are removed demonstrates that in the majority of cases the constructions are removed as a result of the removal of the sentences or paragraphs in which they occur, which suggests that the biggest changes in this register are due to the substantive editing practices typically associated with the editorial role in this register. A small percentage of the two constructions is added, but these are mostly motivated by a reformulation of the sentence (see examples in (5.42) and (5.43)), or are added as a consequence of the addition of text. All of this results in a higher normalised frequency of the two constructions, that is greater for the *of*-genitives, and adjusts the proportions of the two constructions in favour of the *of*-genitive.

- (5.42a) <character name deleted> smiled as he bent down to kiss <character name deleted> on the cheek. [SA_C-006-O]
- (5.42b) <character name deleted> smiled as he brushed <**character name deleted**>'s **cheek** with his lips. [SA_C-006-E]
- (5.43a) That night <character name deleted> dreams she is still lying on the ground by the ruined church, but now she is alone. [SA_C-001-O]
- (5.43b) That night <character name deleted> dreams that she is still lying on **the ground of the ruined church**, but now she is alone. [SA_C-001-E]

The instructional register, as pointed out in Sections 5.3.2.1 and 5.3.2.2, is the one WSAfE register where editorial intervention has the greatest influence on the proportions of the two constructions, particularly in respect of the normalised frequency of the *of*-genitive, which increases substantially and which accounts for the drastic shift in the proportions of the two constructions in favour of the *of*-genitive. The data in Table 5.35 shows that this is due to the addition of more *of*-genitives than *s*-genitives and the removal of more *s*-genitives than *of*-genitives. As noted previously, this register includes a number of school books and teacher's guides targeted at school-level learners, and is characterised by an editorial role that is focused on ensuring that these texts are suited to this audience. Since the *of*-genitive is considered the more analytic, periphrastic and therefore more transparent of the two constructions (Heller et al., 2017b, p. 20), it is likely that the substantial increase in this construction is motivated by the substantive editing of these texts and the preference for the *of*-genitive as a way to ensure the accessibility of the content of these texts, as

exemplified in (5.44), (5.45) and (5.46). The removal of *s*-genitives is mostly located in the instructional texts that are not intended for school-going learners, but which are instruction manuals and informational texts. In these cases, the *s*-genitive is usually deleted to remove unnecessary repetition or wordiness in texts, as exemplified in (5.47) and (5.48).

- (5.44a) Trace the map showing population and the map showing rainfall on page 00. [SA_I-035-O]
- (5.44b) Trace the map showing **the population distribution of South Africa** on these pages and the map showing rainfall... [SA_I-035-E]
- (5.45a) * Asking questions about what they see in a drawing. [SA_I-053-O]
- (5.45b) * Asking questions about **the content of a drawing**. [SA_I-053-E]
- (5.46a) Draw a scale line next to your drawing. Indicate what the scale is. [SA_I-033-O]
- (5.46b) Draw a scale line next to your drawing. The scale line should clearly show what **the scale of your drawing** is. [SA_I-033-E]
- (5.47a) There are municipalities that control their demand by installing tariff miniature circuit breakers (tariff mcb) in **the supplier's meter box**. [SA_I-081-O]
- (5.47b) Some municipalities control domestic demand by installing tariff MCBs in meter boxes. [SA_I-081-E]
- (5.48a) A child is adoptable if: [...] the whereabouts of **the child's parents or guardian** cannot be traced... [SA_I-074-O]
- (5.48b) A child is adoptable if: [...] the whereabouts of its parents or guardian cannot be traced... [SA_I-074-E]

In the popular register, editorial intervention results in a slight increase in the frequency of the *of*-genitive with an accompanying greater comparative decrease in the normalised frequency of the *s*-genitive, so that the proportion shifts slightly in favour of the *of*-genitive. There is a substantial increase in the word count in this register (of just over 20%), and so while the data in Table 5.35 might suggest that editors add more *s*- and *of*-genitives to these texts, given the lower frequency of the *s*-genitive in the unedited texts, the additions are in all likelihood lower than those of the *of*-genitive, where the normalised frequency is much higher. This means that the effect of these additions is counteracted by the increasing word count of the edited texts so that the normalised frequency decreases for the *s*-genitive and increases for the *of*-genitive. An inspection of the *s*- and *of*-genitives that are added confirms this: most of the additions are due to the addition of text, which increases the word count in the popular texts. For example, in (5.49) the *s*-genitive occurs in a sentence that is part of the larger inclusion of a section of text in the edited version. In this specific case, the text is a book that advises parents on raising children in a world with ever-increasing technology. The added section is an example of a 'contract' in which parents and their children agree on how technology can be used, which was added to the text during the editing stage. The example in (5.50) illustrates the same kind of addition for the *of*-genitive, which is taken from a book on travel in South Africa, and which demonstrates the inclusion of information that did not appear in the unedited text. Editors

of popular texts therefore make substantive additions to the texts that edit, which results in an increase in the frequency of both constructions to the marginal benefit of the *of*-genitive.

(5.49a) ----- [SA_P-017-O]

(5.49b) I will obey **the school's cellphone policy**. [SA_P-017-E]

(5.50a) ----- [SA_P-018-E]

(5.50b) Mike was involved in **the construction of uShaka** and has won ... [SA_P-018-E]

The analyses of the types of editorial changes made by editors in different sub-varieties and across different registers show some interesting patterns for the different varieties. Overall, the evidence suggests that AusE editors do not specifically target the genitive constructions in their interventions, and while their interventions do shift the proportions of the two constructions in the academic register in favour of the *s*-genitive and the creative and instructional register in favour of the *of*-genitive, these shifts are largely the by-product of editors either rewriting the sentences in which the constructions occur (triggered by some other issue) or deleting the sentences in which the construction occurs. Because there are no interventions in which either construction is replaced by the other, there seems to be a kind of endorsement by the editors of AusE on how the two variants are used.

The picture in the SAfE sub-varieties is somewhat different, depending on the sub-variety being edited. In AfrE, the editorial focus is split by registers: in the two more formal registers, academic and instructional, editors appear to target the use of the *of*-genitive by AfrE authors by removing these constructions where they are deemed unnecessary or where their use is repetitive or wordy. On the other hand, in the popular and reportage registers, editors of AfrE appear to constrain the use of the *s*-genitive. An analysis of the concordances from these two registers suggests that in these cases, language-internal factors, such as syntactic weight, might influence editorial choices. In WSAfE, editors similarly make many changes, but these appear to be in the opposite direction of editors working on AfrE: in editing texts by WSAfE authors, editors seem to target the underuse of the *of*-genitive by these authors, by (sometimes substantially) increasing the frequency of this construction. In the registers where both of these sub-varieties are represented, it appears that editorial intervention brings the unedited writing of WSAfE and AfrE closer together, something that does not appear to extend to BSAfE. This points to some convergence between WSAfE and AfrE, at least in terms of the use of the genitive constructions. In the academic and instructional BSAfE registers, although editors make substantive additions to and removals from the two texts, the effect of these interventions maintains the proportions of the two constructions, and the very low proportion of *s*-genitives, in comparison to AfrE and WSAfE. This suggests two important findings: first, editors appear to be hesitant to be too interventionist with BSAfE, where most of their interventions relate to the correction of errors made by BSAfE authors. In other words, editors of BSAfE engage less with whether the stylistic expectations of the different registers are met, and concern themselves more with identifying and correcting outright errors in BSAfE. Furthermore, editors validate the proportional uses of the constructions by BSAfE

authors, which provides evidence for the endonormative acceptance of the very low proportional frequency of *s*-genitives in this sub-variety. In the popular and reportage BSAfE registers, editorial intervention boosts the proportion of *s*-genitives, more so in reportage than in the popular register, which suggests that in this register there is less endorsement by the editors for the lower preference for the *s*-genitive. The findings for the differences in the extent of editorial intervention in especially WSAfE and BSAfE link up with those of Kruger (forthcoming), who also found that editors of BSAfE are far more hesitant to make changes in BSAfE than they are in WSAfE.

5.3.3 Statistical evaluation of the interaction of linguistic and language-external factors in genitive alternation in AusE and SAfE

So far the investigation has focused on the influence of the language-external factors of VARIETY and REGISTER on the genitive alternation, and has analysed how editorial intervention shifts the proportions of the two constructions in the varieties and across registers. As discussed in Section 3.4.3.1, linguistic factors also play an important role in genitive alternation. Therefore, the focus of the investigation now shifts to the interaction between the linguistic factors that are known to influence the genitive alternation and the language-external factors.

The analysis will be carried out in two stages. The aim of the first stage is to determine which factors are more important in predicting which genitive construction will be used in unedited and edited writing. The aim of the second stage is to investigate the interaction between the factors shown to be important in conditioning the choice between the two genitive constructions, in order to determine if this interaction differs across the unedited and edited writing of the four sub-varieties investigated.

The first stage of the analysis makes use of random forests (see Section 5.3.3.1). Random forests are non-parametric regression and classification models that seek to predict which of the alternatives of a response variable is most probable given a set of predictors (or independent variables) (Tagliamonte & Baayen, 2012) – see the discussion of random forests in Section 4.3.4 for more detail. Random forests therefore provide an indication of the impact of each predictor variable on the response variable, given all other independent variables. The impact of each independent variable is reflected in its importance score. The interpretation of the importance scores of the different independent variables is straight-forward: the higher the importance score, the stronger the impact of the variable on the response variable.

During the first stage of the analysis, the data are separated by edited status. The aim in doing this is to establish if there are differences in the strength of the effects of the independent variables between unedited and edited writing. A two-pronged approach is used to do this. First, in Section 5.3.3.1.1, random forests are used to explore differences in the overall strength of the various linguistic and language-external factors in the unedited and edited data separately. This analysis helps to determine how editorial intervention influences the strength of the different factors in conditioning the choice of the two constructions. The analysis then zooms in on differences across unedited and edited writing in each of the four sub-varieties

in Sections 5.3.3.1.2 to 5.3.3.1.5, in order to probe how the strength of the factors differs across sub-varieties, and within each sub-variety across unedited and edited writing.

While random forests provide a reliable way of determining what the most important predictors conditioning genitive alternation are, they do not provide much detail on what the interaction between the different factors is in the selection of the two constructions. This can be investigated by constructing a conditional inference tree for the strongest predictors identified through the random forests analysis. The second stage of the analysis, presented in Section 5.3.3.2, therefore investigates the interaction of the most important linguistic and language-external factors using one conditional inference tree per variety in order to explore how the different factors interact with editorial intervention to condition the varying use of two genitive constructions.

5.3.3.1 Random forest analysis of the factors influencing genitive alternation across unedited and edited writing

5.3.3.1.1 Contrastive random forest analysis for the effect of editorial intervention across varieties

The first analysis in this section seeks to identify which factors are the most important predictors for genitive alternation in the unedited and edited texts. The linguistic factors included in the analysis are POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, SEMANTIC RELATION, POSSESSOR DEFINITENESS and FINAL SIBILANCY, while the language-external factors are AUTHOR LANGUAGE (as an indicator of sub-variety), AUTHOR GENDER and REGISTER. Because the analysis is interested in understanding how the different factors influence the choice between the two constructions in unedited and edited writing, and because certain language-external factors are at play in edited writing, but not in unedited writing, two additional language-external factors are included for edited writing: EDITOR LANGUAGE and EDITOR GENDER.

Figure 5.15 and Figure 5.16 reflect the variable importance plots yielded by the random forest analysis for the unedited and edited data respectively. Figure 5.15 shows that all of the predictors, with the exception of AUTHOR GENDER, play a significant role in conditioning the choice between the two genitive constructions in unedited writing (since these predictors have a variable importance score above 0). ANIMACY, POSSESSOR LENGTH, FINAL SIBILANCE and AUTHOR LANGUAGE (or sub-variety) are the four strongest predictors (with importance scores of 0.050, 0.018, 0.008 and 0.008, respectively). The three linguistic factors with the greatest predictive strength identified for unedited writing are also the three factors that are reported to be the most important factors in earlier studies of the genitive alternation (Rosenbach, 2014). Following these three factors, sub-variety (as indicated by AUTHOR LANGUAGE) is the next strongest predictor (with an importance score of 0.008). POSSESSOR DEFINITENESS (0.006) and REGISTER (0.005) are also important predictors, while POSSESSUM LENGTH and SEMANTIC RELATION play a very minor role with an importance score of 0.001 each.

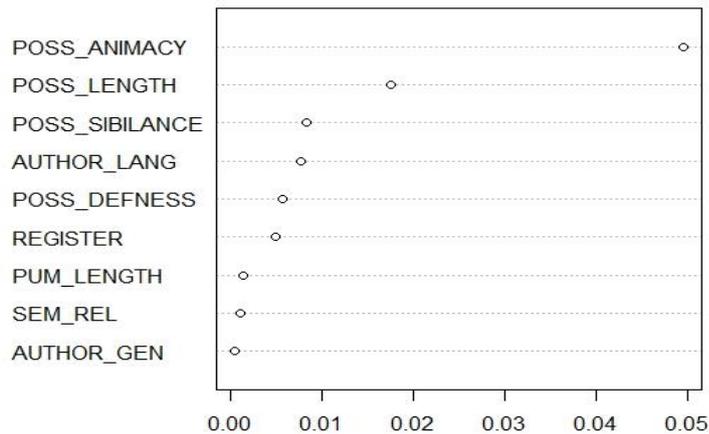


Figure 5.15: Variable importance plot for unedited writing, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables

Figure 5.16, which reflects the variable importance plot for edited writing, shows that ANIMACY, POSSESSOR LENGTH and FINAL SIBILANCE are also the strongest predictors for which genitives are selected in edited writing. The importance score for each of these factors is as follows: ANIMACY (0.042), POSSESSOR LENGTH (0.013) and FINAL SIBILANCE (0.007). The strength of each of these factors is therefore somewhat lower in edited writing than it is in unedited writing, suggesting that editors are somewhat less guided by these factors than authors are. With regard to ANIMACY, the lower importance score for the edited data suggests that because editors are less constrained by this factor, they might be assisting the spread of the *s*-genitive to possessors lower on the animacy scale, potentially implicating them in this ongoing change. Recall from the discussion in Section 3.4.3.1 that POSSESSOR LENGTH and FINAL SIBILANCE are processing and production-driven factors. The lower strength of these factors in edited writing indicates that editors are less influenced by these processing and production constraints. There may be several reasons for this. The first reason relates to the nature of the tasks assumed by editors and authors: authors write the texts, while editors evaluate these texts. The production of texts is a cognitively demanding process, and one could argue that this is more cognitively demanding than evaluating texts. It may therefore be that during more cognitively demanding text-production tasks, the choice of construction is more strongly influenced by processing and production constraints. The lower strength of these factors might also be due to the experience of the person involved in the particular task: for example, editors spend their workday reading through and processing written material (which may not necessarily be the case for authors). Given this more frequent exposure to texts, it may be that editors are less guided by the constraints imposed by long possessors and the presence of a final sibilant on the possessor, and as a result, their choice of genitive construction is less strongly influenced by these processing- and production-driven factors, compared to authors. How this plays out in terms of which feature is selected, is not yet clear, but is explored in more detail in Section 5.3.3.2.

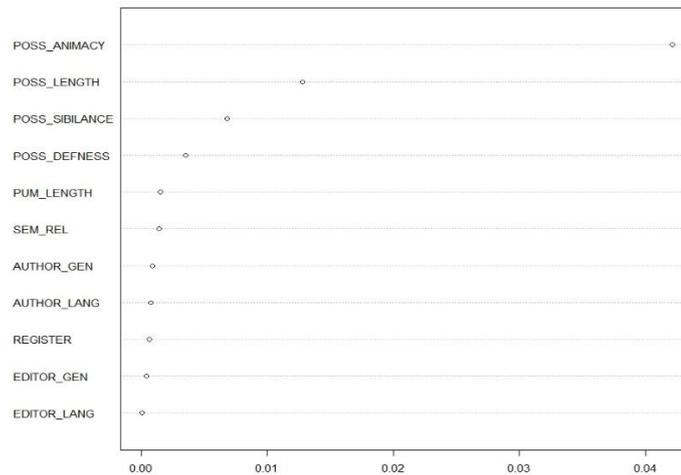


Figure 5.16: Variable importance plot for edited writing, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER, AUTHOR LANGUAGE and EDITOR LANGUAGE as predictor variables.

Two further striking decreases in importance scores in edited writing are for REGISTER and AUTHOR LANGUAGE (which indicates the different sub-varieties). REGISTER is a relatively more important predictor in unedited writing, with an importance score of 0.005; however, in edited writing this factor decreases in strength, so that it is only marginally important in predicting which construction will be used in edited writing, where the importance score is 0.001. REGISTER exerts less influence than other linguistic factors in conditioning genitive alternation in edited writing. Perhaps the most striking finding is for the effect of AUTHOR LANGUAGE. AUTHOR LANGUAGE, that is the sub-variety to which the author belongs (represented as four levels in this study: AusE, AfrE, BSAfE and WSAfE), is the fourth strongest predictor in unedited writing with an importance score of 0.008, but this factor plays no role in conditioning which construction is selected in edited writing. This shows that in edited writing, an author's language does not play an important role in predicting which genitive construction will be used. The overall effect of editorial intervention therefore appears to be to smooth out differences across the different sub-varieties and (to a lesser extent) the registers. In other words, editing appears to have a homogenising effect on varietal and register differences. Furthermore, the conditioning strength of all the linguistic factors is reduced in edited writing, (most notably for ANIMACY, POSSESSOR LENGTH, FINAL SIBILANCE and POSSESSUM LENGTH). Edited writing is therefore less strongly influenced by the various linguistic factors, and therefore editorial intervention might help to bolster the trends associated with the ongoing increase in the use of the *s*-genitive. To explore this in more detail, the analysis now shifts focus to differences across unedited and edited tokens in each of the four sub-varieties in order to establish how the strength of the factors differs across sub-varieties, and within each sub-variety across unedited and edited writing.

5.3.3.1.2 Contrastive random forest analysis for the effect of editorial intervention in AusE

Figure 5.17 and Figure 5.18 reflect the variable importance plots for the factors that condition the choice between the two genitive constructions in unedited and edited AusE, respectively. Figure 5.17 shows that linguistic factors as a group are more important factors than language-external ones for genitive alternation in unedited AusE, followed by the language-external factors. POSSESSOR ANIMACY is by far the most important predictor, with an importance score of 0.063. POSSESSOR LENGTH is also important (with an importance score of 0.013), but it is a much less strong predictor than animacy. FINAL SIBILANCE (0.006) is the third strongest factor, but is a much less important predictor than POSSESSOR LENGTH and ANIMACY. POSSESSOR DEFINITENESS (0.003), and POSSESSUM LENGTH, REGISTER and AUTHOR GENDER, which have the same importance score of 0.002, are the least strong predictors for the choice between the two constructions.

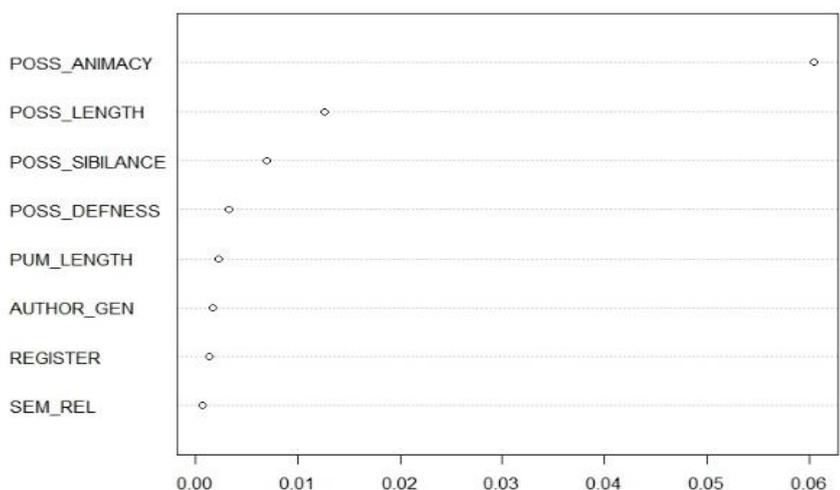


Figure 5.17: Variable importance plot for unedited AusE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables

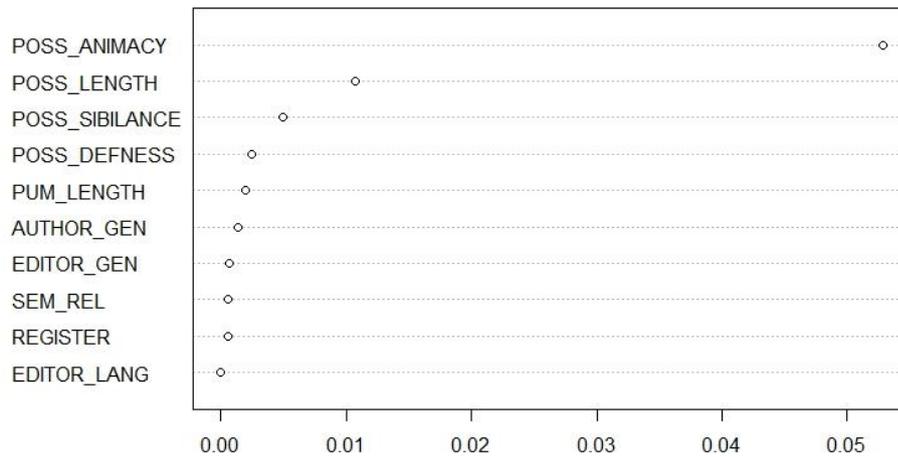


Figure 5.18: Variable importance plot for edited AusE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITOR LANGUAGE as predictor variables

A similar pattern emerges for edited AusE in Figure 5.18: linguistic factors as a group are more important predictors for the choice of genitive construction than language-external factors. The effect of editing in AusE appears to be to reduce the importance scores of the linguistic factors that have the strongest predictive power. For example, POSSESSOR ANIMACY (0.053) is the strongest predictor of genitive alternation in edited AusE, but the importance score of this factor is decreased in edited writing. The strength of POSSESSOR LENGTH (0.011) and FINAL SIBILANCE (0.005) are also weakened in edited AusE, though to a lesser degree than ANIMACY. This demonstrates that although these three factors remain important predictors for which genitive will be selected by AusE editors, editors of AusE are less constrained by them than AusE authors. A loosening of the animacy constraint for possessors lower on the animacy scale is one of the reasons most often provided for the ongoing increase in the *s*-genitive, and it therefore seems that editors of AusE are helping this change along by weakening the strength of animacy in conditioning the choice between the two constructions in the editing process.

Shifts in the importance scores of the remaining factors are all minor by comparison, but it is worth noting that the already low importance score of REGISTER in unedited AusE is further weakened in edited AusE from 0.002 to 0.001, which means that editing leads to more uniform choices across registers. As noted in Section 3.2.1, AusE has reached Phase 5 of the Dynamic Model, and as it continues to undergo normal processes of ongoing change it is likely to distinguish itself from other varieties, not only in its modes of discourse but also in its registers. Section 3.2.1 argued that the pervasive informality, a well-known feature of AusE, has started creeping into even the most formal of written registers, academic writing, as evidenced by the unusually high proportion of *s*-genitives in this register (in comparison to other varieties). It therefore seems that this persistent informal and colloquial character of AusE is narrowing the gap between the different written registers – here, in the direction of a more uniformly informal or colloquial style. The finding that REGISTER plays a minor role in conditioning genitive alternation across unedited and edited

AusE writing, bears this out, but it is noteworthy that AusE editors further reduce the strength of REGISTER in conditioning genitive alternation, which suggests that these editors might be helping to reinforce the narrowing of the differences across written registers, providing clear evidence of endonormativity.

5.3.3.1.3 Contrastive random forest analysis for the effect of editorial intervention in WSAfE

Figure 5.19 and Figure 5.20 reflect the variable importance plots for the factors conditioning genitive alternation in unedited and edited WSAfE respectively. Figure 5.19 shows that with the exception of AUTHOR GENDER all factors are important predictors for genitive alternation in unedited WSAfE. Two linguistic factors and one language-external factor, namely POSSESSOR ANIMACY, POSSESSOR LENGTH and REGISTER, have the most important predictive power for whether the *s-* or *of-*genitive is selected. POSSESSOR ANIMACY is the strongest predictor, with an importance score of 0.052, followed by POSSESSOR LENGTH (0.026), and REGISTER (0.014). POSSESSOR LENGTH and REGISTER clearly have a relatively stronger effect in WSAfE, compared to AusE. POSSESSOR DEFINITENESS and FINAL SIBILANCE are also important factors conditioning genitive alternation in unedited WSAfE, with near similar importance scores of 0.007 and 0.006 respectively. SEMANTIC RELATION and POSSESSUM LENGTH play less important roles (with variable importance scores of 0.003 and 0.001 respectively).

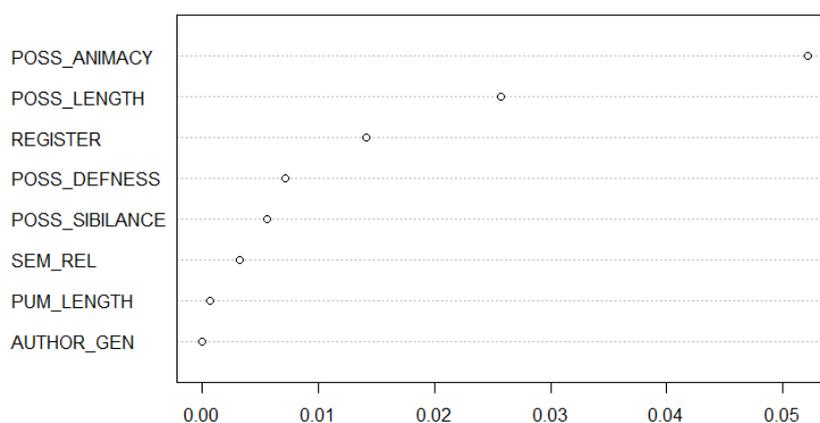


Figure 5.19: Variable importance plot for unedited WSAfE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables.

The picture in edited WSAfE changes somewhat from that in unedited WSAfE, mostly due to some striking shifts in the importance scores of the most important factors. The most notable of these is the drastic reduction in the variable importance score for REGISTER (0.001), which clearly plays a minimal role in conditioning genitive alternation in edited WSAfE, compared to its much stronger role in unedited WSAfE.

When compared to AusE, this shows that there are rather substantial differences in how this factor predicts which genitive construction will be selected by authors, but this difference is reduced by the weakening of this factor in edited WSAfE. This suggests that editors of WSAfE make changes that align usage in the registers away from a local (covert) norm. It should be noted though, that the lower importance score of REGISTER is most likely due to changes made by editors in the creative register noted in Section 5.3.2.1, where the proportional frequencies of the two constructions were reversed.

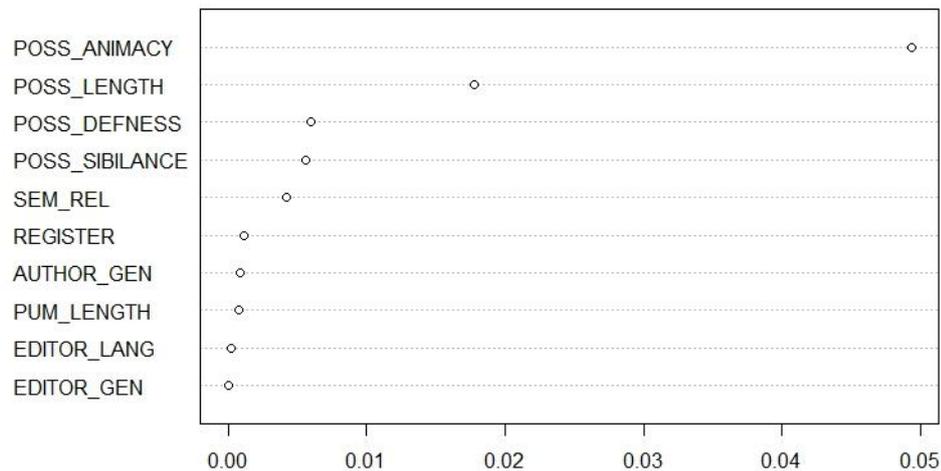


Figure 5.20: Variable importance plot for edited WSAfE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITOR LANGUAGE as predictor variables.

The importance scores for ANIMACY (0.049) and POSSESSOR LENGTH (0.018) are also decreased in edited WSAfE, suggesting that editors of this sub-variety are less constrained by these factors than WSAfE authors are. Why this may be is not clear, but the analysis of the interaction of factors in Section 5.3.3.2 sheds some light on this.

The strength of POSSESSOR DEFINITENESS (0.006) is marginally decreased in edited WSAfE, bringing it in line with FINAL SIBILANCE (0.006), which remains unchanged. While these two factors do play a role in conditioning genitive alternation in unedited and edited WSAfE, they are less strong factors, compared to ANIMACY and POSSESSOR LENGTH. Lastly, the importance scores for SEMANTIC RELATION (0.004) and AUTHOR GENDER (0.001) are increased slightly, while the language-external factors, EDITOR LANGUAGE and EDITOR GENDER, play no role in predicting which of the two constructions will be used in edited WSAfE. With regard to EDITOR LANGUAGE, the unimportance of this factor in edited WSAfE is in all likelihood due to the fact that 73% of the editors of WSAfE are WSAfE users themselves.

5.3.3.1.4 Contrastive random forest analysis for the effect of editorial intervention in AfrE

Figure 5.21 and Figure 5.22 reflect the variable importance plots for the factors that condition the choice between the two genitive constructions in unedited and edited AfrE, respectively. Figure 5.21 shows that, as a group, linguistic factors are the strongest predictors for the choice of genitive construction in unedited AfrE, compared to language-external factors. POSSESSOR ANIMACY is the strongest predictor in unedited AfrE, with an importance score of 0.040 (which is relatively lower than the importance score for unedited AusE and WSAfE). The weaker predictive power of animacy in AfrE supports the argument by Rosenbach (2017) that the animacy constraint in genitive alternation is weaker in AfrE than it is for first-language varieties, including WSAfE, and that this is due to an effect of substrate transfer.¹⁸

FINAL SIBILANCE and POSSESSOR DEFINITENESS are the next strongest factors (with an importance score of 0.014 each), closely followed by POSSESSOR LENGTH (0.013). In comparison to the two first-language varieties, FINAL SIBILANCE plays a relatively more important role in conditioning genitive alternation in unedited AfrE. At this stage it is not clear why this may be, but the most likely explanation seems to be that the presence of a final sibilant places additional processing pressure on second-language users. It may also be that AfrE authors struggle to decide where to place the apostrophe for the *s*-genitive with proper nouns ending in *s* and common plural nouns. It is not clear if this is the case, but the analysis of the interaction of factors in Section 5.3.3.2 provides more detailed analysis.

REGISTER has a relatively low importance score at 0.004, suggesting that REGISTER does play a (minor) role in conditioning the choice between the *s*- and the *of*-genitive in AfrE.

¹⁸ Rosenbach (2017) compares the genitive choices for inanimate possessors of AfrE users with those of WSAfE and BrE users, and finds that AfrE users are more likely to use the *s*-genitive in contexts where WSAfE and BrE users would use *of*-genitives. As noted in Section 3.4.3, she argues that this is most likely due to a substrate transfer effect from Afrikaans (where the animacy constraint is weaker).

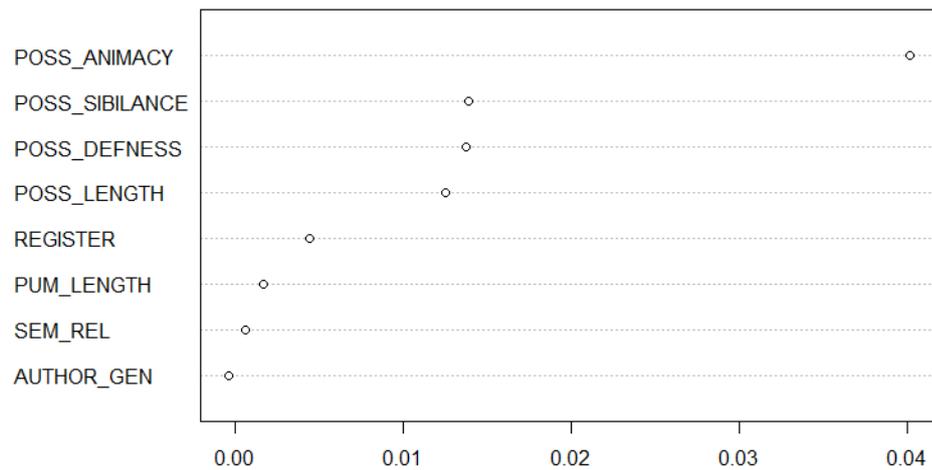


Figure 5.21: Variable importance plot for unedited AfrE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables.

Figure 5.22, which shows the variable importance plot for edited AfrE, reveals that the importance scores for POSSESSOR ANIMACY (0.030), FINAL SIBILANCE (0.010), POSSESSOR DEFINITENESS (0.006), POSSESSOR LENGTH (0.006) and REGISTER (0.001) are all lowered in edited AfrE, but remain the most important predictors in conditioning genitive alternation. Only one factor's importance score is increased, but only marginally: this factor is POSSESSUM LENGTH (0.003).

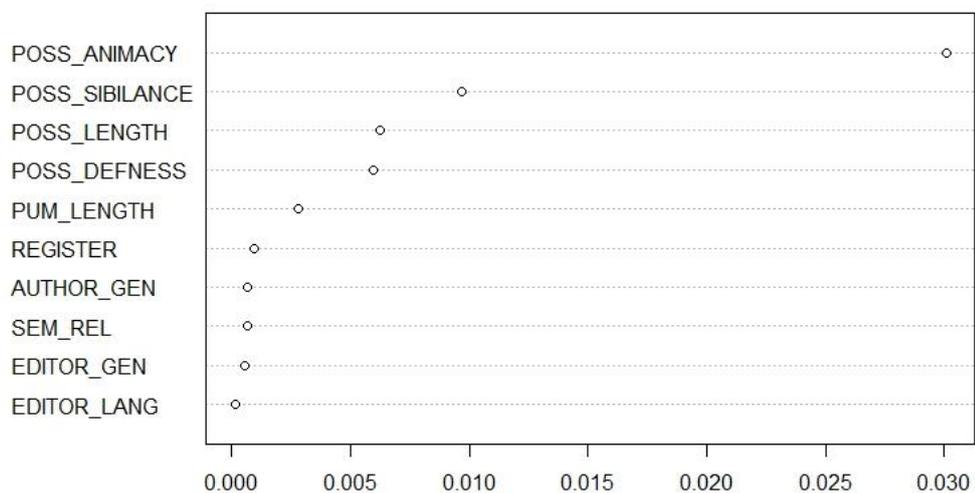


Figure 5.22: Variable importance plot for edited AfrE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER, and EDITOR LANGUAGE as predictor variables.

POSSESSOR ANIMACY is the strongest predictor in edited AfrE, but its importance score is substantively lower in edited AfrE, compared to unedited AfrE. A manual investigation of the data for the languages of the editors of AfrE reveals that approximately 61% of these editors are WSAfE users, while only 4% are

AfrE users (with the background of the remaining 35% unknown). It therefore does not seem that the weaker effect of animacy in edited AfrE can be accounted for by AfrE editors modelling their genitive choices in English on the basis of their first language. Section 5.2.2 demonstrated that most editors of English texts in South Africa are balanced English–Afrikaans bilinguals and it is therefore likely that the WSAfE editors who work on AfrE are also proficient in Afrikaans. This, coupled with the high degree of contact between English and Afrikaans in South Africa, opens up the opportunity for the usage practices of one language to influence another. In other words, the lowering of the strength of animacy as predictor in edited AfrE might point to an influence of Afrikaans on WSAfE users, not only for AfrE writers (as demonstrated in the unedited data) but also for bilingual WSAfE editors who appear to be reinforcing and bolstering the genitive preferences of Afrikaans in the English texts that they edit.

The importance scores of FINAL SIBILANCE (0.010) and POSSESSOR DEFINITENESS (0.006) are also lowered in edited AfrE, and it seems likely that this is due to WSAfE editors bringing the strength of these constraints more in line with strengths of the constraints in WSAfE. POSSESSOR LENGTH and POSSESSUM LENGTH are the fourth and fifth strongest predictors with importance scores of 0.006 and 0.003 respectively. The strength of POSSESSOR LENGTH is therefore reduced, while POSSESSUM LENGTH is increased.

The linguistic factors POSSESSOR ANIMACY, POSSESSOR SIBILANCE, POSSESSOR DEFINITENESS and POSSESSOR LENGTH are the most important predictors of which genitive will be selected in unedited and edited AfrE; however, the strength of these four factors is reduced in edited AfrE. Furthermore, the language-external factor REGISTER is shown to have some predictive strength in conditioning whether the *s-* or *of-*genitive will be selected in unedited AfrE, but as with edited AusE and WSAfE the strength of this factor is reduced in edited AfrE.

5.3.3.1.5 Contrastive random forest analysis for the effect of editorial intervention in BSAfE

The variable importance plots in Figure 5.23 and Figure 5.24 are for unedited and edited BSAfE, respectively. Figure 5.23 shows that only three factors play a notable role in conditioning which genitive is selected in unedited BSAfE, and these are all linguistic factors. Unedited BSAfE stands out as the sub-variety where genitive choice is conditioned by the fewest factors, and it is the only sub-variety in which the language-external factors play no role. POSSESSOR ANIMACY is the strongest predictor for the choice between the *s-* and the *of-*genitive construction, with an importance score of 0.048. POSSESSOR LENGTH follows closely behind (with a notably high importance score of 0.034 – which mirrors the score for the factor in WSAfE). The high importance score of this factor in unedited BSAfE should probably not be attributed to BSAfE authors' status as second-language users (as evidenced by the lower importance score for this factor for AfrE authors, and the high importance score for this factor for WSAfE). It therefore appears that there is some convergence between BSAfE and WSAfE in terms of how possessor length

governs the choice between the genitive constructions, but this can only be confirmed by the interaction analysis (see Section 5.3.3.2.4). FINAL SIBILANCE is the third strongest predictor (0.002), but this factor plays a minor role in comparison to POSSESSOR ANIMACY and POSSESSOR LENGTH. BSAfE authors are therefore not as strongly affected by this constraint as WSAfE authors are, but this is in all likelihood due to the overwhelming preference for the *of*-genitive by BSAfE authors demonstrated in Section 5.3.2. None of the other factors plays a significant role in predicting which of the two constructions will be used by BSAfE authors.

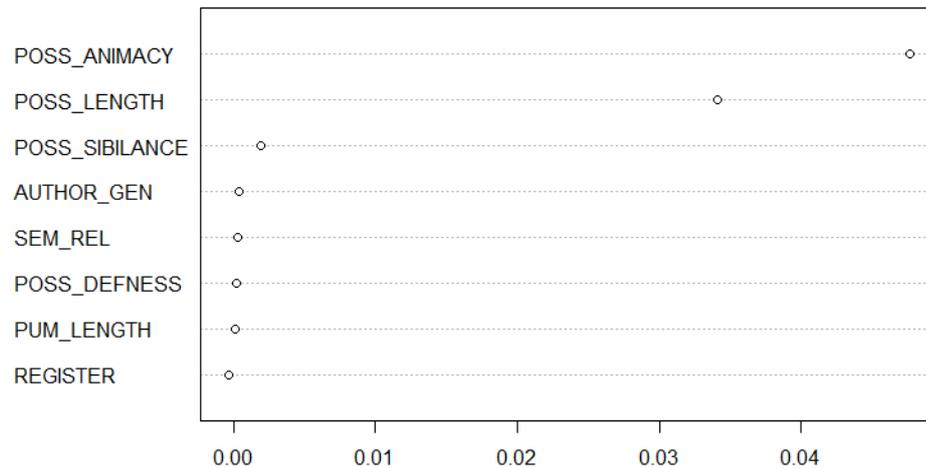


Figure 5.23: Variable importance plot for unedited BSAfE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER and AUTHOR GENDER as predictor variables.

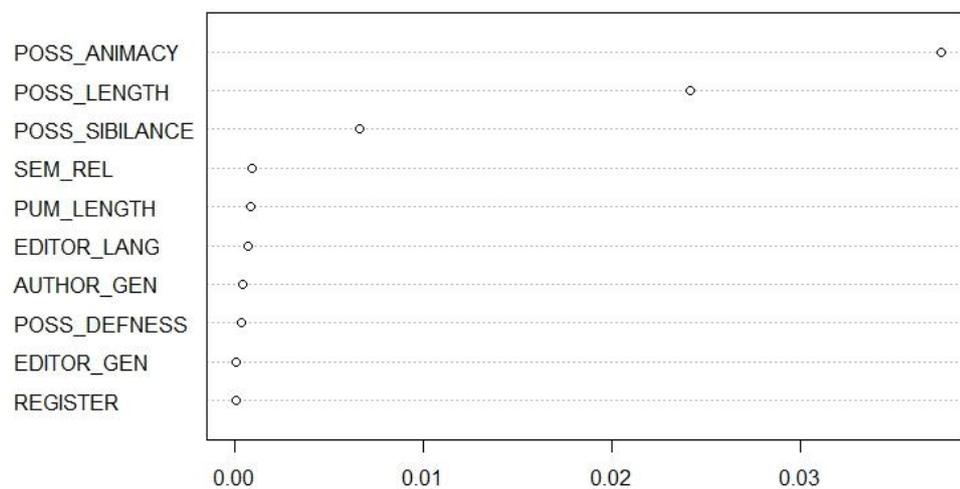


Figure 5.24: Variable importance plot for edited BSAfE, with POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITOR LANGUAGE as predictor variables.

Figure 5.24 reflects the variable importance plot for edited BSAfE, and reflects the extent of editorial intervention in influencing genitive alternation in this sub-variety highlighted in Section 5.3.2.3. While only three factors were shown to condition genitive alternation in unedited BSAfE, six play a role in edited BSAfE. POSSESSOR ANIMACY remains the strongest predictor in edited BSAfE, but the strength of this factor is weaker in edited BSAfE (as reflected in the lower importance score of 0.037). The importance score of POSSESSOR LENGTH is also weaker (0.024), though its importance score remains high. In their linguistic choices, editors of BSAfE therefore reduce the effects of the constraints of possessor animacy and possessor length, while their choice of genitive construction is more strongly guided by FINAL SIBILANCE (0.007) than BSAfE authors are. These changes may be due to the EDITOR LANGUAGE, which is shown to play a minor role in edited BSAfE. An analysis of the edited BSAfE observations shows that 59% of these were edited by WSAfE editors, while 4% were edited by AfrE editors (with the background of the remaining 37% of editors unknown).

Furthermore, two linguistic factors (SEMANTIC RELATION and POSSESSUM LENGTH) that played no role in conditioning genitive alternation in unedited BSAfE, do play minor roles in edited BSAfE. Given that these factors are shown to play a role in unedited and edited AfrE and WSAfE, it is also likely that these factors shape the choices of AfrE and WSAfE editors in their editing of BSAfE writing. REGISTER, AUTHOR GENDER and EDITOR GENDER play no role in predicting which of the two constructions occur in edited BSAfE writing.

5.3.3.1.6 Summary of findings for the contrastive random forest analysis

The discussion of the variable importance strengths of the predictor variables across the unedited and edited writing of the sub-varieties shows that there are striking similarities, but also differences, across the sub-varieties in terms of the factors that are important in unedited and edited writing. When considering the findings for the SAfE sub-varieties, there is convincing evidence to support Van Rooy's (2014a) view that in settings of high contact among many different groups of users, it is important to study the linguistic choices at the level of the local communities, rather than to generalise to the country as a whole. For example, unedited BSAfE stands out as the one sub-variety in which the fewest number of factors play an important role in conditioning genitive alternation, and it is also the only sub-variety in which language-external factors, such as REGISTER and AUTHOR GENDER, do not play any role in predicting which construction will be selected. REGISTER is the third strongest predictor in unedited WSAfE, but plays a very minor or no role in the other sub-varieties. However, editors of WSAfE weaken the strength of this factor considerably, which is the most obvious example of how editorial intervention influences usage practices and works to limit extreme differences across registers. ANIMACY is found to be the strongest predictor in all sub-varieties and across unedited and edited writing; however, the relative strength of this factor is different across the sub-varieties. For all remaining factors, the order (and inclusion) of factors differs across sub-varieties and unedited and edited writing. For example, SEMANTIC RELATION is of minor importance in unedited AfrE and WSAfE, but it plays no role in unedited AusE and BSAfE, and it

is only through editing that this factor comes to play a role in AusE and BSAfE. EDITOR LANGUAGE is unimportant in edited AusE, AfrE and WSAfE, but plays a minor role in BSAfE, and most likely accounts for the inclusion of three additional factors in conditioning genitive alternation in the edited writing of this sub-variety.

The variable importance scores of a number of factors are changed in the edited writing of the different sub-varieties and warrant further discussion (and comparison). Table 5.36 reflects a summary of the variable importance scores for each factor in the different varieties across unedited and edited writing respectively. The ranking of each factor is given in brackets. For the purposes of the discussion to follow, only the most striking differences across sub-varieties will be discussed.

	Unedited				Edited			
	AusE	AfrE	BSAfE	WSAfE	AusE	AfrE	BSAfE	WSAfE
Animacy	0.063 (1)	0.040 (1)	0.048 (1)	0.052 (1)	0.053 (1)	0.024 (1)	0.037 (1)	0.049 (1)
Possessor length	0.013 (2)	0.013 (4)	0.034 (2)	0.026 (2)	0.011 (2)	0.005 (4)	0.024 (2)	0.018 (2)
Final sibilance	0.006 (3)	0.014 (2)	0.002 (3)	0.006 (5)	0.005 (3)	0.008 (2)	0.007 (3)	0.006 (4)
Possessor defness	0.003 (4)	0.014 (3)	0.000	0.007 (4)	0.002 (4)	0.006 (3)	0.000	0.006 (3)
Possessum length	0.002 (5)	0.002 (6)	0.000	0.001 (7)	0.002 (5)	0.002 (5)	0.001 (5)	0.001 (8)
Semantic relation	0.000	0.001 (7)	0.000	0.003 (6)	0.001 (8)	0.001 (7)	0.001 (4)	0.004 (5)
Register	0.002 (6)	0.004 (5)	0.000	0.014 (3)	0.001 (9)	0.001 (6)	0.000	0.001 (6)
Author gender	0.002 (7)	0.000	0.000	0.000	0.001 (6)	0.000	0.000	0.001 (7)
Editor gender	–	–	–	–	0.001 (7)	0.000	0.000	0.000
Editor language	–	–	–	–	0.000	0.000	0.001 (6)	0.000

Table 5.36: Summary of the variable importance scores

The importance scores for ANIMACY, POSSESSOR LENGTH and the language-external factor REGISTER are consistently lowered in the edited writing of all sub-varieties. The importance score for FINAL SIBILANCE decreases in edited AusE, AfrE and WSAfE, but increases substantially in edited BSAfE. POSSESSOR DEFINITENESS plays a minimal role in predicting which construction will be selected in unedited AusE and no role in unedited BSAfE, which is maintained for both sub-varieties in edited writing; however, this factor plays a slightly more important role in unedited WSAfE and AfrE. The strength of all other factors is generally maintained across the unedited and edited writing of the sub-varieties.

As already noted, the strength of ANIMACY as predictor variable is reduced considerably in the edited writing of all sub-varieties and most notably in edited AusE, AfrE and BSAfE. In unedited writing, the variable importance score for ANIMACY is highest in AusE (0.063), then WSAfE (0.052) and BSAfE (0.048), which have near similar importance scores, and lowest in AfrE (0.040). AfrE therefore has the lowest importance score for ANIMACY of the four sub-varieties. This reiterates the findings by Rosenbach (2017) that genitive alternation in AfrE is less constrained by animacy, which is most likely due to a substrate transfer effect of the Afrikaans first language. The variable importance scores for this factor are all lower in edited writing, but the greatest difference in the strength of this factor across unedited and edited writing occurs in AusE, AfrE and BSAfE, where the variable importance scores are 0.053, 0.049 and 0.037, respectively. This demonstrates two very important effects. First, during the editing task, editors'

choices are less constrained by animacy in genitive alternation and in this way they are helping to spread the *s*-genitive to possessors lower on the animacy scale, suggesting a facilitating role for editors in this ongoing change. Second, while unedited AfrE has the lowest importance score for animacy of the four sub-varieties, editors (who are mostly WSAfE users) reinforce and bolster this pattern in edited AfrE.

POSSESSOR LENGTH is the second strongest factor in unedited and edited AusE, BSAfE and WSAfE, but the fourth strongest factor in unedited AfrE. The variable importance score for this factor is the same for unedited AusE and AfrE (0.013), but comparably higher in unedited BSAfE (0.034) and WSAfE (0.026), suggesting that possessor length exerts greater predictive power in conditioning genitive alternation in the writing of BSAfE and WSAfE authors than it does for AusE and AfrE authors. The strength of this factor is weaker in edited AfrE (0.006), BSAfE (0.024) and WSAfE (0.018), but only minimally weaker in edited AusE (0.011), so that in edited writing, the importance score of this factor is lowest in AfrE. At this stage it is unclear why this is the case, but the conditional inference tree analysis for each sub-variety in Section 5.3.3.2 provides more insights.

The presence of a FINAL SIBILANT in the possessor is the second strongest predictor in the choice of genitive construction in unedited AfrE, the third strongest in unedited AusE and BSAfE and the fifth strongest in WSAfE. The presence of a final sibilant on the possessor therefore plays a much more important role for AfrE authors than for AusE and WSAfE authors, and BSAfE authors' selection of genitive construction is only minimally influenced by this production-related factor. With regard to BSAfE and AfrE, the comparably different strengths of the influence of this factor might be linked to the effect of substrate transfer or second language status. For example, in BSAfE the much higher use of the *of*-genitive compared to the *s*-genitive noted in Section 5.3.2 is in all likelihood due to substrate influence and the fact that in the indigenous South African languages, the only genitive construction available is one in which the possessor is realised in postnominal position. This means that the presence of a final sibilant in the possessor might not actually matter in BSAfE writing, where authors are more likely to select the *of*-genitive construction more often simply because this is the construction available in the substrate language. Unlike BSAfE, there are two genitive constructions in Afrikaans that are very similar to the ones in English (see discussion in Section 3.5). According to Redelinguys (forthcoming), final sibilance plays only a minor role in genitive alternation in Afrikaans, and so it may be that the stronger influence of this factor for these second-language users of English results from the production effort associated with using a possessor with a final sibilant in an *s*-genitive. Moreover, in Afrikaans apostrophes are not used in genitive constructions, but are used for other purposes (such as showing contraction). It might be that the use of an apostrophe in the relation marker of the *s*-genitive causes heightened strain for AfrE authors, who may use avoidance of the *s*-genitive as a strategy to mitigate this. The importance score of this factor is maintained in edited WSAfE (0.006), weakened in edited AusE (0.005) and, to a much greater extent, in AfrE (0.010), while it is increased substantially in BSAfE (0.007). The effect of these changes is to bring the factor strengths of the SAfE sub-varieties, in particular, much closer to one another.

POSSESSOR DEFINITENESS plays no role in the genitive choices of BSAfE authors, and the reasons for this may be the same as the reasons provided for the very low importance score for FINAL SIBILANCE for unedited BSAfE. POSSESSOR DEFINITENESS does, to varying degrees, predict which construction will be selected by AusE, AfrE and WSAfE authors. The importance score of this factor is considerably higher in unedited AfrE (0.014), compared to WSAfE (0.007) and AusE (0.003), suggesting that this factor is a stronger predictor for the genitive choices of AfrE authors. The strength of this factor is marginally weaker in edited AusE (0.002) and WSAfE (0.006), but weakened substantively in edited AfrE (0.006), again bringing the two South African varieties much closer together.

REGISTER is not an important factor in predicting genitive choice in either unedited or edited BSAfE (0.000). The unimportance of this factor in unedited and edited BSAfE is in all likelihood due to the overwhelming preference for the *of*-genitive construction as a result of substrate transfer, which editors leave unchanged. While register has a minimal influence in unedited AusE (0.002) and AfrE (0.002), it has a strong influence in unedited WSAfE (0.014). This influence is in all probability due to the higher proportion of *s*-genitives in unedited creative writing by WSAfE authors, and the comparably lower proportion in unedited academic, instructional and popular writing. The strength of this factor is reduced only marginally in AusE (0.001) and AfrE (0.001), but substantially in WSAfE (0.001), which, again, brings the two SAfE sub-varieties closer together.

The findings of this first analysis therefore show that while some factors (such as ANIMACY, POSSESSOR LENGTH, FINAL SIBILANCE and POSSESSOR DEFINITENESS) play a role in conditioning the genitive alternation in all four sub-varieties, other factors do not. Furthermore, the analysis shows that overall, the most important factors are weakened in edited writing, while the strength of many are maintained. This points to a complex pattern of interaction for the different factors in each variety. These interactions cannot be identified using random forests analysis, and thus, having identified the most important factors in conditioning genitive alternation in the different sub-varieties investigated, the focus now moves to an investigation of the interaction of these factors in conditioning the varying use of the *s*- and *of*-genitive constructions, using conditional inference trees.

5.3.3.2 Conditional inference tree analysis of the interaction between linguistic and language-external factors in influencing genitive alternation

In Section 5.3.3.1, it was established that there is a complex pattern in the interaction of the different linguistic and language-external factors in each variety separated by edited status, and that the predictive strength of a number of factors differs across unedited and edited writing. Given these differences and the fact that a number of factors play differing roles in the four sub-varieties, this section considers in more detail the interaction of the different factors in the different sub-varieties. The analysis that follows therefore presents one conditional inference tree per sub-variety that includes only those factors that were shown to be important predictors in genitive alternation in the sub-variety. To avoid watering down the analyses

with multiple complex interactions for only a few observable cases, the approach to the analysis of the conditional inference trees in this section follows Szmrecsanyi et al. (2016b) and presents a discussion of the five uppermost branches of each conditional inference tree.

5.3.3.2.1 Conditional inference tree analysis of the interaction of linguistic and language-external factors in AusE

Figure 5.25 presents the results of the conditional inference tree analysis for AusE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER and EDITED STATUS on the response variable (which has two levels, *s*-genitive and *of*-genitive).

The first noteworthy finding is that EDITED STATUS does not feature in the conditional inference tree for AusE at all: genitive alternation is therefore not affected in any significant way by editorial intervention. This confirms, again, that editors of AusE accept the genitive usage patterns of authors of AusE, and therefore reinforce these usage patterns. In this case, the usage patterns show a strong favouring of the *s*-genitive (as compared to other varieties), which is in line with the generally held view that AusE is receptive to features associated with colloquialisation and informalisation (see Section 3.2.1). This provides evidence of endonormativity for this informality in AusE, at least in terms of genitive alternation. Furthermore, REGISTER plays only a minimal role in conditioning which genitive is used, which suggests that in AusE register preferences for the two genitive constructions are smoothed out, essentially moving the registers closer together. This confirms the findings for the proportional and normalised frequencies of this variety across unedited and edited writing in Section 5.3.2, which showed that although editors do make changes that add and remove the two constructions, these changes do not influence the proportions of the two constructions such that the choices between authors and editors of AusE differ in any significant way. This finding therefore holds even when linguistic factors are accounted for.

When linguistic factors are included alongside language-external factors in an analysis of the genitive alternation in AusE, it is the linguistic factors that most strongly account for alternation between the two constructions, rather than the language-external factors. The most important conditioning variable (as reflected in the first split at node 1) is POSSESSOR ANIMACY, with animate possessors partitioned in the left branch, and inanimate, collective, temporal and locative possessors grouped together in the right branch. The conditional inference tree demonstrates that the *s*-genitive is preferred more than twice as often as the *of*-genitive with animate possessor nouns, with 14 of the 22 end nodes of this branch showing a preference for the *s*-genitive. Collective, temporal, locative and inanimate possessors are further split at node 45 for POSSESSOR ANIMACY, with inanimate nouns partitioned in the left branch and collective, temporal and locative possessors partitioned to the right. All of the end nodes for inanimate possessors show a stronger preference for the *of*-genitive, while *s*-genitives are preferred in 3 of the 14 end nodes for collective, locative

and temporal possessors. This shows that the animacy constraint is quite strong in AusE (as reflected in the variable importance scores for this factor in Section 5.3.3.1), with nouns lower on the animacy scale being realised with *of*-genitives much more often than they are with *s*-genitives, and with inanimate nouns being exclusively realised by the *of*-genitive.

With regard to the animate possessors in the left branch of node 1, the first important factor to interact with animacy is POSSESSOR LENGTH at node 2: the *s*-genitive is favoured with animate possessors that are short (containing two or fewer words) – as evidenced by 12 of the 16 end nodes showing a higher proportion of *s*-genitives; while the *of*-genitive is favoured with animate possessors that are longer than two words in length. The next strongest factor is FINAL SIBILANCE. The branches to the left of node 3 show that the *s*-genitive is favoured with short animate possessors, irrespective of whether they end in a final sibilant or not. Therefore, the *s*-genitive is strongly favoured with animate possessors that are shorter than two words in length.

FINAL SIBILANCE also plays a role in cases of animate possessors that are longer than two words in length. The right branch of node 34 shows that the *of*-genitive is the preferred construction for animate possessors longer than two words in length and that end in a final sibilant (node 43, $n = 15$; node 44, $n = 64$). Furthermore, the *s*-genitive is preferred with animate possessors that do not contain a final sibilant and which are two to three words in length (node 36, $n = 126$); in all other cases, the *of*-genitive is favoured. This shows a clear interaction between ANIMACY, POSSESSOR LENGTH and FINAL SIBILANCE, and demonstrates that the strength of the animacy constraint for animate nouns is tempered by the length of the possessor and the presence of a final sibilant on the possessor: when animate possessors are longer than two words in length and end in a sibilant, then the *of*-genitive is favoured.

With regard to the right branch of node 1, genitives with collective, locative, temporal and inanimate nouns are further split according to animacy at node 45, so that inanimate possessors are partitioned to the left of node 45, while collective, temporal and locative possessors are partitioned to the right. An analysis of the end nodes for all branches of inanimate possessors shows that these possessors overwhelmingly prefer *of*-genitives. The only contexts in which the choice of genitive with these nouns prefers the *s*-form is when the nouns are short definite proper nouns in academic, creative and reportage writing (node 54, $n = 37$), or when they are short definite nouns or proper nouns, edited by female editors in instructional and popular writing (node 56, $n = 143$). With collective, temporal and locative possessors, the preference is overwhelmingly in favour of the *of*-genitive.

Overall, the analysis of the interaction of the different linguistic and language-external factors in AusE shows EDITED STATUS does not play any role in conditioning which genitive is selected, and register only plays a minor role lower down in the conditional inference tree. The proportional frequencies of the two constructions in Section 5.3.2.1 bear this out, with minimal shifts across unedited and edited writing in different registers. Linguistic factors are therefore the most important conditioning variables in the choice

of genitive construction in AusE writing. POSSESSOR ANIMACY is the strongest predictor of genitive choice. Animate possessors (which account for 2 206 of all cases) overwhelmingly favour the *s*-genitive, while inanimate possessors (which account for 2 468 of all cases) favour the *of*-genitive. This means that the high number of animate possessors helps to boost the higher proportional preference for the *s*-genitive noted in Section 5.3.1. This is slightly tempered by collective, temporal and locative possessors which prefer the *of*-genitive, except in a few cases, where they do not have a final sibilant, are definite or proper nouns and have short possessors and/or possessums. This suggests that the extension of the *s*-genitive to nouns lower on the animacy scale remains fairly restricted in AusE, and that the higher proportion of *s*-genitives noted for the sub-variety is due to the high number of animate possessors that occur in AusE.

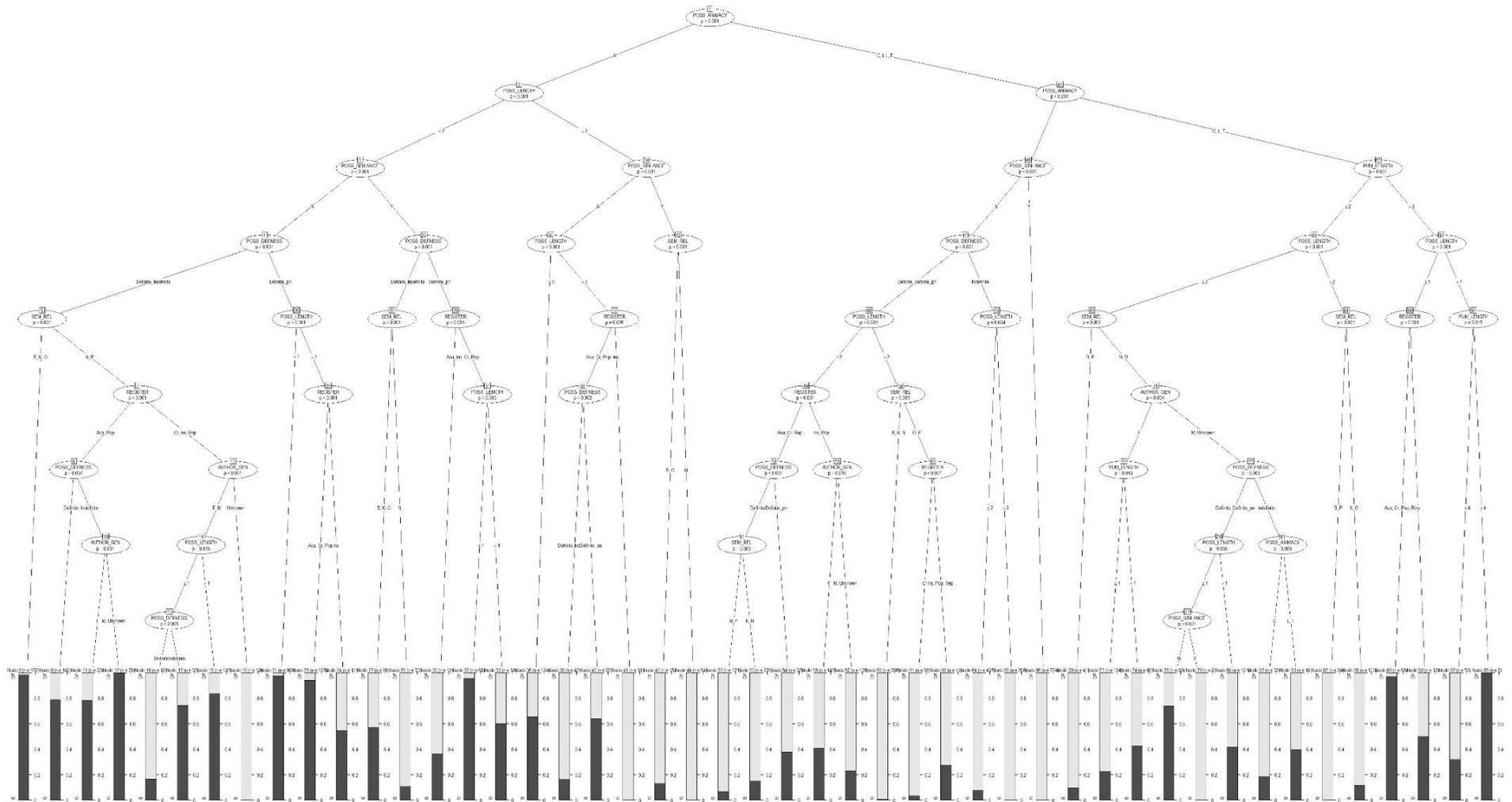


Figure 5.25: Conditional inference tree for AusE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER and EDITED STATUS in conditioning genitive alternation

5.3.3.2.2 Conditional inference tree analysis of the interaction of linguistic and language-external factors in WSAfE

Figure 5.26 presents the results of the conditional inference tree analysis for WSAfE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER and EDITED STATUS on the response variable (which has two levels, *s*-genitive and *of*-genitive). As with AusE, the absence of the variable EDITED STATUS implies that editors do not affect the eventual text in any significant way in WSAfE. This is generally in line with the findings in Section 5.3.2, except for the case of the creative register, where the proportions of the two constructions were reversed across unedited and edited writing (which implies that some other factor comes into play in edited creative writing). This shows that editors accept the genitive usage patterns of authors of WSAfE, and therefore reinforce these usage patterns, providing evidence of endonormativity in WSAfE, at least in terms of genitive alternation. An inspection of Figure 5.26 shows that linguistic factors play the most important role in conditioning the choice between the variants in WSAfE.

POSSESSOR ANIMACY is the most important factor in the choice of genitive construction, as evidenced by the first split at node 1: animate, collective and locative possessors are grouped together in the left branch, while inanimate and temporal possessors are grouped together in the right branch.

Animate, collective and locative possessors are then further split on the basis of POSSESSOR LENGTH at node 2, with short animate possessors strongly favouring the *s*-genitive and short collective and locative possessors favouring the *of*-genitive. The branch to the right of node 2 shows that animate, collective and locative nouns with possessors longer than two words favour the *of*-genitive, except when the possessum is longer (in which case they exclusively take *s*-genitives; node 36, $n = 7$). Two further exceptions are when animate and collective possessors are three words in length, but do not contain a final sibilant (node 30, $n = 26$; node 31, $n = 41$), in which case they favour the *s*-genitive. These findings show that animate nouns are mostly used with the *s*-genitive in WSAfE. Collective and locative nouns do sometimes occur with the *s*-form, but this only happens when possessors are short and do not contain final sibilants, or when their possessums are longer than their possessors.

Inanimate and temporal nouns are split on the basis of POSSESSOR DEFINITENESS at node 37. The branch to the right of node 37 shows that inanimate and temporal nouns that are definite proper nouns favour the *s*-genitive in 68% of 48 cases at node 61. Furthermore, definite inanimate and proper nouns also favour the *s*-genitive, but only when used in instructional and popular writing. In all other cases, inanimate and temporal possessors favour the *of*-genitive. This demonstrates that in WSAfE, definiteness cuts across animacy categories, with most definite and proper noun possessors favouring the *s*-genitive, irrespective of the animacy status of the possessor. This shows that in WSAfE definiteness of the possessor facilitates the extension of the *s*-genitive to possessors lower down on the animacy scale.

Given that EDITED STATUS plays no role in conditioning the choice between the two constructions, it may be concluded that the extension of the *s*-genitive in WSAfE to definite and proper noun possessors lower down on the animacy scale enjoys acceptance among editors of this sub-variety, and provides evidence for editors facilitating ongoing change in WSAfE and also evidence for endonormativity in WSAfE, at least in terms of genitive alternation.

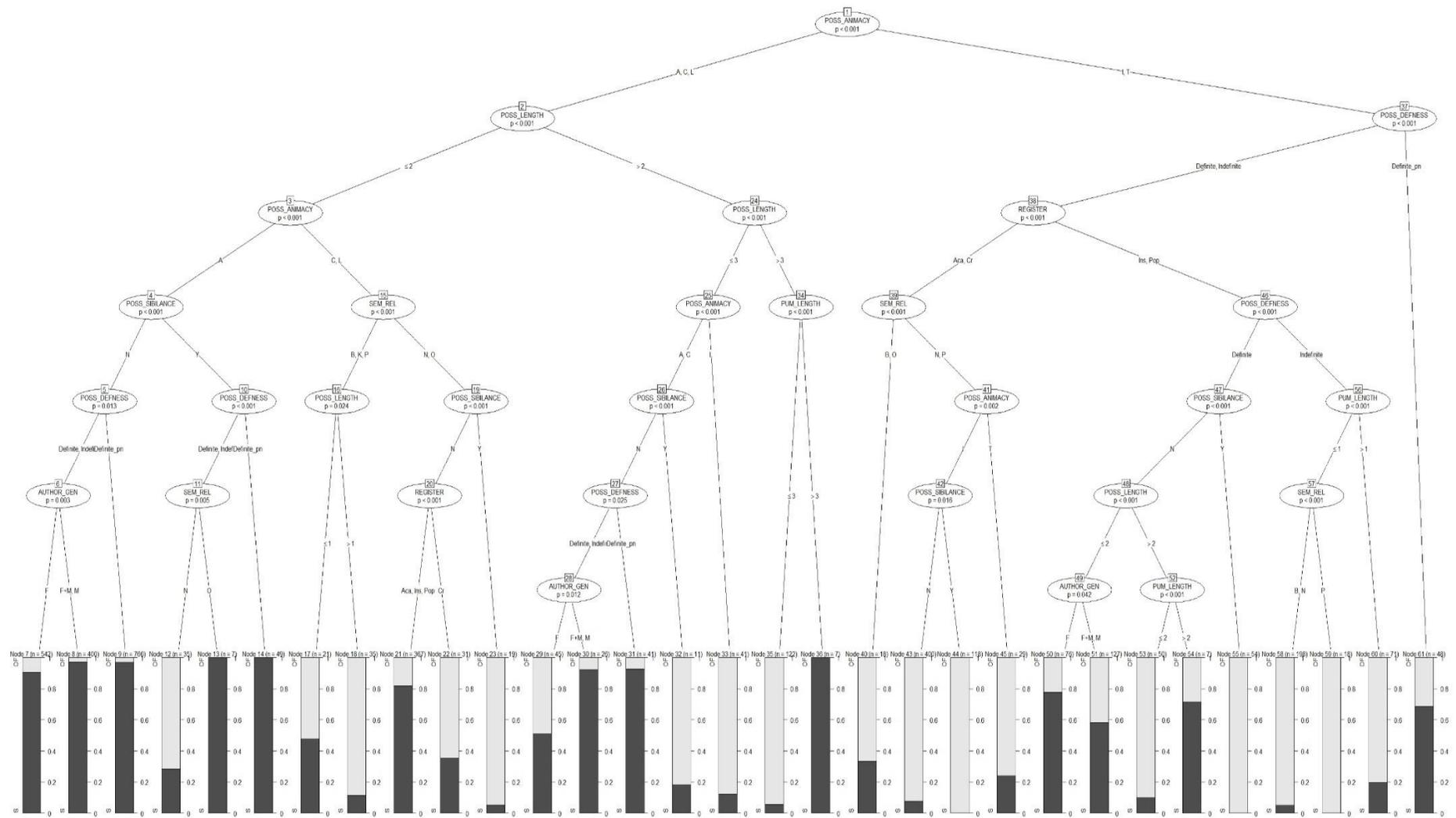


Figure 5.26: Conditional inference tree for WSafE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER and EDITED STATUS in conditioning genitive alternation

5.3.3.2.3 Conditional inference tree analysis of the interaction of linguistic and language-external factors in AfrE

Figure 5.27 presents the results of the conditional inference tree analysis for AfrE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITED STATUS on the response variable (which has two levels, *s*-genitive and *of*-genitive). EDITED STATUS, again, plays no role in conditioning the alternation between the two constructions, which suggests that the genitive usage patterns of AfrE authors are accepted by editors, providing evidence for endonormativity (for genitive usage patterns) in this sub-variety. The findings in Section 5.3.2 confirm this and show very little difference in the proportional frequencies of the two genitive constructions across unedited and edited AfrE.

An inspection of Figure 5.27 shows that POSSESSOR ANIMACY at node 1 is the most important factor conditioning genitive alternation in AfrE. Node 1 partitions animate, collective and locative possessors in the left branch and inanimate and temporal possessors in the right branch.

An inspection of the end nodes for the left branch shows that the *s*-genitive is the preferred construction with animate, collective and locative possessors in 16 of the 24 nodes. The possessors are split for FINAL SIBILANCE at node 2. The *of*-genitive is mostly preferred where these possessors contain a final sibilant, except in academic and instructional writing when these possessors are definite proper nouns (node 47, $n = 57$).

Animate, collective and locative possessors that do not end in a sibilant are further split at node 3 for POSSESSOR LENGTH. With regard to the right branch of node 4, short animate, collective and locative possessors overwhelmingly prefer the *s*-genitive. Long possessors prefer *of*-genitives in popular and reportage writing, while long possessors ending in a final sibilant prefer the *s*-genitive in academic and instructional writing. These differences may be somewhat unexpected but they may be linked to the kinds of readerships associated with these registers. Recall from Section 3.3 that popular and reportage texts are read by a more general and less specialised audience than the type of readership typical of academic and instructional writing. Given that long possessors and possessors ending in a final sibilant are more difficult to process in *s*-genitives, it seems likely that the preference of these types of possessors with the *of*-genitive in popular and reportage writing is associated with the broader and less specialised audience typical of these registers. This confirms the findings of Biber and Gray (2012) and demonstrates the way in which the broader processes of language change interact with the communicative functions of different registers to condition their varying openness to features associated with ongoing change.

Inanimate and temporal possessors are split at node 51 for DEFINITENESS. Indefinite inanimate and temporal possessors overwhelmingly favour the *of*-genitive, as do definite and definite proper noun

possessors ending in a final sibilant. However, definite and proper noun possessors that do not end in a final sibilant favour the *s*-genitive (more so in academic and instructional writing than in popular and reportage writing).

The conditional inference tree analysis of AfrE shows that EDITED STATUS plays no role in conditioning the choice between the two constructions, which suggests acceptance on the part of editors for the genitive usage patterns of AfrE authors and provides evidence for endonormativity for this feature in AfrE. Perhaps the most significant finding for AfrE is the fact that animate, collective and locative possessors favour the *s*-genitive, which provides evidence of the extension of the *s*-genitive to possessors lower down on the animacy scale. This extension is further evident in the fact that inanimate and temporal definite and proper noun possessors also favour the *s*-genitive (except when they end in a final sibilant). This demonstrates that while POSSESSOR ANIMACY is the most important predictor for which genitive will be selected, the strength of this factor is lower in AfrE than in the other varieties (as demonstrated by Rosenbach (2017)). This, again, provides evidence for a possible substrate transfer effect of Afrikaans on AfrE. The fact that EDITED STATUS plays no role in conditioning which genitive is selected in AfrE, suggests that the extended use of the *s*-genitive to possessors lower down on the animacy scale in AfrE enjoys acceptance by editors, and therefore provides evidence of endonormativity. Furthermore, for possessors that fall into the lowest two categories of the animacy scale (inanimate and temporal), FINAL SIBILANCE interacts with REGISTER to facilitate the further extension of these possessors to *s*-genitives. Specifically, it appears that the more compact *s*-genitive is preferred with inanimate and temporal possessors in registers targeted at more specialised readerships. Academic and instructional registers are therefore the locus of the extension of the *s*-genitive to the least animate possessors in AfrE – which is accepted by editors.

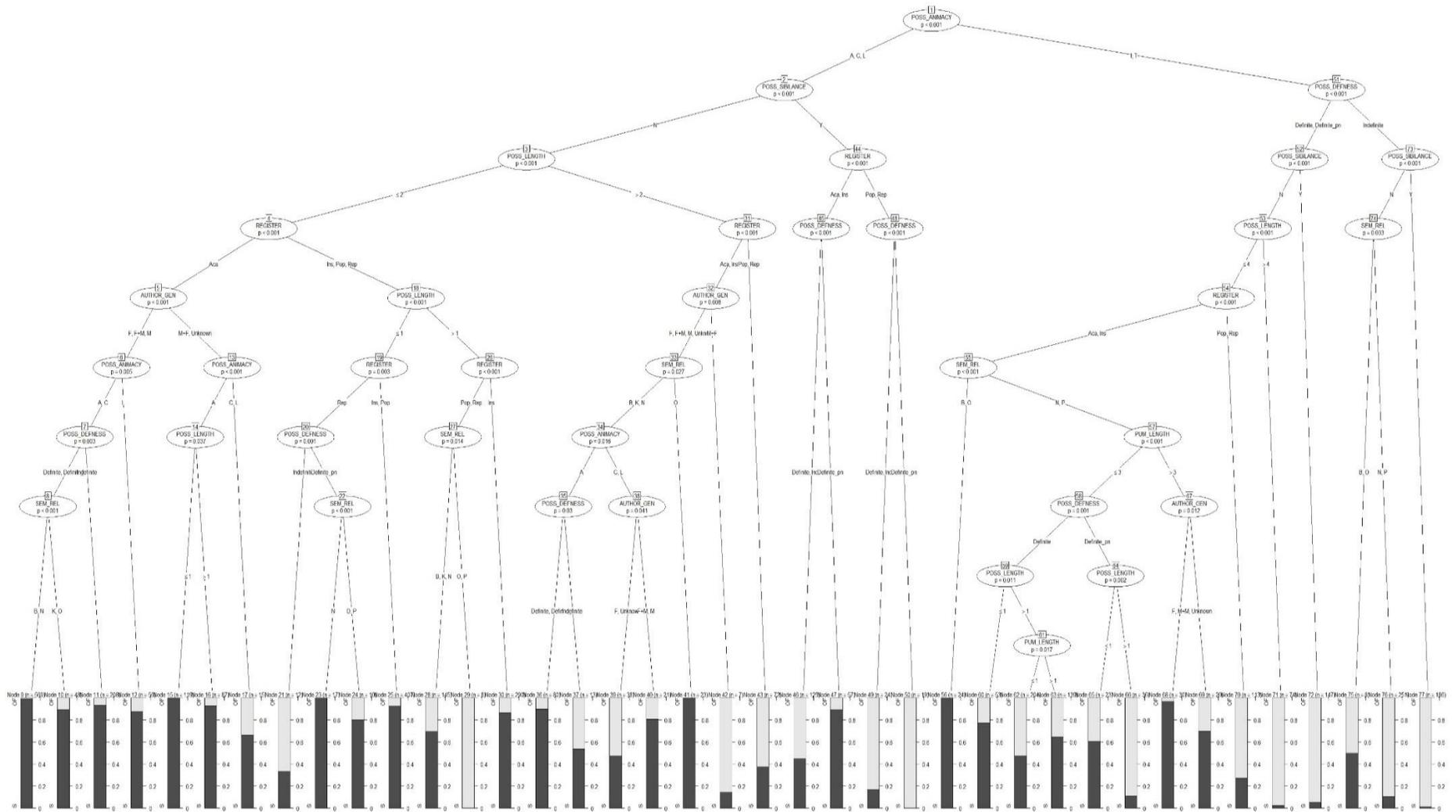


Figure 5.27: Conditional inference tree for AfriE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, POSSESSOR DEFINITENESS, SEMANTIC RELATION, REGISTER, AUTHOR GENDER, EDITOR GENDER and EDITED STATUS in conditioning genitive alternation

5.3.3.2.4 Conditional inference tree analysis of the interaction of linguistic and language-external factors in BSafE

Figure 5.28 presents the results of the conditional inference tree analysis for BSafE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, SEMANTIC RELATION, EDITOR LANGUAGE and EDITED STATUS on the response variable (which has two levels, *s*-genitive and *of*-genitive). EDITED STATUS plays a minor role in conditioning the choice of genitive construction in BSafE. This is noteworthy, particularly given that EDITED STATUS does not play a role in conditioning genitive alternation in AusE, WSafE and AfrE. While the findings in Section 5.3.2 show very little difference in the proportional frequencies of the two genitive constructions across unedited and edited BSafE, the analysis of the types of editorial changes in Section 5.3.2.3 shows that editors do make more types of changes to the genitive constructions. Furthermore, the conditional inference tree for BSafE shows that this is the only sub-variety in which EDITOR LANGUAGE plays a (minor) role. This is also a noteworthy finding, particularly in light of the fact that this is the only sub-variety where the language profile of the editors differs substantially from that of the authors. As indicated in Section 4.3.2, BSafE editors are not represented in the corpus at all. This is not unusual, since previous studies have shown that editors of English texts in South Africa are either WSafE or AfrE users (two varieties that have enjoyed much closer contact in the course of their historical development and which have progressed further along the stages of the Dynamic Model than BSafE (refer to the discussion in Section 3.2.2)). Therefore, the choices of BSafE authors are subjected to a more stringent acceptability test, than are the choices of WSafE and AfrE authors where the language profile of editors is much more similar to authors.

POSSESSOR ANIMACY is the most important predictor in genitive alternation in BSafE. Node 1 splits the animacy categories so that animate and temporal possessors are grouped together in the left branch and collective, inanimate and locative possessors in the right branch.

Of the five end nodes for animate and temporal possessors, two favour the *of*-genitive. This is firstly for possessors that are longer than two words (node 10, $n = 11$) and secondly for shorter possessors that end in a FINAL SIBILANT (node 9, $n = 40$). Animate and temporal possessors that are short and do not end in a final sibilant overwhelmingly favour the *s*-genitives (node 5, $n = 85$; node 7, $n = 13$; node 8, $n = 140$), with very minor differences across EDITOR LANGUAGE (node 4).

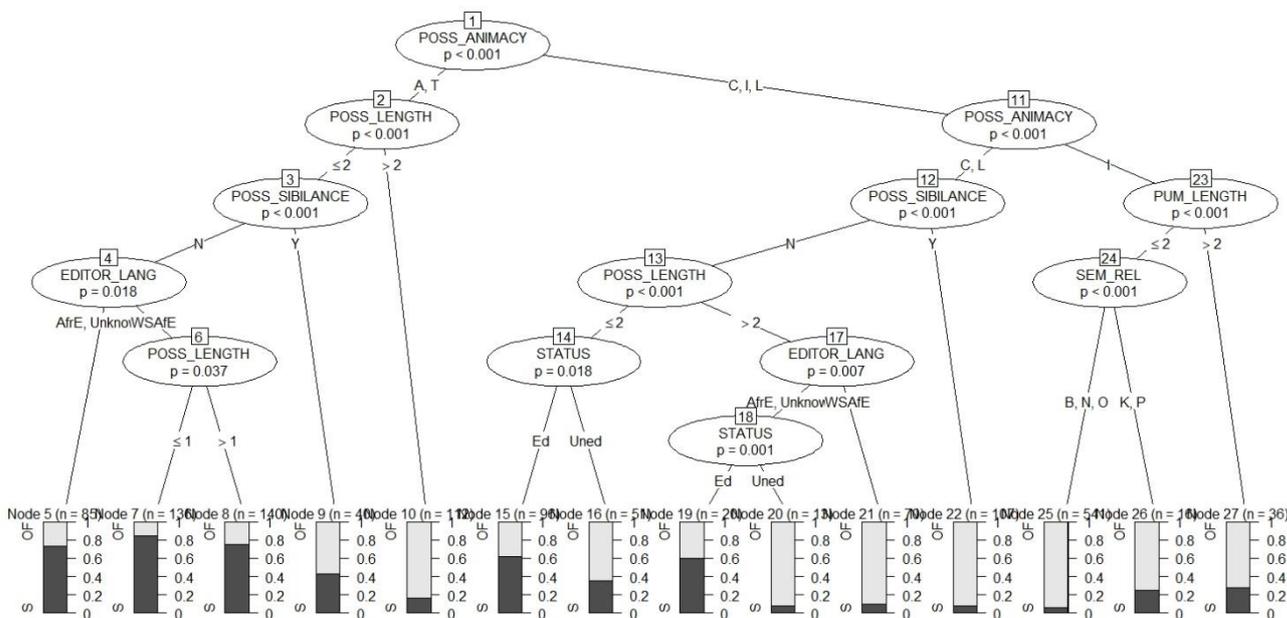


Figure 5.28: Results of the conditional inference tree analysis for BSAfE for the interaction of the predictor variables POSSESSOR ANIMACY, POSSESSOR LENGTH, POSSESSUM LENGTH, FINAL SIBILANCE, SEMANTIC RELATION, EDITOR LANGUAGE and EDITED STATUS with the response variable (with two levels *s*-genitive and *of*-genitive)

The right branch of node 1 is of particular interest for this study, especially given the very strong preference for the *of*-genitive found for BSAfE in Section 5.3.2. Of the nine end nodes for this branch, seven show a very strong preference for the *of*-genitive, and is most likely a contributing factor to the very high proportion of *of*-genitives in this sub-variety. The *s*-genitive is the more favoured construction for collective and locative nouns in only two contexts, and in both cases this appears to be due to editorial intervention. When collective and locative possessors do not contain a FINAL SIBILANT and have short POSSESSOR LENGTHS, then they favour the *s*-genitive more in edited texts than in unedited texts. Furthermore, EDITOR LANGUAGE interacts with EDITED STATUS for collective and locative possessors that have long possessors: AfrE editors are much more likely to use these possessors in *s*-genitives, than BSAfE authors and WSAfE editors are. This is in all likelihood due to the lower strength of POSSESSOR LENGTH in AfrE and the relatively stronger importance of POSSESSOR LENGTH in WSAfE highlighted in Section 5.3.3.1. Given this, it seems that in BSAfE, the strength of POSSESSOR LENGTH is variable across the sub-varieties, with AfrE and WSAfE editors imposing their own preferences when editing BSAfE, which, in the case of AfrE works to boost the *s*-genitive with collective and locative nouns, and in the case of WSAfE works to constrain the extension of the *s*-genitive to these possessors. This shows that WSAfE and AfrE editors are not yet happy to accept the choices of BSAfE authors especially with regard to their choice of genitive construction with collective and locative possessors. But the effect of these two editor language profiles is to move the preference for genitive constructions in opposite directions, so that the constraining effect of WSAfE is counteracted by the more innovating effect of AfrE, resulting in very little difference in the proportional frequencies of the two constructions across unedited and edited BSAfE.

The analysis of the conditional inference trees show that the effects of editorial intervention raised in the frequency-based analyses are a collateral consequence of other factors changed in the editing process. For example, for AusE, WSAfE and AfrE EDITED STATUS plays no role in conditioning genitive alternation across unedited and edited writing (as was suggested in Section 5.3.2) but it does play a minor role for BSAfE. In the frequency based analysis, it seemed that editors largely accept the innovative usages of BSAfE authors, but the conditional inference tree analysis showed that such a broad approach masks the actual patterns: for example, WSAfE editors maintain the lower *of*-genitive in BSAfE writing, but AfrE editors do not, and increase the proportion of *s*-genitives. Given that the animacy constraint is lower in AfrE (due to substrate influence), this means that there is a complex interplay between covert and overt norms when editors edit the texts of authors who come from different linguistic backgrounds. Furthermore, the conditional inference tree analysis showed that while the proportion of *s*-genitives is higher in AusE, this is not due to an extension of this genitive to nouns lower on the animacy scale, but is due to the higher use of animate possessors in the variety coupled with a much stronger animacy constraint.

5.4 Summary of the findings

This chapter has presented, discussed and interpreted the results from the questionnaire and corpus analysis that are the foundation of the empirical component of this thesis, and which seek to answer research questions 2.1, 2.2 and 2.3.

The analysis of the questionnaire data revealed some similarities between editors of AusE and SAfE along the lines of age, gender, education level and work experience, but some striking differences in their view of their editorial role, their language profiles and their use of norm-providing sources.

Editors of AusE and SAfE are usually older women who are highly educated and therefore belong to a group of language users that generally regarded as conservative language users with a preference for standard language usage. In addition to near similar profiles in terms of age, gender and education, the respondents had similar working backgrounds: most work on a freelance basis and consequently have experience in a large range of text types. The effect of this is a more diffuse normative environment with the accompanying diversity of stylistic and textual conventions associated with the many types of texts that freelance editors work on. This less focused normative environment opens up the opportunity for the diversity of usages preferences to creep in to editors' choices, and may result in varying choices in the texts that they edit.

The respondents in the two countries view their role as having both a normative and communicative focus, but they differ somewhat in the weight afforded to each. Specifically, the AusE editors indicated that they prioritise the normative dimension in their work, while this dimension carries less weight for the South African editors who expressed a much more balanced view in terms of the importance of each dimension. The respondents' motivations for these views are telling, and speak to the different linguistic landscapes in

which their editorial work is carried out. For the South African editors, the fact that many readers of English texts in South Africa are second-language users of English seems to result in a view in which both dimensions of the editorial role are considered equally important, yet for the Australian editors the fact that most readers of the texts they edit are native users of English seems to bolster the normative dimension. This is probably also due to differences in the availability of codification instruments for English in the two countries: in the South African context, South African editors might (in the absence of codified norms) turn to the communicative dimension to ensure that texts are optimised for their audiences, while in the Australian context there is a clear set of norms to aim for and against which a text can be assessed, providing Australian editors with the opportunity to take a more focused normative approach to their editorial work.

Perhaps the most striking finding with regard to the demographic profiles of editors in the two countries is differences in the language backgrounds of the editors from the two countries. The Australian editors are all first- or home-language users of English who nearly all edit in English only, which further reinforces the more focused normative environment in which they work. This homogeneity in language profiles together with a much more clearly codified variety of English, points to convergence between overt and covert norms for AusE editors, which results in linguistic editorial choices that converge on the Australian norm. The South African respondents' language profiles are much more complex. There is a high degree of bidirectional bilingualism among South African editors (roughly half of the editors are English–Afrikaans bilinguals, while the other half are Afrikaans–English bilinguals). This, together with the much more diffuse normative environment most likely opens up the opportunity for the choices of SAfE editors to be more divergent than those of AusE, under the pressures of less a focused normative environment, the potential influence of the substrate Afrikaans for AfrE editors (or even bilingual WSAfE editors) and the possibility of contact effects among the many different sub-varieties used in the country. Furthermore, it is striking that only two strands of English are represented in the language profiles of the South African editors, and that none of the editors represent the dominant IDG strand, BSAfE. This suggests that BSAfE users are not represented in this type of linguistic gatekeeper role, and the texts produced by BSAfE users are edited by users of other strands, which places particularly stringent acceptability requirements on the choices of BSAfE authors (whose work must be judged as acceptable by users of other sub-varieties).

This contrast between a more homogenous and a more diffuse normative environment is also evident in the kinds of norm-providing sources that the respondents from the two samples draw on as well as the orientation of these sources. For the Australian respondents, there is a strong preference for endonormative sources, with a minor preference for AmE sources (especially for dictionaries and style guides). In the absence of codification instruments for SAfE, the South African editors have a clear preference for exonormative sources, with the appeal of BrE sources being particularly strong. In addition to a preference for BrE sources, the South African editors also have a preference for bilingual English–Afrikaans and monolingual Afrikaans sources, which is probably due to the high level of English–Afrikaans bilingualism among the South African editor respondents, as reflected in the data on their language profiles, but it may

also suggest that in the absence of explicit SAfE norm-providing sources, these bilingual editors utilise their high levels of bilingual proficiency and draw on sources produced within the South African context, which opens up the possibility that their editorial decisions are influenced by Afrikaans norms.

The number of Australian respondents who indicated that they use usage guides, is higher than the number of SAfE editors who make use of usage guides. Australian editors seem to be aware of and draw on the usage guides developed for their own variety of English; however, these are used by less than half of the respondents, and compete with BrE usage guides. The South African editors far prefer making use of language reference works, compared to usage guides, which is in all likelihood due to the fact that comprehensive usage guides are not yet available for this variety. Given the high number of English sub-varieties in the country, usage in South Africa is quite diverse, which might lead to editors having to consult online sources in order to guide their decisions when ‘questionable’ usage arises in the texts that they edit. Since it is in usage guides where normative recommendations for the use of the *s-* and *of-*genitive are set out, the low use of usage guides among SAfE editors suggests that decisions regarding which genitive are used are based on editors own judgements, and are most likely informed by their exposure to the communal usage patterns in the community.

One last noteworthy finding is the popularity of free online sources among both groups of editors. For example, in both the Australian and the South African data, dictionaries that can be accessed for free were some of the most popular sources among the respondents to the questionnaire. This raises the possibility that the selection and use of norm-providing sources might be less guided by explicit decisions about appropriate norms, than they are by questions of cost. It is noteworthy that systematic patterns for preferred norm-orientation may still be found across the two samples: Australian editors make use of free AmE dictionaries, while South African editors make use of BrE dictionaries.

The data from the questionnaire provided some important information on the kinds of norm-providing sources that editors draw which gives an indication of the kinds of overt norms that they invoke in their editing work. The information on the editors’ language profiles provides some clues as to the possible influence of covert norms on their choices, but this kind of reported data does not provide a reliable indication of the extent to which either of these norms find their way into the actual editorial choices that editors make. To assess this, a corpus-based analysis of the effects of editorial intervention across the two varieties and different written registers in conditioning genitive alternation was conducted.

Although the data from the questionnaire cannot be correlated with the data from the corpus (since it is not known whether the editors who completed the questionnaire are the same ones represented in the corpus), it is possible to draw parallels since the language profiles of the editors represented in the corpus are broadly in line with those who completed the questionnaire. All editors of AusE are first-language users of English, with nearly all being AusE users. Most editors in the SAfE corpus are WSAfE, with some representation of AfrE.

The first stage of the corpus-based investigation, which focused on the distributional analysis of the proportional frequencies of the genitive constructions across sub-varieties and the edited registers, showed that there are some differences across sub-varieties and registers. The statistical evaluation of the effect of variety and register showed that register is a stronger predictor than sub-variety for which genitive construction is used and confirmed the findings of the general analysis of the proportions and normalised frequencies of the two genitive constructions across registers and sub-varieties (Section 5.3.1.1 and 5.3.1.2). The academic and instructional registers both favour the use of the *of*-genitive, compared to the creative, popular and instructional registers, where preference for this *s*-genitive is stronger. The statistical evaluation showed that sub-variety does not play a role in conditioning the genitive alternation in the edited creative, popular and reportage registers, suggesting that in these registers, there is very little difference in the proportions of the two constructions across the sub-varieties. Cross-varietal differences only emerge for the academic and instructional registers, where the sub-varieties cluster together along country lines. AusE has the strongest preference for the *s*-genitive in these two registers, compared to the SAfE varieties. This demonstrates that academic and instructional AusE is particularly more open to the more informal and compact feature than the South African sub-varieties. Within the South African sub-varieties, AfrE and WSAfE cluster together in their preference for the *s*-genitive in academic and instructional writing, providing evidence that the two sub-varieties function like a dialect cluster, with BSAfE distinguishing itself as significantly less open to the use of the *s*-genitive in these two registers, by comparison. This confirms the findings by various scholars that the preference for the *s*-genitive appears to pattern along the lines of how far a variety has progressed through the stages of the Dynamic Model (Heller et al., 2017b; Szmrecsanyi et al., 2016b)

Having established that there are differences in the proportional frequencies of *s*- and *of*-genitives across registers and sub-varieties in edited texts, the analysis shifted focus to the influence of editorial intervention on the proportions and frequencies of the two constructions across the registers and sub-varieties. This analysis showed that editors do play an important role in shaping the proportional frequencies of the two constructions across the sub-varieties and registers, but that the effect of editorial intervention differs across the varieties, and within each variety, across the different registers. In AusE, editorial intervention does not play all that big a role in shifting the proportions of the two constructions in either direction: editors mostly accept the proportions of the two constructions across the registers, which reflects the endonormative stabilisation of this variety. The higher overall proportion of *s*-genitives in AusE, and the acceptance of this by editors of AusE, shows that the characteristic colloquial nature of AusE has gained acceptance among linguistic gatekeepers, even in the most formal of registers, academic writing, where the proportion of *s*-genitives is remarkably high.

Editors of the SAfE sub-varieties play a much more interventionist role in shifting the frequencies of the two constructions, but this differs across the sub-varieties. Specifically, editors of WSAfE constrain the increase of the *s*-genitive in unedited instructional writing bringing this register in line with conventional

expectations of instructional writing. Editors of WSAfE are therefore partly responsible for maintaining the distinction between the more formal and the less formal registers in terms of their preferences for the two constructions. Unedited BSAfE overall disfavours the *s*-genitive (which is in all likelihood due to the slower progress of this sub-variety along the stages of the Dynamic Model, but also to a possible influence of substrate language, since a comparable *s*-genitive construction is unavailable in the Bantu languages). Edited status does not play any significant effect in shifting the proportion of *s*-genitives in the registers of this sub-variety: in other words, editors of BSAfE accept the very low usage of *s*-genitives across all registers in BSAfE. In AfrE, editorial intervention only plays a significant role in the academic and instructional registers, where editors increase the proportional frequencies of the *s*-genitive in these two registers, compared to the popular and reportage registers, where the decrease in the proportion of *s*-genitives in edited writing is not significant. This demonstrates that editors of AfrE boost the proportion of *s*-genitives in these two most formal registers, which bring the proportions of the two constructions much more in line with WSAfE and provide evidence for convergence between the two sub-varieties which is aided by editorial intervention.

In the final analysis, the focus shifted to the interaction of linguistic factors with language-external factors in conditioning genitive alternation in each of the four sub-varieties. This analysis showed the group of linguistic factors are the most important conditioning variables in all four varieties, but the effect of these factors differ across the sub-varieties, confirming the finding by Grafmiller, Szmrecsanyi, Röthlisberger, and Heller (2018) that shifting usage frequencies in language-internal variation lead to differences across (sub-)varieties. Language-external factors were shown to play only a minimal role in the different sub-varieties, but there are differences in terms of the variable strength of these factors across the different varieties.

The most relevant finding for this study is that, once linguistic predictors are taken into account in the model, editorial intervention does not play a role in conditioning which genitive construction is selected in AusE, WSAfE and AfrE, but it does play a role in BSAfE (where the first language of the editor plays a role too). This suggests that the genitive usage practices of AusE, WSAfE and AfrE authors enjoy acceptance by editors of these varieties, who reinforce these usage patterns by accepting them and allowing them to disseminate further.

In the case of AusE, the usage patterns show a strong favouring of the *s*-genitive with animate possessors, while inanimate possessors are favoured by *of*-genitives. This suggests that the extension of the *s*-genitive to nouns lower on the animacy scale is largely restricted in AusE, but the higher proportion of *s*-genitives noted for the sub-variety, is due to a higher number of animate nouns coupled with a strong animacy constraint. This suggests that AusE might not be as open to ongoing extension of the *s*-genitive as would be assumed by the higher proportion of *s*-genitives; rather, it is simply that there are an extremely high number of animate possessors used in the texts in this variety, and this, coupled with the strong animacy constraint accounts for the higher proportion of *s*-genitives.

In WSAfE, there is evidence to show that the *s*-genitive is being extended to possessors lower on the animacy scale, and especially collective and locative possessors. The extension of the *s*-genitive to inanimate and temporal possessors is more restricted, but definite and proper noun possessors are favoured by the *s*-genitive. This provides evidence that WSAfE does participate in the ongoing change in this grammatical feature, and that this is facilitated by a gradient loosening of the animacy constraint. Furthermore, editors of WSAfE accept these innovative usages, legitimising them and contributing to their further dissemination. Editorial intervention in WSAfE therefore facilitates the ongoing change in WSAfE, which provides evidence for endonormativity in WSAfE, at least in terms of genitive alternation.

A similar result is found for AfrE: AfrE authors favour the *s*-genitive with animate, collective and locative possessors, which provides evidence of the extension of the *s*-genitive to possessors lower down on the animacy scale for this sub-variety. This extension is further evident in the fact that definite and definite proper noun inanimate and temporal possessors also favour the *s*-genitive. Furthermore, editors of AfrE accept these usages, providing evidence of endonormativity. Therefore, the two varieties' participation in the ongoing change in genitive alternation is the same and is facilitated by the same linguistic factors, and is accepted by editors of both sub-varieties. One last noteworthy finding for AfrE is that for possessors that fall into the lowest two categories of the animacy scale (inanimate and temporal), FINAL SIBILANCE interacts with REGISTER to facilitate the further extension of these possessors to *s*-genitives. Specifically, the more compact *s*-genitive is preferred with inanimate and temporal possessors in registers targeted at more specialised readerships – the academic and instructional registers. Therefore, the locus of the extension of the *s*-genitive to the least animate possessors in AfrE is in the more formal of the two registers, but also the two registers where readers are able to easily process the kind of increased lexical density associated with the more compact construction.

A different picture emerges for BSAfE, which is the only sub-variety in which EDITED STATUS and EDITOR LANGUAGE play a role. This is also a noteworthy finding, particularly in light of the fact that this is the only sub-variety where the language profile of the editors differs substantially from that of the authors. This means that the choices of BSAfE authors are subjected to a more stringent acceptability test, than are the choices for WSAfE and AfrE where the language profile of editors is much more similar to authors. Specifically, it was found that the effect of these two editor language profiles is to move the preference for genitive constructions in opposite directions, so that the constraining effect of WSAfE is counteracted by the more innovating effect of AfrE resulting in very little difference in the proportional frequencies of the two constructions across unedited and edited BSAfE.

5.5 Conclusion

This chapter has presented and analysed the findings for the empirical investigation of this thesis. Section 5.2 presented the findings for the questionnaire that sought to gather information on the demographic profiles of editors of English texts in Australia and South Africa. More importantly, Section 5.2 analysed

the information gathered on the linguistic profiles of editors of these two varieties and the norm-providing sources used by these two sets of editors. Section 5.3 presented the findings of the corpus-based case study of how editorial intervention influences the alternation of the genitive constructions in AusE and SAfE, in order to determine if and what role editorial intervention plays in the ongoing change of this grammatical feature in the two varieties, and how this role differs in the two varieties. This analysis made use of quantitative corpus linguistic techniques and drew on distributional analyses and statistical evaluations in order to investigate this. The next chapter, Chapter 6, summarises and synthesises the findings of the theoretical and empirical investigations in the study with reference to the research questions. Thereafter, suggestions for further research are made, before the study is concluded.

Chapter 6

Synthesis of findings, recommendations for future research and conclusion

6.1 Introduction

The role of editorial intervention in the processes of ongoing language change has not received much attention in investigations of the different varieties of World Englishes and across different written published registers. As a consequence, very little is known about how editors contribute to the processes through which linguistic features and usage patterns associated with change are disseminated and come to be conventionalised within the different varieties and across different written registers, nor how this intervention might help to move varieties along the stages of the Dynamic Model.

This study sought to address this gap. In a first step (see Chapter 2), it considered how the role of editorial intervention may be conceptualised within a usage-based framework of language that integrates cognitive and social perspectives for understanding language structure, variation and change within the context of World Englishes and register studies. Subsequent to this, it used a double-pronged empirical method combining survey research and corpus analysis to (a) investigate the sociolinguistic background and explicit normative orientation of editors in South Africa and Australia (thus exploring the relationship between overt and covert norms in the two varieties) (see Section 5.2), and (b) analyse the editorial role in the changing frequencies of two genitive constructions across AusE and SAfE and five different written registers (see Section 5.3).

This chapter concludes the study by summarising the findings of the investigation with reference to the research questions posed in Chapter 1. The summary of findings for the first two research questions are presented in Section 6.2. Thereafter, in Section 6.3, the synthesis of the theoretical and empirical findings for these two questions is considered in terms of processes of language change, and ongoing change in World Englishes and written published registers, thus answering research question 3. Section 6.4 reflects on the limitations of the study and makes recommendations for further research based on these limitations. Section 6.5 provides some concluding remarks.

6.2 Summary of findings

The introductory chapter of this study posed three primary questions and three sub-questions. Research question 1 was answered by means of the theoretical discussion in Chapter 2, while research question 2 and its sub-questions were answered by the empirical work presented in Chapter 5, against the background of the literature survey in Chapter 3. Research question 3 was provisionally answered in Section 5.4, through the triangulation of the empirical data and the interpretation of the findings from this data within the

theoretical framework proposed in Chapter 2, but will be discussed further in this section. This section summarises the findings with reference to each research question.

6.2.1 Findings for research question 1

How do broad trends of language change, the different contextual forces of the varieties of English, register effects and editorial practice interact to influence language variation and change?

Chapter 2 answered research question 1 by developing a theoretical framework in which the role of editorial intervention in ongoing language change in different varieties of World Englishes and across different registers is conceptualised within a usage-based approach that integrates cognitive and social perspectives for understanding how language is used and how this use gives rise to variation and drives change. The positioning of the editorial role within these processes can be understood in light of several crucial considerations.

First, within the context of the interwoven processes of language variation and change, selection is a process that is driven by the interaction of individual psycholinguistic factors and communal social factors (see Sections 2.2 and 2.3). Language users build up their knowledge of language (that is, their individual cognitive representations of the conventions of the language community to which they belong) by exposure to language use in those communities. This cognitive representation of the communal patterns of use is sensitive to the effects of frequency: the more a language user encounters a particular variant in language-usage events, the more entrenched this variant becomes and the more likely it is that this variant will be selected in subsequent language-usage events (see Section 2.2.1). This cognitively-motivated individual selection and subsequent use forms part of the input that other language users receive, increasing the frequency of their exposure to the variant, and leading to the entrenchment of the feature for them too. Thus, the diffusion of a variant within a language community is driven by the collective (cognitively-conditioned) choices of interacting individual language users in contact with one another.

However, language users' encounters with variant constructions take place within the context of humans communicating with one another in a range of social contexts (in both speech and writing), and information about the various aspects of these contexts of use are encoded in their individual cognitive representations of language (as discussed in Section 2.2.1). Thus, language users also encode information about the conventionalised meanings of constructions, possible variants and the social and discursive contexts in which they are used. The social dimension of language use means that linguistic variants are invariably assigned social values, which typically centre on issues of salience and prestige, and specifically, norms. Thus, a linguistic variant may be socially evaluated as reflecting an overt prestige norm (in which case its usage is sanctioned by norm-providing authorities) or a covert solidarity norm (in which case it reflects a communal pattern of use that may deviate from the norm sanctioned in norm-providing sources). The variants that language users select may be motivated either by the conscious selection of a variant on the

basis of trying to achieve some social goal, or may be a result of the subconscious process of entrenchment due to exposure to use (see Section 2.3).

Within the context of World Englishes, a language user may be exposed to a number of language communities, depending on the nature of the country's sociolinguistic landscape, the number of different sub-varieties represented in the country and the degree of contact between the sub-varieties – all of which shape the emergence of covert norms (see Section 2.4). In countries where there are multiple interacting sub-varieties, the communal usage patterns will be more diffuse, compared to where fewer interacting sub-varieties are represented and where the interaction among them is limited. Furthermore, depending on a variety's progression along the stages of the Dynamic Model (see Section 2.4), the orientation of the overt norms of the language community may either be exonormative or endonormative. Thus, for countries that have progressed further along the stages of the Dynamic Model, with one variety having attained endonormative stabilisation and where contact between different sub-varieties is limited, the covert norm is likely to be more homogenous and more closely aligned to the overt norm. For countries that have not attained endonormative stabilisation and where multiple sub-varieties are in contact with one another, the covert norm is likely to be more heterogenous and competition may therefore arise between covert norms and the overt extraterritorial English norms that are codified in different norm-providing sources (see Section 2.6.1). Furthermore, due to the particular characteristics of the local language environment, different varieties participate in global trends of change differentially, giving rise to different patterns of intra- and inter-varietal variation.

Editorial intervention in written published texts reflects the language use of editors, whose knowledge of language is built up in the same way as it is for other language users (through exposure to use) and is subject to the same cognitive and social processes, mechanisms and constraints that affect other language users (see Section 2.6.2). Editors are also unique in comparison to other language users, because their frequent exposure to prescriptive overt norms, as a consequence of their work, also feeds into their knowledge of language. Furthermore, editors are exposed to large amounts of unedited written language produced by various language users, which, depending on the nature of this written language as well as whether authors and editors share the same linguistic background or not, increases their exposure to varied communal patterns of use. This has the potential to entrench particular variant or innovative usages in editors' cognitive representations of language, relative to the number of sub-varieties represented in this writing. This continual exposure increases the likelihood that these usages will be accepted by editors when they encounter them in the texts that they edit (due to the cognitive effects of frequency). Therefore, while overt norms are more saliently represented in editors' mental representations of language than they are for other language users, editors are also frequently exposed to changing usage patterns through the process of reading and evaluating unedited writing. This potentially gives rise to competition between overt norms and covert norms, where a disconnect between the two sets of norms exist.

Furthermore, editors do not consult a norm-providing source for every amendment that they make to the texts that they edit, but instead mostly rely on their expertise in language, which is their entrenched knowledge of language built up through their exposure to overt norms sanctioned in norm-providing sources and covert norms that assert themselves through communal usage patterns. This increases the likelihood that covert norms will be accepted by editors but also demonstrates that editors' minds are an arena for the complex interplay between overt and covert norms.

Thus, a measure of the acceptability of features associated with ongoing change is the (conscious or sub-conscious) acceptance or removal of the feature by editors in different varieties of English. Such acceptance allows the feature into the feature pool, legitimises its use and contributes to its further dissemination in the language community, aiding its conventionalisation and helping the variety move towards endonormativity.

6.2.2 Findings for research question 2

How have AusE and SAfE progressed along the stages of the Dynamic Model, and what are the implications of this for editorial work in AusE and SAfE?

Research question 2 is a complex question, which necessitated the identification of three sub-questions, namely:

- 2.1) How do the language profiles of AusE and SAfE editors differ and how are these profiles similar or different to the language profiles of the authors that they edit, reflecting similarities and differences in covert norm orientation?
- 2.2) What overt norm-providing instruments do editors of AusE and SAfE use to guide their editorial choices, and how and why does this differ in each of these two varieties?
- 2.3) How do editors of AusE and SAfE respond to a feature known to be undergoing language change, namely the genitive alternation, and how are these responses shaped by the interplay between overt and covert norm orientation?

To lay the foundation for answering the sub-questions of research question 2, Chapter 3 presented an overview of the varieties of AusE and SAfE, the five registers investigated in the study, and the linguistic feature investigated, namely the genitive alternation. Section 3.2 explored the formation of the two varieties, their progression along the stages of the Dynamic Model and the current position and status of English in the two countries. The language profiles of editors and authors in each of the two varieties also received attention, and some existing evidence for the potential influence of editorial intervention in these two varieties was discussed.

Section 3.3 discussed the registers investigated, focusing on a description of each in terms of the various situational characteristics that give rise to the differences among them. The situational characteristics discussed include the function of the text, the nature of the topic, the type of readership for which the text is created, and the production circumstances. In addition, existing scholarship on the differential participation of the five registers in two broader trends of ongoing change, namely colloquialisation and densification, was surveyed in order to explore how and why the registers participate in two trends of change in different ways, relative to their underlying functional motivations and the types of readers for which they are created. Section 3.3 also sketched the variable role and nature of editorial intervention in the different registers with reference to the unique and varying publishing–production circumstances of each.

Thereafter in Section 3.4, the genitive alternation, the grammatical feature selected for analysis, was discussed. This discussion focused on the two major adnominal constructions that enter into this alternation, sketching out their historical development and presenting evidence of this alternation by drawing on existing scholarship from within the variationist paradigm to demonstrate the ongoing change in this construction. The discussion of this alternation explored the various language-internal factors known to condition the selection from the two choices available to language users, and reflected on how these linguistic factors are known to interact with variety and register (two language-external factors) to give rise to variation in this construction across different varieties of World English and across different registers. The discussion also showed how the use of the constructions might be conditioned by the broader trends of change and the different varieties' and registers' openness to these trends. The discussion in Chapter 3 therefore laid the foundation for the interpretation of the empirical data collected for this study.

The collection of the data used to answer research question 2 followed a two-pronged approach. The first approach sought to answer research sub-questions 2.1 and 2.2, and did so by surveying editors of English in Australia and South Africa on their language profiles and the kinds of norm-providing sources they draw on to guide their editorial choices. The analysis and interpretation of the questionnaire data was framed against the development and current status of English in each of the two countries in terms of the Dynamic Model (as discussed in Section 3.2), and tensions between overt and covert norm orientation, thus answering questions 2.1 and 2.2. The second approach sought to answer research question 2.3, and made use of a corpus-based variationist method in order to investigate how editorial intervention interacts with broader processes of language, specific language change in AusE and SAfE, and register effects.

6.2.2.1 Findings for research questions 2.1 and 2.2

As discussed in Section 3.2.1, AusE is an Inner Circle variety in which the STL strand has prevailed, and the variety has achieved endonormativity, as reflected in the codification of the variety in several national dictionaries and other sources like style manuals and usage guides. AusE appears to be making inroads into Phase 5 of the Dynamic Model, mostly reflected in sociolectal differences (conditioned by age, education

and gender) in phonology and lexicogrammar, with scholars noting that further internal differentiation in AusE registers and modes of discourse may emerge as part of the (normal) processes of ongoing change and the linguistic changes in the relationships between registers in modern English.

The complex sociolinguistic constellation in South Africa has given rise to a diversified internal variation structure in which multiple sub-varieties of English, such as WSAfE, AfrE and BSAfE, are currently developing in contact with each other at different rates in the same context (see Section 3.2.2). WSAfE is a visible and influential sub-variety in the country, but it is only used by a minority of South Africans, with the overwhelming majority of English users in South Africa belonging to the second-language indigenised sub-variety, BSAfE. WSAfE and AfrE have been in continual contact for an extended period of time, and some evidence of linguistic convergence between the two sub-varieties can be found at the level of lexis and grammar. As a result of apartheid policies, there has been less contact between BSAfE and the other two sub-varieties, which has had significant effects on the development of this sub-variety. Nevertheless, there is some recent evidence for convergence between WSAfE and BSAfE at the level of pronunciation and lexis. Given this complex sociolinguistic context, SAfE has not progressed as far along the stages of the Dynamic Model (with the different sub-varieties progressing through the model at different rates), and has not enjoyed the same level of codification as AusE. Although a handful of dictionaries have been produced for the variety, the set of norms governing the use of English in South Africa appears to be exonormative, with the orientation mostly towards BrE.

The contrast between the more homogenous linguistic landscape in Australia and the heterogenous linguistic landscape in South Africa carries through to editors' language profiles and choice of norm-providing sources as determined by the survey of these two groups of editors in this study. As demonstrated in Section 5.2.2, there are some striking differences between the two countries. Editors of English in Australia are first- or home-language users of English, and while some are multilingual (as a result of acquiring a heritage language or learning a European language at school) these additional languages do not extend to their editing work. Furthermore, AusE editors have a strong preference for local norm-providing sources, particularly local dictionaries, style guides and usage guides, and there is therefore a clear endonormative orientation in the sources used by these editors (see Section 5.2.3.2). Some exonormative sources are used: AmE sources feature strongly among the sources associated with lexis and style (such as dictionaries and style guides), while BrE sources feature more strongly among sources associated with grammar and usage. However, the preference for these sources is not nearly as strong as for AusE sources and appears to be partly motivated by the fact that they are freely accessible online – raising the possibility that sometimes the selection of norm-providing sources is less guided by explicit decisions about norm alignment, than it is by questions of cost or accessibility. The homogeneity in the language profiles of AusE editors together with the availability of and preference for local norm-providing sources means that the linguistic representations of AusE editors are most likely reflective of the Australian norm. Furthermore, given the dominance of English in AusE and its status as an Inner Circle variety, the exposure of AusE

editors to communal usages in unedited writing most likely reflect their own covert norms, resulting in linguistic editorial choices that converge on the Australian norm.

The language profiles of South African editors of English are, in comparison, more diffuse and represent a combination of profiles that include first- and second-language users of English (specifically WSAfE and AfrE), most of whom edit in English and Afrikaans (as demonstrated in Section 5.2.2). Both English and Afrikaans are therefore cognitively active in the largely bilingual language representation of many editors of SAfE, which acts as a point of contact between the two languages (in addition to contact between the two language communities), opening up opportunities for the two linguistic systems to influence one another in the process of editorial interventions. All editors of English in South Africa, at least those who responded to the questionnaire, are therefore WSAfE and AfrE users. BSAfE users are not represented in the corps of editors in this country, despite the fact that BSAfE is the most widely used sub-variety in the country. This means that the usage patterns of written BSAfE are evaluated by users of other strands of English, namely WSAfE and AfrE, placing a stricter acceptability requirement on the usages of BSAfE. Furthermore, the authors of English texts in South Africa represent the full spectrum of the sub-varieties of English, while editors represent two sub-varieties. This has two effects for the sub-varieties: first, there is a stricter acceptability requirement for innovative usages because acceptability must be attained across sub-varieties; and second, the linguistic representations of editors of English in South Africa are likely highly varied as a result of their exposure to a diffuse set of covert norms.

This heterogeneity in covert norms is also reflected in the kinds of overt norm-providing sources that editors of SAfE draw on and the orientation of these sources (as demonstrated in Section 5.2.3.3). In the absence of codification instruments for SAfE, the appeal of BrE is strong with most editors demonstrating a clear preference these sources. One striking finding for editors of SAfE, which mirrors their language profiles, is the preference for bilingual English–Afrikaans and monolingual Afrikaans sources. The use of these sources reflects the high level of English–Afrikaans bilingualism among editors of English in South Africa, but it also suggests that in the absence of explicit norm-providing sources for SAfE, bilingual editors utilise their bilingual proficiency and draw on sources written within the South African context (even if those sources are in or include another language). This opens up the editorial decisions of these editors to potential influence from Afrikaans, providing a strong case for the potential of cross-linguistic transfer effects, but also adds an additional layer of complexity to the competition between overt norms (in which the norms of two different languages are used) and covert norms.

Therefore, there is convergence between the overt and covert norms for both AusE authors and editors; however, in South Africa discrepancies arise not only between the covert norms of different author and editor groups, but also between overt norms where, in the absence of local norms for English in South Africa – competition arises between British and Afrikaans overt norms.

6.2.2.2 Findings for research question 2.3

The analysis of the corpus data sought to determine how these (complex) language profiles and the interaction between (different) overt and covert norms play out with respect to actual editing choices, focusing on one feature, namely genitive alternation. This was done by analysing differences in the frequency and selection of the two genitive variants in unedited and edited writing in a register-differentiated corpus of texts representing AusE, WSAfE, AfrE and BSAfE. The analysis was carried out in a step-wise fashion in order to demonstrate the importance of including editorial intervention in investigations of ongoing language change.

The first important finding relates to the overall distributional patterns of the two constructions in the edited (and therefore published) written texts for each register in each of the sub-varieties. The statistical evaluation in Section 5.3.1.3 showed that register, rather than variety, is the strongest predictor in conditioning genitive alternation: the *s*-genitive occurs in a significantly higher proportion in the creative, popular and reportage registers, compared to the academic and instructional registers. Cross-varietal differences only emerge for the academic and instructional registers, where the sub-varieties cluster together along country lines. AusE has the strongest preference for the *s*-genitive in these two registers, compared to the SAfE varieties. Given that AusE has been shown to be particularly receptive to features associated with colloquialisation (Collins, 2014), the higher proportion of *s*-genitives in the academic and instructional registers suggests that this receptivity extends even to the most formal of the registers. Within the South African sub-varieties, AfrE and WSAfE cluster together in their preference for the *s*-genitive in academic and instructional writing, providing evidence that the two sub-varieties function like a dialect cluster. BSAfE distinguishes itself as significantly less open to the use of the *s*-genitive in these two most formal registers, by comparison. Since very little is known about genitive alternation within the context of AusE and SAfE, and across different published written registers, these findings contribute new data to the study of genitive alternation from a cross-varietal and cross-register perspective and provide support for the argument by Hundt (2007) that patterns of inter- and intra-varietal patterns of variation of global trends are complex and subject to a number of factors that converge on questions of the interaction of stylistic preferences or pressures, substrate influence and contact phenomena in the internal developments of the different (sub-)varieties of English.

Section 5.3.2 shifted the focus to editorial intervention, and carried out a contrastive analysis of the preference for the *s*-genitive in the two varieties and across the different registers, but added the role of editorial intervention as an additional factor in influencing this preference (and therefore alternation). Drawing on a conditional inference tree that included REGISTER, SUB-VARIETY and EDITED STATUS as predictor variables for the proportion of *s*-genitives as response variable, the findings demonstrate an effect of editorial intervention across the sub-varieties and registers, but the effect differs across the varieties, and within each variety across the different registers. A striking difference that emerges when editorial intervention is added as a predictor to the model is that the most important (first-level) predictor

is SUB-VARIETY – unlike in the model for the edited writing only, in which the most important predictor is REGISTER. This suggests, in the very first instance, that editing removes some of the strong varietal distinctions in favour of a more homogenous norm in which register is a stronger predictor. In AusE, editors mostly accept the proportions of the two constructions across the registers, which points to the endonormative stabilisation of this variety. One striking finding in AusE is AusE editors' acceptance of the very high proportion of *s*-genitives in unedited academic writing, providing evidence for the acceptance of the wide-spread use of this informal construction in the very formal register, and reflecting convergence on the informal covert norm. In WSAfE, the effect of editorial intervention is greatest for the instructional register, where editors significantly reduce the proportion of *s*-genitives, and bring it in line with academic writing. Editors of WSAfE therefore smooth out differences in the usage patterns across the two formal registers in line with conventional expectations of these formal registers. Editorial intervention also plays a significant role in academic and instructional AfrE, but here the effect is in the opposite direction to WSAfE: editors increase the proportional frequencies of the *s*-genitive in these two registers, bringing the proportions of the two constructions much more in line with WSAfE, providing evidence for convergence between the two sub-varieties aided by editorial intervention. The very strong preference for the more formal construction in unedited BSAfE is reflective of the slower progress of this sub-variety along the stages of the Dynamic Model, but also points to an effect of substrate influence (where the *s*-genitive construction is unavailable in the Bantu languages). Edited status does not play any significant effect in shifting the proportion of *s*-genitives in the registers of BSAfE: in other words, editors of BSAfE, who are users of other sub-varieties, accept the genitive usage patterns of BSAfE authors, which provides evidence for endonormativity for this sub-variety.

The focus then shifted to include the linguistic factors known to influence the alternation between the two constructions in order to consider how different external/stylistic and linguistic factors and editorial intervention interact in conditioning the use of the two variants. The language-external factors considered include the different sub-varieties, the five registers, and the edited and unedited status of the texts, while the linguistic factors considered include the semantic relation between the possessor and possessum, animacy of the possessor, syntactic weight of the possessor and possessum, definiteness of the possessor and final sibilance.

The findings in Sections 5.3.3.1 and 5.3.3.2 show that linguistic factors are the most important conditioning variables in all four varieties, but that the effect of these factors differs across the sub-varieties, relative to their progression along the stages of the Dynamic Model, but also due to the influence of substrate languages for BSAfE and AfrE. For example, the animacy constrain is lowest in AfrE, which is in all likelihood a transfer effect from Afrikaans (where the animacy constraint is known to be low). The language-external factors play only a minimal role in the different sub-varieties: editing intervention plays no significant role in conditioning the alternation in AusE, WSAfE and AfrE (once the linguistic predictors are taken into account in the model), but it does play a role in BSAfE, where the language of the editor

influences which construction is selected. This suggests that the genitive usage practices of AusE, WSAfE and AfrE authors enjoy acceptance by editors of these varieties, who reinforce the usage patterns by accepting them and allowing them to disseminate further.

The animacy constraint is high in AusE, with a strong favouring for the *s*-genitive with animate possessors, while inanimate possessors are favoured by *of*-genitives. The evidence shows that the extension of the *s*-genitive to nouns lower on the animacy scale remains largely restricted in Australian English. This demonstrates that, in fact, AusE is not as open to the ongoing extension of the *s*-genitive as would be assumed by the higher proportion of *s*-genitives; instead, the high number of animate possessors (which is typically associated with colloquialisation) interacts with a strong animacy constraint, giving rise to a higher proportion of *s*-genitives, compared to the other sub-varieties investigated.

WSAfE and AfrE both participate in the ongoing extension of the *s*-genitive to possessors lower on the animacy scale in very similar ways (demonstrating convergence between the two most likely as a result of the prolonged contact between English and Afrikaans users in South Africa). Furthermore, this usage is accepted by editors. In both varieties, this extension is primarily to collective and locative possessors, and is restricted with inanimate and temporal possessors, except in the case of definite and proper noun possessors, which favour the *s*-genitive. This provides evidence that WSAfE and AfrE participate in the ongoing change in this grammatical feature, and that this is facilitated by a gradient and parallel loosening of the animacy constraint in the two sub-varieties. Furthermore, editors accept these innovative usages, legitimising them and contributing to their further dissemination. A noteworthy finding for AfrE is that for inanimate and temporal possessors, the presence of a final sibilant interacts with register to facilitate the use of these possessors with *s*-genitives. Specifically, the more compact *s*-genitive is preferred with inanimate and temporal possessors in registers targeted at more specialised readerships – the academic and instructional registers. Therefore, the extension of the *s*-genitive to the least animate possessors is creeping into formal AfrE in the more formal of the two registers, largely in line with the findings by (Biber & Gray, 2010, 2012) that readers of these highly specialised informationally-dense formal registers are able to process the kind of increased lexical density associated with the more compact construction.

A different picture emerges for BSAfE, which is the only sub-variety in which the edited status of the text and the language of the editor play a role. This is noteworthy, particularly because this is the only sub-variety where the language profile of the editors differs substantially from that of the authors. The choices of BSAfE authors are subjected to a more stringent acceptability test than are the choices for WSAfE and AfrE where the language profile of editors is much more similar to authors'. Of particular interest is the fact that the effect of these two editor language profiles is to move the preference for genitive constructions in slightly different directions particularly for collective and locative possessors: WSAfE editors accept the BSAfE authors' preference for the *of*-genitive with these possessors, but AfrE editors do not and instead replace many of these *of*-genitives with *s*-genitives, so that the overall proportion of *s*-genitives is higher in edited BSAfE than in unedited writing. This shows that AfrE editors impose their own norms for the use

of the *s*-genitive (which is conditioned by the effect of the Afrikaans substrate), so that the constraining effect of WSAfE editors is counteracted by the more innovating effect of AfrE editors resulting in a very slight (non-significant) difference in the proportional frequencies of the two constructions across unedited and edited BSAfE.

There is therefore evidence that editorial intervention interacts with linguistic factors to legitimise the genitive usage patterns of AusE, WSAfE and AfrE authors, helping in the further dissemination and conventionalisation of these usages. In BSAfE, the genitive usage patterns enjoy acceptance by WSAfE editors, but AfrE editors do not accept these usages, particularly with nouns slightly lower on the animacy scale. In other words, the extension of the *s*-genitive to collective and locative nouns in AfrE is transferred to BSAfE via the editorial interventions of AfrE editors. This points to a complex interplay between covert and overt norms in different sub-varieties of English.

6.3 Findings for research question 3: Theoretical implications for studies of language change, World Englishes and across written registers

Based on the integration of theoretical and empirical work, how does editorial intervention interact with broader processes of language change, specific language change in the varieties of English and register effects to influence different opportunities for and constraints on the processes of dissemination and conventionalisation in a variety's progression towards endonormativity?

The findings of the study show that editorial intervention can (and does) interact with the broader processes of language change, specific change in varieties of English and register effects. This interaction can only be understood if it is (first) accepted that editorial intervention is an important mechanism in the diffusion and conventionalisation of features associated with ongoing change, and (second) understood that editorial intervention is carried out by language users who are like any other, and whose knowledge of language is built up in the same way and is subject to the same cognitive and social processes, mechanisms and constraints that affect other language users. What makes editors unique is the salience of overt norms in their cognitive representations of language, but these norms are not necessarily more strongly represented than the covert norms that are entrenched through their exposure to language use, and especially to unedited forms of language use (potentially produced by authors from another sub-variety). In other words, while overt norms are more saliently represented in editors' mental representations of language than they are for other language users, editors are also frequently exposed to changing and differing usage patterns through the process of reading and evaluating unedited writing.

The findings further show the importance of covert norms in influencing editors' acceptability judgements of features associated with variation and change in different varieties of English, but demonstrate that this is not simply a matter of a broad, national communal usage pattern. Instead, the findings clearly indicate that stylistically distinct usage patterns of features associated with ongoing change may arise at the level of

a variety more broadly, but may also arise in contexts where multiple sub-varieties interact with each other in the same context, providing evidence to support the argument of Van Rooy (2014a) that in high contact varieties it is necessary to study the development of the sub-varieties at a local level. Furthermore, these stylistically distinct patterns of features associated with ongoing change are informed by the unique interaction of many factors in different varieties of English, including differing stylistic preferences across sub-varieties; substrate influence; the amount, type and duration of contact among the language users of the different sub-varieties; the differing strength of linguistic factors; the progression of the sub-variety along the stages of the Dynamic Model; and the kinds of overt norms available.

The investigation of the interaction between linguistic and language-external factors revealed how complex this picture is but showed that editing, generally, endorses these distinctive usage patterns at the level of sub-varieties. A more fine-grained analysis at the level of registers within these sub-varieties in SAfE shows that in some cases, editing acts to subtly pull the preferences in particular directions, relative to the (sub-)variety being edited and the differing or similar preferences of the sub-variety to which the editor belongs, as is the case for the editing of BSAfE by WSAfE and AfrE users. Endonormativity is thus, in some cases, achieved at the level of the different sub-varieties in South Africa, but depending on the historical nature of contact between the sub-varieties and the usage patterns of the particular sub-varieties, such endonormativity reflects either convergence or divergence between the sub-varieties.

The findings of the empirical investigation thus provide support for the theoretical framework proposed and show that editorial intervention does play an important role in the broader processes of ongoing language change. In the case of the feature investigated in this study, editorial intervention is an important mechanism in the propagation and conventionalisation of shifting preferences for innovative usage patterns in the different varieties of English, rather than a mechanism that blocks their dissemination and conventionalisation in a community of language users, as is assumed by various scholars who comment on the potential influence of editorial intervention. This demonstrates how important evidence-based evaluations of the influence of different factors in studies of language variation and change are if we are to truly study the total linguistic fact (Dąbrowska, 2015, 2016; Geeraerts & Kristiansen, 2015; Hans-Jörg, 2016; Harder, 2010, 2015).

6.4 Limitations of the study and recommendations for further research

The findings of the study should be interpreted against the limitations associated with the use of a case study. The two greatest limitations relate to the investigation of only two varieties and one grammatical feature, and further research that replicates the current study but which includes other varieties and features needs to be carried out. Given that the feature analysed in this study is one that is not particularly saliently marked in norm-providing sources, future investigations could include features that are frequently noted by prescriptive commentators. The inclusion of such normatively marked features will allow for a more detailed investigation into how overt and covert norms interact, and may well yield results that differ from

the findings of the current study. Such features could include shibboleth features of a particular sub-variety or features that apply to broader trends more generally. In addition, the role of editorial intervention should be investigated in more varieties of English, to allow for a direct comparison of larger sets of varieties.

Some further limitations of the study include the size and balance of the corpus, and further research on larger, more balanced corpora is needed. In addition, this study has made use of a synchronic corpus and therefore provides a snapshot of current editing practices, but future work needs to address the issue of how editorial practices have shaped the development of the varieties over time. Such an approach could include a parallel investigation of changing prescriptive guidelines, tracing how editorial intervention responds to prescriptive recommendations over time and how this interacts with covert norms.

Although this study did collect and analyse information on the kinds of changes that editors make, more detailed investigations of *why* they make specific choices or *why* they prefer some normative sources over others (particularly in varieties that have not yet been codified) would add a rich layer to understanding their choices. Such information could, for example, be gleaned through the use of interviews which would provide much more rich data on the motivations behind editors' choices and the interaction between their overt and covert norms.

6.5 Concluding remarks

This study has investigated the role of editorial intervention in the processes of ongoing language change in genitive alternation in AusE and SAfE and across five published written registers. It has described how the editorial role can be positioned within the processes of language change by drawing on and integrating cognitive and sociolinguistic usage-based theories of language and language change, models of World Englishes and register studies. The study has demonstrated that editors play an important role shaping the language of written texts, and that this role largely contributes to the dissemination and conventionalisation of the features of change, relative to particular varieties of English. In doing so it has contributed to the understanding of the processes by which languages and varieties change, and has drawn attention to one mechanism of change, editorial intervention, that has largely been ignored in investigations of variability in World Englishes and in written language in particular.

Bibliography

- Society for Editors and Proofreaders (SfEP). (2015). A tribute to Judith Butcher. Retrieved from <https://blog.sfep.org.uk/a-tribute-to-judith-butcher/>
- Akinlotan, M. (2016). Genitive alternation in New Englishes: The case of Nigerian English. *Token: A Journal of English Linguistics*, 5, 59–73.
- Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of child language*, 42(2), 239–273. doi:10.1017/S030500091400049X
- Auer, A., Peersman, C., Pickl, S., Rutten, G., & Vosters, R. (2015). Historical sociolinguistics: The field and its future. *Journal of Historical Sociolinguistics*, 1(1), 1–12. doi:10.1515/jhsl-2015-0001
- Australian National Dictionary. (1988). (1st ed.). Melbourne: Oxford University Press.
- Backus, A. (2015). Rethinking Weinreich, Labov & Herzog from a usage-based perspective. *Taal en Tongval*, 67(2), 275–306. doi:10.5117/TET2015.2.BACK
- Backus, A., & Spotti, M. (2012). Normativity and change: Introduction to the special issue on *Agency and power in multilingual discourse*. *Sociolinguistic Studies*, 6(2), 185–208. doi:10.1558/sols.v6i2.185
- Bamgbose, A. (1998). Torn between the norms: innovations in world Englishes. *World Englishes*, 17(1), 1–14. doi:10.1111/1467-971X.00078
- Baxter, G., & Croft, W. (2016). Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change*, 28(2), 129–173. doi:10.1017/S0954394516000077
- Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *Linguistics*, 47(2), 383–411. doi:10.1515/LING.2009.014
- Behrens, H., & Pfänder, S. (Eds.). (2016). *Frequency Counts: Frequency Effects in Language*. Berlin: De Gruyter.
- Bekker, I. (2012). The story of South African English: A brief linguistic overview. *International Journal of Language, Translation and Intercultural Communication*(1), 139–150. doi:10.12681/ijltic.16
- Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, 10(3), 210–223. doi:10.2167/beb441.0
- Biber, D. (1988). *Variation across Speech and Writing*. New York: Cambridge University Press.

- Biber, D. (1995). *Dimensions of Register Variation: A Cross-linguistic Comparison*. Cambridge: Cambridge University Press.
- Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory*, 8(1), 9–37.
- Biber, D. (2015). Corpus-based and corpus-driven analyses of language variation and use. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press. Retrieved from <http://www.oxfordhandbooks.com/10.1093/oxfordhb/9780199677078.001.0001/oxfordhb-9780199677078-e-8>. doi:10.1093/oxfordhb/9780199677078.013.0008
- Biber, D., & Conrad, S. (2009). *Register, Genre and Style*. Cambridge: Cambridge University Press.
- Biber, D., & Conrad, S. (2013). Introduction: Multi-dimensional analysis and the study of register variation. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 3–13). Oxon: Routledge.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9, 2–20.
- Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: The influence of written language use. *English Language and Linguistics*, 15(2), 223–250. doi:10.1017/s1360674311000025
- Biber, D., & Gray, B. (2012). The competing demands of popularization vs. economy: Written language in the age of literacy. In T. Nevalainen & E. C. Tragot (Eds.), *The Oxford Handbook of the History of English* (pp. 314–328). New York: Oxford University Press.
- Biber, D., & Gray, B. (2013). Being specific about historical change: The influence of sub-register. *Journal of English Linguistics*, 41(2), 104–134.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finnegan, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson.
- Bisaillon, J. (2007). Professional editing strategies used by six editors. *Written Communication*, 24(4), 295–322. doi:10.1177/0741088307305977
- Branford, W. (1996). English in South African Society. In V. de Klerk (Ed.), *Focus on South Africa* (Vol. 15, pp. 35–51). Amsterdam: John Benjamins.
- Branford, W., & Claughton, J. S. (1995). Mutual lexical borrowings among some languages of Southern Africa: Xhosa, Afrikaans and English. In R. Mesthrie (Ed.), *Language and Social History: Studies in South African Sociolinguistics* (pp. 209–221). Cape Town: David Phillip Publishers.

Bibliography

- Bresnan, J., & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, 86(1), 168–213.
- Bresnan, J., Rosenbach, A., Szmrecsanyi, B., Tagliamonte, S. A., & Simon, T. (2017). *Syntactic Alternations Data: Dative and Genitives in Four Varieties of English*. Retrieved from: <http://purl.stanford.edu/qj187zs3852>
- Butcher, J., Drake, C., & Leach, M. (2006). *Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors and Proofreaders* (3rd ed.). Cambridge: Cambridge University Press.
- Butterfield, J. (Ed.) (2015). *Fowler's Dictionary of Modern English Usage* (4th ed.). Oxford: Oxford University Press.
- Bybee, J. (2007a). Diachronic linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 945–987). Oxford: Oxford University Press.
- Bybee, J. (2007b). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Bybee, J. (2013). Usage-based theory and exemplar representations of constructions. In T. Hoffman & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 49–69). Oxford: Oxford University Press.
- Bybee, J. (2015). *Language Change*. Cambridge: Cambridge University Press.
- Bybee, J., & Beckner, C. (2014). Language use, cognitive processes and linguistic change. In C. Bowerman & B. Evans (Eds.), *The Routledge Handbook of Historical Linguistics* (pp. 503–518). London: Routledge.
- Bybee, J., & Beckner, C. (2015). Usage-based theory. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (2nd ed.). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199677078.013.0032
- Cameron, D. (2012). *Verbal Hygiene* (2nd ed.). London: Routledge.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences*, 7(7), 300–307. doi:10.1016/S1364-6613(03)00136-0
- Clark, L., & Trousdale, G. (2010). A cognitive approach to quantitative sociolinguistic variation: Evidence from th-fronting in Central Scotland. In D. Geeraerts, G. Kristiansen, & Y. Peirsman (Eds.), *Advances in Cognitive Sociolinguistics* (Vol. 45, pp. 291–321). Berlin: De Gruyter Mouton.

- Coetzee-Van Rooy, S. (2013). Afrikaans in contact with English: Endangered language or case of exceptional bilingualism? *International Journal of the Sociology of Language*, 224, pp. 179–207. doi 10.1515/ijsl-2013-0060.
- Coetzee-Van Rooy, S. (2014). Explaining the ordinary magic of stable African multilingualism in the Vaal Triangle region in South Africa. *Journal of Multilingual and Multicultural Development*, 35(2), 121–138. doi:10.1080/01434632.2013.818678.
- Collins, P. (2012). Grammatical variation in English worldwide: The role of colloquialization. *Linguistics & the Human Sciences*, 8(3), 289–306. doi:10.1558/lhs.v8i3.289.
- Collins, P. (2014). Australian English. In H. Koch & R. Nordlinger (Eds.), *The Languages and Linguistics of Australia: A Comprehensive Guide* (pp. 449–484): De Gruyter. Retrieved from <https://www.degruyter.com/view/books/9783110279771/9783110279771.449/9783110279771.449.xml>. doi:10.1515/9783110279771.449
- Cornelius, E. (2015). Defining 'plain language' in contemporary South Africa. *Stellenbosch Papers in Linguistics*, 44, 1–18.
- Cox, F., & Palethorpe, S. (2012). Standard Australian English: The sociostylistic broadness continuum. In R. Hickey (Ed.), *Standards of English: Codified Varieties Around the World* (pp. 294–317). Cambridge: Cambridge University Press.
- Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. Harlow: Longman.
- Croft, W. (2008). Evolutionary models and functional-typological theories of language change. In *The Handbook of the History of English* (pp. 68–91): Blackwell Publishing Ltd.
- Croft, W. (2010). Language Evolution: Contact, Competition and Change by Salikoko S. Mufwene. *World Englishes*, 29(2), 306–311. doi:10.1111/j.1467-971X.2010.01650.x
- Croft, W. (2013). Evolution: Language use and the evolution of languages. In P. Binder & K. Smith (Eds.), *The Language Phenomenon: Human Communication from Milliseconds to Millennia* (pp. 93–120). Heidelberg: Springer Berlin Heidelberg.
- Dąbrowska, E. (2015). Language in the mind and in the community. In J. Daems, E. Zenner, K. Heylen, D. Speelman, & H. Cuyckens (Eds.), *Change of paradigms – New paradoxes: Recontextualizing language and linguistics* (pp. 221–235). Berlin: De Gruyter.
- Dąbrowska, E. (2016). Cognitive linguistics' seven deadly sins. *Cognitive linguistics*, 27(4), 479–491. doi:10.1515/cog-2016-0059
- de Klerk, V. (1996). Introduction. In V. de Klerk (Ed.), *Focus on South Africa* (pp. 1–17). Amsterdam: John Benjamins.
- de Klerk, V. (1999). Black South African English: Where to from here? *World Englishes*, 18(3), 311–324.

Bibliography

- de Klerk, V., & Gough, D. (2002). Black South African English. In R. Mesthrie (Ed.), *Language in South Africa* (pp. 356–378). Cambridge: Cambridge University Press.
- Delbridge, A. (1999). Standard Australian English. *World Englishes*, 18(2), 259–270. doi:10.1111/1467-971X.00138
- Deutscher, G. (2005). *The Unfolding of Language: An Evolutionary Tour of Mankind's Greatest Invention*. New York: Henry Holt.
- Diessel, H. (2011). Language, usage and cognition (review). *Language*, 87(4), 830–844.
- Diessel, H. (2015). Usage-based construction grammar. In E. Dąbrowska & D. Dagmar (Eds.), *Handbook of Cognitive Linguistics* (pp. 296–321). Boston: De Gruyter.
- Diessel, H. (2016). Frequency and lexical specificity in grammar: A critical review. In H. Behrens & S. Pfänder (Eds.), *Experience Counts: Frequency Effects in Language* (pp. 209–237). Berlin: De Gruyter.
- Diessel, H., & Hilpert, M. (2016). Frequency effects in grammar. In M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. doi 10.1093/acrefore/9780199384655.013.120
- Divjak, D., & Arppe, A. (2013). Extracting prototypes from exemplars: What can corpus data tell us about concept representation? *Cognitive linguistics*, 24(2), 221–274. doi:10.1515/cog-2013-0008
- DoE. (2008). *CESM: Classification of Educational Subject Matter*. Pretoria: Government Printer. Retrieved from http://www.che.ac.za/sites/default/files/publications/CESM_August2008.pdf.
- Dykman, D. J., Geldenhuys, J. D. U., & Viljoen-Smook, E. E. (2008). *The Write Stuff: The Style Guide with a Difference*. Cape Town: Pharos.
- Ellis, N. C. (2011). The emergence of language as a complex adaptive system. In J. Simpson (Ed.), *The Routledge Handbook of Applied Linguistics* (pp. 654–667). Abingdon: Routledge.
- Evans, V. (2017). Cognitive linguistics. In S. E. F. Chipman (Ed.), *The Oxford Handbook of Cognitive Science*. Oxford: Oxford University Press.
- Feist, J. (2012). What controls the "genitive variation" in Present-Day English? *Studies in Language*, 36(2), 261–299. doi:10.1075/sl.36.2.03fei
- Ford, M., & Bresnan, J. (2015). Generating data as a proxy for unavailable corpus data: The contextualised sentence completion task. *Corpus Linguistics and Linguistic Theory*, 11(1), 187–224. doi:10.1515/cllt-2015-0018
- Freeman, R. (2010). Black and White: In search of an 'apt' response to Indigenous writing. *Text: journal of writing and writing courses*, 14(2), 1–17.

- GCIS. (2013). *Editorial Style Guide*. Retrieved from https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/guidelines/Editorial_Style_Guide.pdf
- Geeraerts, D., & Kristiansen, G. (2015). Variationist linguistics. In E. Dabrowska & D. Dagmar (Eds.), *Handbook of Cognitive Linguistics* (pp. 366–389). Boston: De Gruyter.
- Giliomee, H. (2003). The making of the Apartheid plan, 1929-1948. *Journal of Southern African Studies*, 29(2), 373–392. doi:10.1080/03057070306211
- Goldberg, A. E. (2013). Constructionist approaches. In T. Hoffman & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 15–31). Oxford: Oxford University Press.
- Grafmiller, J. (2014). Variation in English genitives across modality and genres. *English Language and Linguistics*, 18(3), 471–496. doi:10.1017/S1360674314000136
- Grafmiller, J., Szmrecsanyi, B., Röthlisberger, M., & Heller, B. (2018). General introduction: A comparative perspective on probabilistic variation in grammar. *Glossa: A Journal of General Linguistics*, 3(1). doi:10.5334/gjgl.690
- Gries, S. T., & Stefanowitsch, A. (2004). Extending collocation analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics*, 9(1), 97–129. doi:10.1075/ijcl.9.1.06gri
- Guy, G. R. (1991). Australia. In J. Cheshire (Ed.), *English around the World: Sociolinguistic Perspectives* (pp. 213–226). Cambridge: Cambridge University Press.
- Hans-Jörg, S. (2016). Why Cognitive linguistics must embrace the social and pragmatic dimensions of language and how it could do so more seriously. *Cognitive linguistics*, 27(4), 543–557.
- Harder, P. (2010). *Meaning in Mind and Society: A Functional Contribution to the Social Turn in Cognitive Linguistics*. Berlin: De Gruyter Mouton.
- Harder, P. (2015). Cognitive sociolinguistics, language systems and the fall of empires. In J. Daems, E. Zenner, K. Heylen, D. Speelman, & H. Cuyckens (Eds.), *Change of Paradigms – New Paradoxes: Recontextualizing Language and Linguistics* (pp. 237–251). Berlin: Mouton de Gruyter.
- Heller, B. (2018a). *Stability and Fluidity in Syntactic Variation World-Wide: The Genitive Alternation Across Varieties of English*. (PhD), KU Leuven, Leuven.
- Heller, B. (2018b). *When context shapes grammar: Stylistic flexibility in the English genitive alternation*. Paper presented at the International Congress of Linguists ICL20, Cape Town.
- Heller, B., Bernaisch, T., & Gries, S. T. (2017a). Empirical perspectives on two potential epicentres: The genitive alternation in Asian Englishes. *ICAME Journal*, 41, 111–144. doi:10.1515/icame-2017-0005

Bibliography

- Heller, B., Szmrecsanyi, B., & Grafmiller, J. (2017b). Stability and Fluidity in Syntactic Variation World-Wide: The Genitive Alternation Across Varieties of English. *Journal of English Linguistics*, 45(1), 3–27. doi:10.1177/0075424216685405
- Hendry, O. (2011). *The (Essential) Companion for Writers and Editors*. Cape Town: Pharos.
- Hilpert, M. (2014). *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.
- Hinrichs, L., & Szmrecsanyi, B. (2007). Recent changes in the function and frequency of Standard English genitive constructions: a multivariate analysis of tagged corpora. *English Language and Linguistics*, 11(3), 437–474. doi:10.1017/S136067430700234
- Hirson, B. (1981). Language in control and resistance in South Africa. *African Affairs*, 80(319), 219–237.
- Hothorn, T., & Zeileis, A. (2015). partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, 16, 3905–3909.
- Hruschka, D. J., Christiansen, M. H., Blythe, R., Croft, W., Heggarty, P., Mufwene, S. S., . . . Poplack, S. (2009). Building social cognitive models of language change. *Trends in Cognitive Sciences*, 13(11), 464–469.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hull, D. L. (1988). *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.
- Hundt, M. (2007). Global feature – local norms? A case study of the progressive passive. In T. Hoffman & L. Siebers (Eds.), *World Englishes – Problems, Properties and Prospects* (pp. 287–308). Amsterdam: John Benjamins.
- Hundt, M., & Mair, C. (1999). "Agile" and "Uptight" genres: The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics*, 4(2), 221–242.
- Hundt, M., & Szmrecsanyi, B. (2012). Animacy in Early New Zealand English. *English World-Wide*, 33(3), 241–263. doi: 10.1075/eww.33.3.01hun
- IPEd. (2016). *Report of the IPEd National Survey of Editors*. Retrieved from <http://iped-editors.org/site/DefaultSite/filesystem/documents/IPEd%20National%20Survey%20of%20Editors%202016%20%20FINAL%2020072017.pdf>
- Iwasaki, S. (2015). A multiple-grammar model of speakers' linguistic knowledge. *Cognitive linguistics*, 26(2), 161–210. doi:10.1515/cog-2014-0101
- Jankowski, B. L. (2013). *A Variationist Approach to Cross-register Language Variation and Change*. (PhD), University of Toronto, Toronto.

- Jankowski, B. L., & Tagliamonte, S. A. (2014). On the genitive's trail: Data and method from a sociolinguistic perspective. *English Language and Linguistics*, 18(2), 305–329. doi:10.1017/S1360674314000045
- Janse van Rensburg, C. (2016). Ambagsgereedskap: Die Redigeerder en proefleser se gereedskapkis. Retrieved from <https://www.editors.org.za/GenericPage.aspx?code=22>
- Kachru, B. (1985). Standards, codification, and sociolinguistic realism: the English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the World: Teaching and Learning the Language and Literatures* (pp. 11–30). Cambridge: Cambridge University Press for the British Council.
- Kachru, B. (1990). World Englishes and applied linguistics. *World Englishes*, 9(1), 3–20.
- Kachru, B. (1992). World Englishes: approaches, issues and resources. *Language Teaching*, 25(1), 1–14. doi:10.1017/S0261444800006583
- Kachru, B. (1996). World Englishes: Agony and ecstasy. *Journal of Aesthetic Education*, 30(2), 135–155. doi:10.2307/3333196
- Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow & S. Kremmer (Eds.), *Usage-based Models of Language* (pp. 7–28). Stanford, CA: CSLI Publications.
- Kiesling, S. F. (2008). English in Australia and New Zealand. In B. B. Kachru, Y. Kachru, & C. L. Nelson (Eds.), *The Handbook of World Englishes* (pp. 74–89). Maldan, MA: Blackwell Publishing Ltd.
- Korhonen, M. (2017). *Perspectives on the Americanisation of Australian English: A Sociolinguistic Study of Variation*. (PhD), University of Helsinki, Helsinki.
- Kruger, H. (2017). The effects of editorial intervention: Implications for studies of the features of translated language. In G. de Sutter, I. Delaere, & M.-A. Lefer (Eds.), *Empirical translation studies: New methodological and theoretical traditions* (pp. 113–156). De Gruyter: Berlin.
- Kruger, H. (forthcoming). Does editing matter? Editorial work, endonormativity and convergence in written Englishes in South Africa. In R. Hickey (Ed.), *English in Multilingual South Africa*. Cambridge: Cambridge University Press.
- Kruger, H., & Bevan-Dye, A. (2010). Guidelines for the editing of dissertations and theses: A survey of editors' perceptions. *Southern African Linguistics and Applied Language Studies*, 28(2), 153–169.
- Kruger, H., & Bevan-Dye, A. (2013). The language editor's role in postgraduate research: A survey of supervisors' perceptions. *South African Journal of Higher Education*, 27(4).

Bibliography

- Kruger, H., & Smith, A. (2018). Colloquialization versus densification in Australian English: A Multidimensional analysis of the Australian Diachronic Hansard Corpus (ADHC). *Australian Journal of Linguistics*, 38(3), 293–328. doi:10.1080/07268602.2018.1470452
- Kruger, H., & Van Rooy, B. (2016). Syntactic and pragmatic transfer effects in reported-speech constructions in three contact varieties of English influenced by Afrikaans. *Language Sciences*, 56, 118–131. doi:<https://doi.org/10.1016/j.langsci.2016.04.003>
- Kruger, H., & Van Rooy, B. (2017). Editorial practice and the progressive in Black South African English. *World Englishes*, 36, 20–41. doi:10.1111/weng.12202
- Labov, W. (1963). The Social Motivation of a Sound Change. *WORD*, 19(3), 273–309. doi:10.1080/00437956.1963.11659799
- Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (1994). *Principles of Linguistic Change: Internal Factors* (Vol. 1). Cambridge, MA: Blackwell.
- Labov, W. (2001). *Principles of Linguistic Change: Social Factors* (Vol. 2). Malden, MA: Wiley-Blackwell.
- Labov, W. (2010). *Principles of Linguistic Change: Cognitive and Cultural Factors*. (Vol. 3.). Malden, MA: Wiley-Blackwell
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical prerequisites* (Vol. 1). Stanford, CA: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press.
- Lanham, L. W. (1996). A History of English in South Africa. In V. de Klerk (Ed.), *Focus on South Africa* (Vol. 15, pp. 19–34). Amsterdam: John Benjamins.
- Lanham, L. W., & MacDonald, C. A. (1979). *The Standard in South African English and its Social History*. Heidelberg: Groos.
- Laugesen, A. (2014). Dictionaries for a nation: the making of the Macquarie Dictionary (1981) and the Australian National Dictionary (1988). *Journal of Australian Studies*, 38(1), 52–67. doi:10.1080/14443058.2013.865659
- Law, M. A. (2011). The development of core standards for editing in South Africa. *Southern African Linguistics and Applied Language Studies*, 29(3), 275–292. doi:10.2989/16073614.2011.647491
- Law, M. A. (2014). Factors influencing editorial work within the sectors of the South African editing industry. *Southern African Linguistics and Applied Language Studies*, 32(3), 285–299. doi:10.2989/16073614.2013.837614

- Law, M. A., & Kruger, H. (2008). Towards the professionalisation of editing in South Africa. *Southern African Linguistics and Applied Language Studies*, 26(4), 479–493. doi:10.2989/SALALS.2008.26.4.6.678
- Le Roux, E., & Cassells, L. (2018). *Annual Book Publishing Industry Survey Report 2016*. Retrieved from Pretoria: <http://www.publishsa.co.za/file/1519203677awr-2016publishingindustrysurvey.pdf>
- Leech, G., Hundt, M., Mair, C., & Smith, N. (2009). *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Leech, G., Smith, N., & Rayson, P. (2012). English style on the move: Variation and change in stylistic norms in the twentieth century. In M. Kytö (Ed.), *English Corpus Linguistics: Crossing Paths* (pp. 69–98). Amsterdam: Rodopi.
- Leitner, G. (2004). *Australia's Many Voices: Australian English – The National Language* (Vol. 1). Berlin: Mouton de Gruyter.
- Levshina, N. (2015). *How to do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam: John Benjamins.
- Linnegar, J. (2009). *Engleish, our Engleish: Common Errors in South African English and How to Resolve them*. Cape Town: Pharos.
- Linnegar, J. (2016). Tools of the Trade: Your Editorial Armoury. Retrieved from <https://www.editors.org.za/GenericPage.aspx?code=22>
- Lukač, M. (2016). Grammar Advice in the Age of Web 2.0: Introducing the new (and keeing the old) language authorities. *English Today*, 32(2), 2–3. doi:10.1017/S0266078416000134
- Lukač, M. (2017). From usage guides to language blogs. In I. Tiekens-Boon van Ostade (Ed.), *English Usage Guides: History, Advice, Attitudes*. Oxford: Oxford University Press. Retrieved from <http://www.oxfordscholarship.com/10.1093/oso/9780198808206.001.0001/oso-9780198808206-chapter-7>. doi:10.1093/oso/9780198808206.003.0007
- Mackenzie, J. (2011). *The Editor's Companion* (2nd ed.). New York: Cambridge University Press.
- Macquarie Dictionary. (1981). (1st ed.). Sydney: Macquarie Library.
- Mair, C. (2006). *Twentieth-century English: History, Variation and Standardization*. Cambridge: Cambridge University Press.
- Makalela, L. (2004). Making sense of BSAE for linguistic democracy in South Africa. *World Englishes*, 23(3), 355–366. doi:10.1111/j.0883-2919.2004.00363.x
- Mesthrie, R. (2002). Introduction. In R. Mesthrie (Ed.), *Language in South Africa* (pp. 1–8). Cambridge: Cambridge University Press.

Bibliography

- Mesthrie, R. (2009). Contact Linguistics and World Englishes. In B. B. Kachru, Y. Kachru, & C. L. Nelson (Eds.), *The Handbook of World Englishes* (pp. 273–288). Oxford: Wiley-Blackwell.
- Mesthrie, R. (2010). Sociolinguistics in South Africa: A critical overview of current research. In M. Ball (Ed.), *The Routledge Handbook of Sociolinguistics Around the World* (pp. 187–202). Oxon: Routledge.
- Mesthrie, R., & Bhatt, R. M. (2008). *World Englishes: The Study of New Linguistic Varieties*. Cambridge: Cambridge University Press.
- Milroy, L., & Milroy, J. (1992). Social network and social class: Toward an integrated sociolinguistic model. *Language in Society*, 21(01), 1–26.
- Minow, V. (2010). *Variation in the Grammar of Black South African English*. Frankfurt: Peter Lang.
- Mossop, B. (2014). *Revising and editing for translators* (4th ed.): Routledge.
- Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.
- Mufwene, S. S. (2008). *Language evolution: Contact, competition and change*. London: Continuum.
- Mufwene, S. S., & Vigouroux, C. B. (2012). Individuals, populations, and timespace: Perspectives on the ecology of language. *Cahiers de Linguistique*, 38(2), 111–138.
- Murray-Smith, S. (1987). *Right Words*. Melbourne: Viking.
- Nelson, G. (2002). ICE Markup Manual for Written Texts. Retrieved from <http://ice-corpora.net/ice/manuals.htm>
- Nelson, G., Wallis, S., & Aarts, B. (2002). *Exploring Natural Language: Working with the British Component of the International Corpus of English* (Vol. G29). Amsterdam: John Benjamins.
- Nevalainen, T. (2008). Historical Sociolinguistics and Language Change. In *The Handbook of the History of English* (pp. 558–588): Blackwell Publishing Ltd.
- Noble, R. (2014). Foreword. In E. Flann, B. Hill, & L. Wang (Eds.), *The Australian Editing Handbook* (3rd ed.). Milton, Qld.: John Wiley & Sons.
- Onysko, A. (2016). Modeling world Englishes from the perspective of language contact. *World Englishes*, 35(2), 196–220. doi:10.1111/weng.12191
- PEG. (2018). *Results of 2018 rates survey*. Retrieved from <https://www.editors.org.za/Documents/PEG%20rates%20survey%202018.pdf>
- Peters, P. (2006). English usage: Prescription and description. In B. Arts & A. McMahon (Eds.), *The Handbook of English Linguistics* (pp. 759–780). Malden, MA: Blackwell.
- Peters, P. (2007). *The Cambridge Guide to Australian English Usage*. Melbourne: Cambridge University Press.

- Peters, P. (2009). Epilogue: Collective Findings and Conclusions. In P. Peters, P. Collins, & A. Smith (Eds.), *Comparative Studies in Australian and New Zealand ENGLISH: Grammar and Beyond* (pp. 387–399). Amsterdam: John Benjamins.
- Peters, P. (2014a). Differentiation in Australian English. In S. Buschfeld, T. Hoffman, M. Huber, & A. Kautzsch (Eds.), *The Evolution of Englishes: The Dynamic Model and Beyond* (pp. 107–125). Amsterdam: John Benjamins.
- Peters, P. (2014b). Usage guides and usage trends in Australian and British English. *Australian Journal of Linguistics*, 34(4), 581–598. doi:10.1080/07268602.2014.929082
- Peters, P. (2017). The lexicography of English usage: Describing usage variation and change. In I. Tiekens-Boon van Ostade (Ed.), *English Usage Guides: History, Advice, Attitudes*. Oxford: Oxford University Press. Retrieved from <http://www.oxfordscholarship.com/10.1093/oso/9780198808206.001.0001/oso-9780198808206-chapter-3>. doi:10.1093/oso/9780198808206.003.0003
- Peters, P., Collins, P., & Smith, A. (2009). *Comparative Studies in Australian and New Zealand English: Grammar and Beyond*. Amsterdam: John Benjamins.
- Peterson, P. G. (2009). Commas and connective adverbs. In P. Peters, P. Collins, & A. Smith (Eds.), *Comparative Studies in Australian and New Zealand English: Grammar and Beyond* (pp. 277–292). Amsterdam: John Benjamins.
- Pettman, C. (1913). *Africanderisms*. London: Longmans, Green & Co.
- Pfänder, S., & Behrens, H. (2016). Experience counts: An introduction to frequency effects in language. In H. Behrens & S. Pfänder (Eds.), *Experience Counts: Frequency Effects in Language* (Vol. 54, pp. 1–20). Berlin: De Gruyter.
- Pinker, S. (1995). *The Language Instinct: How the Mind Creates Language*. New York: HarperPerennial.
- R Core Team. (2018). R: A Language Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org>.
- Redelinghuys, K. (forthcoming). *Language Contact and Change through Translation in Afrikaans and South African English: A Diachronic Corpus-based Study*. PhD (Cotutelle), North-West University and Macquarie University, Vanderbijlpark and Sydney.
- Ritter, R. M. (Ed.) (2012). *New Oxford Style Manual* (2nd ed.). Oxford: Oxford University Press.
- Romaine, S. (1982). *Socio-historical Linguistics: Its Status and Methodology*. Cambridge: Cambridge University Press.
- Romaine, S. (2000). *Language in Society: An Introduction to Sociolinguistics* (2nd ed.). Oxford: Oxford University Press.

Bibliography

- Romaine, S. (2003). Variation in language and gender. In J. Holmes & M. Meyerhoff (Eds.), *The Handbook of Language and Gender* (pp. 98–118). Maiden, MA: Blackwell Publishing Ltd.
- Rosenbach, A. (2002). *Genitive Variation in English*. Berlin/New York: Mouton de Gruyter.
- Rosenbach, A. (2003). Aspects of iconicity and economy in the choice between the *s*-genitive and the *of*-genitive in English. In G. Rohdenburg & B. Mondorf (Eds.), *Determinants of Grammatical Variation in English*. Berlin: Mouton de Gruyter.
- Rosenbach, A. (2005). Animacy versus weight as determinants of grammatical variation in English. *Language*, 81(3), 613–644. doi:10.1353/lan.2005.0149
- Rosenbach, A. (2006). Descriptive genitives in English: A case study of constructional gradience. *English Language and Linguistics*, 10(1), 77–118.
- Rosenbach, A. (2007). Emerging variation: Determiner genitives and noun modifiers in English. *English Language and Linguistics*, 11(1), 143–189. doi:10.1017/S1360674306002140
- Rosenbach, A. (2008). Animacy and grammatical variation—Findings from English genitive variation. *Lingua*, 118, 151–171. doi:10.1016/j.lingua.2007.02.002
- Rosenbach, A. (2014). English genitive variation – the state of the art. *English Language and Linguistics*, 18(2), 215–262. doi:10.1017/S1360674314000021
- Rosenbach, A. (2017). Constraints in contact: Animacy in English and Afrikaans genitive variation – a cross-linguistic perspective. *Glossa: A Journal of General Linguistics*, 2(1), 1–21. doi:<https://doi.org/10.5334/gjgl.292>
- Rosenbach, A., & Vezzosi, L. (2000). Genitive constructions in Early Modern English: New evidence from a corpus analysis. In R. Sornicola, E. Poppe, & A. Shisha-Havely (Eds.), *Stability, Variation and Change of Word-order Patterns over Time* (pp. 285–307). Amsterdam: John Benjamins.
- Russi, C. (2016). Introduction. In C. Russi (Ed.), *Current Trends in Historical Sociolinguistics*. Warsaw, Poland: De Gruyter Open.
- Salazar, D. (2014). Towards improved coverage of Southeast Asian Englishes in the Oxford English Dictionary. *Lexicography*, 1(1), 95–108. doi:10.1007/s40607-014-0003-2
- Saldanha, G., & O'Brien, S. (2013). *Research Methodologies in Translation Studies*. Oxon: Routledge.
- SATI. (2015). Find a language practitioner. Retrieved from http://translators.org.za/sati/cms/index.php?frontend_action=search&menu_item_id=23
- Schmid, H.-J. (2012). Linguistic Theories, Approaches, and Methods. In M. Middeke, T. Müller, C. Wald, & H. Zapf (Eds.), *English and American Studies: Theory and Practice* (pp. 371–394). Stuttgart: J.B. Metzler.

- Schneider, E. W. (2003a). The Dynamics of New Englishes: From Identity Construction to Dialect Birth. *Language*, 79(2), 233–281.
- Schneider, E. W. (2003b). Evolutionary patterns of New Englishes and the special case of Malaysian English. *Asian Englishes*, 6(2), 44–63. doi:10.1080/13488678.2003.10801118
- Schneider, E. W. (2007). *Postcolonial English: Varieties Around the World*. Cambridge: Cambridge University Press.
- Siegel, J. S. (2018). *Demographic and Sociolinguistic Basis of Ethnolinguistics*. Cham: Springer International Publishing.
- Silva, P. (1997). *South African English: Oppressor or Liberator*. Paper presented at the The Major Varieties of English: MAVEN 97, Vaxjo. Retrieved from <https://core.ac.uk/download/pdf/145046916.pdf>
- Spencer, B. (2011a). International sporting events in South Africa, identity re-alignment, and Schneider's EVENT X. *African Identities*, 9(3), 267–278. doi:10.1080/14725843.2011.591225
- Spencer, B. (2011b). Linguistic penetration at Schneider's Phase 4: Acceptability ratings of entrenched features of Black South African English by South Africans outside the originating culture of the variety. *Journal for Language Teaching*, 45(2), 133–149.
- Statistics South Africa. (2012). *Census 2011: Census in brief* (03-01-41). Retrieved from http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf
- Stefanowitsch, A. (2003). Constructional semantics as a limit to grammatical variation. In G. Rohdenburg & B. Mondorf (Eds.), *Determinants of Grammatical Variation in English* (pp. 155–173). Berlin & New York: Mouton de Gruyter.
- Straaijer, R. (2017). The usage guide: Evolution of a genre. In I. Tieken-Boon van Ostade (Ed.), *English Usage Guides: History, Advice, Attitudes*. Oxford: Oxford University Press. Retrieved from <http://www.oxfordscholarship.com/10.1093/oso/9780198808206.001.0001/oso-9780198808206-chapter-2>. doi:10.1093/oso/9780198808206.003.0002
- Szmrecsanyi, B. (2010). The English genitive alternation in a cognitive sociolinguistics perspective. In D. Geeraerts, G. Kristiansen, & Y. Peirsman (Eds.), *Advances in Cognitive Sociolinguistics* (pp. 141–166). Berlin: Mouton de Gruyter.
- Szmrecsanyi, B., Biber, D., Egbert, J., & Franco, K. (2016a). Toward more accountability: Modeling ternary genitive variation in Late Modern English. *Language Variation and Change*, 28, 1–29. doi:10.1017/S0954394515000198
- Szmrecsanyi, B., Grafmiller, J., Heller, B., & Röthlisberger, M. (2016b). Around the World in Three Alternations: Modeling Syntactic Variation in Varieties of English. *English World-Wide*, 37(2), 109–137. doi:doi 10.1075/eww.37.2.01szm

Bibliography

- Szmrecsanyi, B., & Hinrichs, L. (2008). Probabilist determinants of genitive variation in spoken and written English: A multivariate comparison across time, space and genres. In T. Nevalainen, I. Taavitsainen, P. Pahta, & M. Korhonen (Eds.), *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present* (pp. 291–310). Amsterdam/Philadelphia: John Benjamins.
- Tagliamonte, S. A. (2012). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.
- Tagliamonte, S. A., & Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, 24, 135–178.
- Team, R. C. (2018). R: A Language and Environment for Statistical Computing. from R Foundation for Statistical Computing <https://www.R-project.org/>
- Tredinnick, M. (2008). *The Little Green Grammar Book*. Sydney: University of New South Wales Press.
- Tummers, J., Heylen, K., & Geeraerts, D. (2005). Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*, 1(2), 225–261.
- Universities Australia. (2017). Data snapshot. Retrieved from <https://www.universitiesaustralia.edu.au/australias-universities/key-facts-and-data#.XChzhlwza70>.
- Van Rooy, B. (2010). Social and linguistic perspectives on variability in world Englishes. *World Englishes*, 29(1), 3–20. doi:10.1111/j.1467-971X.2009.01621.x
- Van Rooy, B. (2011). A principled distinction between error and conventionalized innovation in African Englishes. In J. Mukherjee & M. Hundt (Eds.), *Exploring Second-Language Varieties of English and Learner Englishes: Bridging the Paradigm Gap* (pp. 189–208). Amsterdam: John Benjamins.
- Van Rooy, B. (2014a). Convergence and Endonormativity at Phase 4 of the Dynamic Model. In S. Buschfeld, T. Hoffman, M. Huber, & A. Kautzsch (Eds.), *The Evolution of Englishes: The Dynamic Model and beyond* (pp. 21–38). Amsterdam: John Benjamins.
- Van Rooy, B. (2014b). English in South Africa. In M. Filppula, J. Klemola, & D. Sharma (Eds.), *The Oxford Handbook of World Englishes*. New York: Oxford University Press. doi:10.1093/oxfordhb/9780199777716.013.017
- Van Rooy, B., & Kruger, H. (2015). The case for an emergentist approach. *Stellenbosch Papers in Linguistics Plus*, 48, 41–67. doi:10.5842/48-0-676
- Van Rooy, B., & Kruger, H. (2016). The innovative progressive aspect of Black South African English: The role of language proficiency and normative pressures. *International Journal of Learner Corpus Research*, 2(2), 205–228. doi:10.1075/ijlcr.2.2.04van

- Van Rooy, B., & Terblanche, L. (2010). Complexity in word-formation processes in new varieties of South African English. *Southern African Linguistics and Applied Language Studies*, 28(4), 357–374. doi:10.2989/16073614.2010.548022
- Van Rooy, B., Terblanche, L., Haase, C., & Schmied, J. (2010). Register differentiation in East African English: A multidimensional study. *English World-Wide*, 31(3), 311–349.
- Von Mengden, F., & Croussé, E. (2014). Introduction. In E. Croussé & F. Von Mengden (Eds.), *Usage-based Approaches to Language Change* (pp. 1–19). Amsterdam: John Benjamins.
- Wasserman, G. P. (2014a). 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. doi:10.1177/0075424213514588
- Wasserman, G. P. (2014b). *Modality on Trek: Diachronic Changes in Written South African English across Text and Context*. (PhD in Linguistics), North-West University, Vanderbijlpark.
- Weinreich, U., Labov, W., & Herzog, M. I. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for Historical Linguistics* (pp. 95–195). Austin: University of Texas Press.
- Wodak, R., Johnstone, B., & Kerswill, P. (2011). Introduction. In R. Wodak, B. Johnstone, & P. Kerswill (Eds.), *The SAGE Handbook of Sociolinguistics* (pp. 1–8). London: SAGE.
- Wolk, C., Bresnan, J., Rosenbach, A., & Szmrecsanyi, B. (2013). Dative and genitive variability in Late Modern English: Exploring cross-constructural variation and change. *Diachronica*, 30(3), 382–419. doi: 10.1075/dia.30.3.04wol

Appendix A: Draft questionnaire for pilot

Survey: Editors' norm-providing sources

Chief Investigator's / Supervisor's Name & Title: Dr Haidee Kruger and Prof. Bertus van Rooy

Participant Information and Consent Form

Name of Project: The role of editorial intervention in ongoing language variation and change in South African and Australian English

You are invited to participate in a study of editorial intervention in published English texts. The purpose of the study is to determine what role editors play in the processes of language variation and change, and how their editorial decisions are guided by norm-providing sources. We would like to examine this in two varieties of English: Australian English and South African English.

The study is being conducted by Melanie Ann Law of Macquarie University, Sydney, Australia (email: melanie.law@students.mq.edu.au) and North-West University, Vaal Triangle Campus, South Africa (email: melanie.law@nwu.ac.za). The study will be conducted to meet the requirements of the degrees Doctor of Philosophy in Linguistics at Macquarie University, Sydney, Australia and Doctor of Philosophy in Linguistics and Literary Studies at North-West University, Vaal Triangle Campus, South Africa, under a Joint Doctoral Supervision Agreement between Macquarie University and North-West University, and under the supervision of:

- Dr Haidee Kruger (haidee.kruger@mq.edu.au) of the Department of Linguistics, Macquarie University
- Prof. Bertus van Rooy (bertus.vanrooy@nwu.ac.za) of the School of Languages, UPSET Research Focus Area, North-West University

If you decide to participate, you will be asked to complete the online questionnaire that follows this information letter. This questionnaire has three sections:

- Background information
- Sources used during editing
- Strategies for solving editing problems

The questionnaire should take approximately **30 minutes** to complete.

There is **no risk or discomfort** involved in the completion of the questionnaire.

Any information or personal details gathered in the course of the study are confidential, except as required by law. No individual will be identified in any publication of the results. Only the investigators named above will have access to the data. A summary of the results of the data can be made available to you on request by contacting Melanie Law on one of the email addresses above.

Participation in this study is entirely voluntary: you are not obliged to participate and if you decide to participate, you are free to withdraw at any time and without having to give a reason and without consequence.

The ethical aspects of this study have been approved by the Macquarie University Human Research Ethics Committee (ethics clearance reference number: 5201600843D) and the North-West University Research Ethics Committee for Language Matters (ethics clearance reference number: NWU-00558-16-A8). If you have any complaints or reservations about any ethical aspect of your participation in this research, you may contact the Macquarie University Human Research Ethics Committee through the Director, Research Ethics & Integrity (telephone +61 298 507 854; email ethics@mq.edu.au), or the North-West University

Research Ethics Committee for Language Matters through the Chairperson, Prof. Bertus van Rooy, (telephone +27 16 910 3482; email bertus.vanrooy@nwu.ac.za). Any complaint you make will be treated in confidence and investigated, and you will be informed of the outcome.

Please indicate whether you give your consent to participate in this study:

- I consent to participate in this study (1)
- I do not consent to participate in this study (2)

Instructions for completing the questionnaire

1. Please answer all questions.
2. Please read all questions carefully.
3. Complete the questionnaire using the selection options or text boxes provided.

Section 1: Background information

This section comprises several questions related to your demographic information (including your age, language profile, qualification, working languages, areas of experience and years' editing experience).

Please complete all questions in this section.

A1 Please select your age from the options below.

- 18-29 years old
- 30-39 years old
- 40-49 years old
- 50-59 years old
- 60 years and older

A2 What language do you regard as your first or home language?

- Afrikaans
- An indigenous Australian language
- An indigenous South African language
- English
- Other

Display the following question:

A2.1 Please use the text box below to indicate your first or home language.

If *What language do you regard as your first or home language?* "Other" is selected

Or *What language do you regard as your first or home language?* "An indigenous Australian Language" is selected

Or *What language do you regard as your first or home language?* "One of the nine indigenous South African languages" is selected

A3 To which gender identity do you most identify?

- Prefer not to answer
- Female
- Male
- Other (please specify) _____

A4 In the text boxes below, please list all the languages you know, up to a maximum of five. List the languages in order, from your strongest (1) to your weakest (5) language.

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

A5 In the text boxes below, please list all the languages in which you do editing work, up to a maximum of five. List the languages in order, from the language you edit in most often (1) to the language you edit in least often (5).

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

A6 From the options available, please select the country in which you currently reside.

- Australia
- South Africa
- Other

Display the following question:
A7 Please use the text box to indicate the country in which you currently reside.

If *From the options available, please select the country you currently reside in. "Other" is selected*

A8 Have you always lived in the country in which you currently reside?

- Yes
- No

Display the following question:
A8.1 How many years have you lived in the country in which you currently reside?

- 0-10 years (1)
- 11-20 years (2)
- 21-30 years (3)
- 31-40 years (4)
- 41-50 years (5)
- 51 years or more (6)

If *Have you always lived in the country in which you currently reside? "No" is selected*

Display the following question:

A9 Please use the first text box to indicate the country in which you were born and the number of years in which you lived in this country. If you have lived in countries other than the one(s) in which you were born and currently reside, please also list these countries and the number of years spent in each, in the second text box below.

I was born in: _____

I have lived in: _____

If *Have you always lived in the country in which you currently reside?* “No” is selected

A10 How many years’ editing experience do you have?

- 1-10 years (1)
- 11-20 years (2)
- 21-30 years (3)
- 31-40 years (4)
- 41 years and more (5)

A11 Do you work as a full-time editor or part-time editor?

- Full-time (1)
- Part-time (2)

Display the following question:

A11.1 In addition to editing, please list the other kind(s) of work you do in the text box below.

If *Do you work as a full-time editor or part-time editor?* “Part-time” is selected

A12 Do you work as a freelance or self-employed editor?

- Yes (1)
- No (2)

Display the following question:

A12.1 Please provide a short description of the organisation or company where you are employed. (For example, a community newspaper.)

If *Do you work as a freelance or self-employed editor?* “No” is selected

A13 What is your highest level of qualification?

- Year 12 / Matric
- First degree or equivalent
- Honours degree
- Masters degree
- Doctoral degree
- Other

Display the following question:

A13.1 Please use the text box to indicate your highest level of qualification.

If *What is your highest qualification?* “Other” is selected

Display the following question:

A14 In what discipline is this qualification? (For example, Accounting, Languages, Engineering, etc.)

If *What is your highest qualification?* “First degree or equivalent” is selected

Or *What is your highest qualification?* “Honours degree” is selected

Or *What is your highest qualification?* “Masters degree” is selected

Or *What is your highest qualification?* “Doctoral” is selected

Or *What is your highest qualification?* “Other” is selected

A15 From the options below, please select all text types in which you have editing experience.

- Academic articles appearing in scholarly journals
- Academic books
- Administrative texts (including minutes, institutional communication, etc.)
- Contracts and legal documents
- Dissertations and theses
- Instruction manuals
- Magazine articles
- Medical and pharmaceutical texts
- Newsletters
- Newspaper reports: online
- Newspaper reports: print
- Novels, novellas and short stories for leisure reading
- Plays
- Poetry
- Popular non-fiction books
- Schoolbooks
- Study guides
- Teacher's guides
- Website material
- Other

Display the following question:

A15.1 Please indicate any other text types in which you have editing experience in the text boxes below.

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____
- 6 _____
- 7 _____
- 8 _____

9 _____

10 _____

If From the options below, please select all text types in which you have editing experience. "Other" is selected

Carry Forward Entered Choices from A15 and Entered Text from A15.1

A16 In which text type do you have the most editing experience?

- Academic articles appearing in scholarly journals
- Academic books
- Administrative texts (including minutes, institutional communication, etc.)
- Contracts and legal documents
- Dissertations and theses
- Instruction manuals
- Magazine articles
- Medical and pharmaceutical texts
- Newsletters
- Newspaper reports: online
- Newspaper reports: print
- Novels, novellas and short stories for leisure reading
- Plays
- Poetry
- Popular non-fiction books
- Schoolbooks
- Study guides
- Teacher's guides
- Website material
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____
 - 6 _____
 - 7 _____
 - 8 _____
 - 9 _____
 - 10 _____

Carry Forward Entered Choices from A15 and Entered Text from A15.1

A17 In which text type is your most recent or current editing experience?

- Academic articles appearing in scholarly journals
- Academic books
- Administrative texts (including minutes, institutional communication, etc.)
- Contracts and legal documents
- Dissertations and theses
- Instruction manuals

Appendix A: Draft questionnaire for pilot

- Magazine articles
- Medical and pharmaceutical texts
- Newsletters
- Newspaper report: online
- Newspaper reports: print
- Novels, novellas and short stories for leisure reading
- Plays
- Poetry
- Popular non-fiction books
- Schoolbooks
- Study guides
- Teacher's guides
- Website material
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____
 - 6 _____
 - 7 _____
 - 8 _____
 - 9 _____
 - 10 _____

For the next question, you will be asked to select the option that most accurately describes how you view the function or aim of your editorial role. The description for each function is as follows:

Normative: Normative work aims to bring a text into conformance with predefined rules of usage as set out in dictionaries, house styles, style manuals, usage guides and other prescriptive sources.

Communicative: Communicative work aims to ensure that the text communicates optimally with its intended reader, and may make use of language that may not (yet) be set out in dictionaries, house styles, style manuals usage guides and other prescriptive sources.

A18 I view my editorial role as fulfilling

- an exclusively normative function. (1)
- an exclusively communicative function. (2)
- an equally normative and communicative function. (3)
- a primarily normative function, with communicative elements. (4)
- a primarily communicative function, with normative elements. (5)

A19 Please use the text box below to explain or motivate your answer.

Section 2: Sources used during editing

Editors rely on many **different sources** when editing. The questions that follow seek to gather information on the kinds of sources you use in the course of your editorial work.

If you do not make use of a particular source, please select the "I do not make use of (type of source) in the course of my editorial work" option.

For each source type you may select up to five sources.

B1 From the options provided below, please select up to five online dictionaries you use most often in the course of your editorial work.

- I do not use online dictionaries in the course of my editorial work.
- Australian National Dictionary
<http://australiannationaldictionary.com.au/index.php>
- Cambridge Free English Dictionary and Thesaurus
<http://dictionary.cambridge.org/dictionary/english/>
- Collins English Dictionary
<http://www.collinsdictionary.com/dictionary/english>
- Dictionary of South African English on Historical Principles
<http://dsae.co.za/>
- The Free Dictionary by Farlex
www.thefreedictionary.com
- Longman Dictionary of Contemporary English
<http://www.ldoceonline.com/>
- Macmillan Dictionary
<http://www.macmillandictionary.com/>
- The Macquarie Dictionary
<https://www.macquariedictionary.com.au/>
- Merriam-Webster Dictionary and Thesaurus
<http://www.merriam-webster.com/>
- Oxford English Dictionary (OED)
<http://www.oed.com/>
- Oxford Dictionaries
<https://en.oxforddictionaries.com/definition/>
- Oxford Reference Online
www.oxfordreference.com
- Wiktionary
https://en.wiktionary.org/wiki/Wiktionary:Main_Page
- Other

Display the following question:

B1.1 Please list any other online dictionaries that you use in the course of your work:

- 1 _____
- 2 _____
- 3 _____
- 4 _____

5 _____

If From the options provided below, please select the online dictionaries you use most often in the course of your editorial work. "Other" is selected.

Carry Forward Entered Choices from B1. and Entered Text B1.1

B1.2 Of the online dictionaries you selected or listed, which one do you use most often in the course of your work?

- Australian National Dictionary
<http://australiannationaldictionary.com.au/index.php>
- Cambridge Free English Dictionary and Thesaurus
<http://dictionary.cambridge.org/dictionary/english/>
- Collins English Dictionary
<http://www.collinsdictionary.com/dictionary/english>
- Dictionary of South African English on Historical Principles
<http://dsae.co.za/>
- The Free Dictionary by Farlex
www.thefreedictionary.com
Longman Dictionary of Contemporary English
- <http://www.ldoceonline.com/>
- Macmillan Dictionary
<http://www.macmillandictionary.com/>
- The Macquarie Dictionary
<https://www.macquariedictionary.com.au/>
- Merriam-Webster Dictionary and Thesaurus
<http://www.merriam-webster.com/>
- Oxford English Dictionary (OED)
<http://www.oed.com/>
- Oxford Dictionaries
<https://en.oxforddictionaries.com/definition/>
- Oxford Reference Online
www.oxfordreference.com
- Wiktionary
https://en.wiktionary.org/wiki/Wiktionary:Main_Page
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

B2 From the options below, please select up to five print dictionaries you use most often in the course of your editorial work. Please indicate the edition number(s) of your selections in the text boxes.

- I do not use print dictionaries in the course of my editorial work.
- Australian National Dictionary, Oxford University Press. Edition number(s):

- Australian Oxford Dictionary, Oxford University Press. Edition number(s):

- Cambridge Advanced Learner’s Dictionary, Cambridge University Press. Edition number(s):

- Collins English Dictionary – Complete and Unabridged, HarperCollins Publishers. Edition number(s): _____
- Collins Dictionary for Writers and Editors, HarperCollins Publishers. Edition number(s):

- Collins South African English Dictionary, HarperCollins Publishers. Edition number(s):

- English Dictionary for South Africa, Pharos. Edition number(s): _____
- Longman Dictionary of Contemporary English, Longman/Pearson. Edition number(s):

- The Macquarie Dictionary, Macquarie Library Pty Ltd, Macquarie University. Edition number(s):

- Merriam-Webster’s Collegiate Dictionary, Merriam Webster Inc. Edition number(s):

- New Oxford Dictionary for Writers and Editors, Oxford University Press. Edition number(s):

- Oxford South African Concise Dictionary, Oxford University Press. Edition number(s):

- Other

Display the following question:

B2.1 Please list any other dictionaries that you use in the course of your work (including information about editions) in the text boxes:

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

If *From the options below, please select the print dictionaries you use most often in the course of your editorial work. “Other” is selected*

Carry Forward Entered Choices from B2 and Entered Text from B2.1

B2.2 Of the print dictionaries you selected or listed, which one do you use most often in the course of your work?

- Australian National Dictionary, Oxford University Press. Edition number(s):

- Australian Oxford Dictionary, Oxford University Press. Edition number(s):

- Cambridge Advanced Learner’s Dictionary, Cambridge University Press. Edition number(s):

- Collins English Dictionary – Complete and Unabridged, HarperCollins Publishers. Edition number(s): _____

Appendix A: Draft questionnaire for pilot

- Collins Dictionary for Writers and Editors, HarperCollins Publishers. Edition number(s):

- Collins South African English Dictionary, HarperCollins Publishers. Edition number(s):

- English Dictionary for South Africa, Pharos. Edition number(s): _____
- Longman Dictionary of Contemporary English, Longman/Pearson. Edition number(s):

- The Macquarie Dictionary, Macquarie Library Pty Ltd, Macquarie University. Edition number(s):

- Merriam-Webster's Collegiate Dictionary, Merriam Webster Inc. Edition number(s):

- New Oxford Dictionary for Writers and Editors, Oxford University Press. Edition number(s):

- Oxford South African Concise Dictionary, Oxford University Press. Edition number(s):

- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

B3 From the options provided below, please select up to five usage guides you use most often in the course of your editorial work. Please indicate the edition number(s) of your selections in the text boxes.

- I do not use usage guides in the course of my editorial work.
- The Cambridge Guide to Australian English Usage, Peters, P, Cambridge University Press. Edition number(s): _____
- The Cambridge Guide to English Usage, (2004), Peters, P, Cambridge University Press.
- Collins Gem English Grammar, Hardie, RG, HarperCollins Publishers. Edition number(s):

- English, our English: Common Errors in South African English and How to Solve Them, (2009), Linnegar, J, Pharos.
- Fowler's Modern English Usage, Oxford University Press. Edition number(s):

- Garner's Modern English Usage, Oxford University Press. Edition number(s):

- The Little Green Grammar Book, (2008), Tredinnick, M, UNSW Press.
- Modern Australian Usage, Oxford University Press/Allen & Unwin. Edition number(s):

- The Oxford A-Z of Grammar & Punctuation, Seely, J, Oxford University Press. Edition number(s):

- Right Words: A Guide to English Usage in Australia, Murray-Smith, S, Viking/Penguin. Edition number(s): _____
- Usage and Abusage: A Guide to Good English, Partridge, E, Penguin Australia. Edition number(s):

- Other

Display the following question:

B3.1 Please list any other usage guides that you use in the course of your work (including information about editions) in the text boxes:

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

If From the options provided below, please select the usage guides you use most often in the course of your editorial work. "Other" is selected

Carry Forward Entered Choices from B3 and Entered Text from B3.1

B3.2 Of the usage guides you selected or listed, which one do you use most often in the course of your work?

- I do not use usage guides in the course of my editorial work.
- The Cambridge Guide to Australian English Usage, Peters, P, Cambridge University Press. Edition number(s): _____
- The Cambridge Guide to English Usage, (2004), Peters, P, Cambridge University Press.
- Collins Gem English Grammar, Hardie, RG, HarperCollins Publishers. Edition number(s): _____
- English, our English: Common Errors in South African English and How to Solve Them, (2009), Linnegar, J, Pharos.
- Fowler's Modern English Usage, Oxford University Press. Edition number(s): _____
- Garner's Modern English Usage, Oxford University Press. Edition number(s): _____
- The Little Green Grammar Book, (2008), Tredinnick, M, UNSW Press.
- Modern Australian Usage, Oxford University Press/Allen & Unwin. Edition number(s): _____
- The Oxford A–Z of Grammar & Punctuation, Seely, J, Oxford University Press. Edition number(s): _____
- Right Words: A Guide to English Usage in Australia, Murray-Smith, S, Viking/Penguin. Edition number(s): _____
- Usage and Abusage: A Guide to Good English, Partridge, E, Penguin Australia. Edition number(s): _____
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

B4 From the options provided below, please select up to five style manuals you use most often in the course of your editorial work. Please indicate the edition number(s) of your selections in the text boxes.

- I do not use style manuals in the course of my editorial work.
- AMA Manual of Style: A Guide for Authors and Editors, Oxford University Press. Edition number(s): _____
- The Australian Government Style Manual: For Authors, Editors and Printers, John Wiley. Edition number(s): _____
- Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors and Proofreaders, Cambridge University Press. Edition number(s): _____
- Chicago Manual of Style, University of Chicago Press. Edition number(s): _____
- Collins Good Writing Guide, King, G, HarperCollins Publishers. Edition number(s): _____
- Effective Writing: Plain English at Work, Murphy, EM, Lacuna. Edition number(s): _____
- The Elements of Style, Strunk, W and White, E, Allyn & Bacon. Edition number(s): _____
- MLA Handbook for Writers of Research Papers, Modern Language Association. Edition number(s): _____
- The New Oxford Style Manual, Ritter, RM, Oxford University Press. Edition number(s): _____
- Publication Manual of the American Psychological Association, American Psychological Association. Edition number(s): _____
- Working words, Murphy, EM, Canberra Society of Editors. Edition number(s): _____
- Other

Display the following question:

B4.1 Please list any other style manuals that you use in the course of your work (including information about editions) in the text boxes:

1 _____
2 _____
3 _____
4 _____
5 _____

If From the options provided below, please select the style manuals you use most often in the course of your editorial work. "Other" is selected

Carry Forward Entered Choices from B4 and Entered Text from B4.1

B4.2 Of the style manuals you selected or listed, which one do you use most often in the course of your work?

- AMA Manual of Style: A Guide for Authors and Editors, Oxford University Press. Edition number(s): _____

- The Australian Government Style Manual: For Authors, Editors and Printers, John Wiley. Edition number(s): _____
- Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors and Proofreaders, Cambridge University Press. Edition number(s): _____
- Chicago Manual of Style, University of Chicago Press. Edition number(s):

- Collins Good Writing Guide, King, G, HarperCollins Publishers. Edition number(s):

- Effective Writing: Plain English at Work, Murphy, EM, Lacuna. Edition number(s):

- The Elements of Style, Strunk, W and White, E, Allyn & Bacon. Edition number(s):

- MLA Handbook for Writers of Research Papers, Modern Language Association. Edition number(s): _____
- The New Oxford Style Manual, Ritter, RM, Oxford University Press. Edition number(s):

- Publication Manual of the American Psychological Association, American Psychological Association. Edition number(s): _____
- Working words, Murphy, EM, Canberra Society of Editors. Edition number(s):

- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

- B5 Do you make use of any other sources in the course of your editorial work, not covered by the categories mentioned in the questions in this section (online dictionaries, print dictionaries, usage guides, style manuals), such as organisations, people, house style sheets, corpora?
- Yes (1)
 - No (2)

Display the following question:

B5.1 Please provide the names and/or bibliographic details of these sources in order, from the one you use most often (1) to the one you use least often (5), in the text boxes below.

1 _____

2 _____

3 _____

4 _____

5 _____

If Do you make use of any other sources in the course of your editorial work, not covered by the categories in questions B3, B5, B7, and B9 (such as organisations, people, house style sheets)? “Yes” is selected

B6 In the course of your editorial work, do you encounter particular linguistic features for which you cannot find sufficient advice in the language sources you use?

- Yes (1)
- No (2)

Display the following question:

B7 Please use the text box to provide more information or examples.

If *In the course of your editorial work, do you frequently encounter particular linguistic features for which you cannot find sufficient advice in the language sources you use? "Yes" is selected*

Section 3: Strategies for solving editing problems

Editors rely on many **different strategies to solve the problems** they encounter in texts. Some of these problem-solving strategies (and examples of their use) are listed in the questions that follow.

Please read each strategy and **provide up to three examples** of editing problems you have solved using each strategy in the text boxes provided. If you do not make use of a particular strategy, please type n/a in the first text box.

B8.1 Strategy: Use a print dictionary

Example: *While editing an academic article you encounter the word "chiasmus". To ensure the correct spelling of the word, and to verify its meaning, you consult a print dictionary.*

- 1 _____
- 2 _____
- 3 _____

B8.2 Strategy: Use an online dictionary

Example: *You are editing a novel and come across the word "Pakeha". Because you are unsure what this word means, and if it even exists, you consult an online dictionary to verify its existence, meaning and spelling.*

- 1 _____
- 2 _____
- 3 _____

B8.3 Strategy: Use a style manual

Example: *A company newsletter includes the following sentence: "The data was gathered by the marketing department." You are unsure if "data" should be treated as a singular or plural noun, and consult the company's preferred style manual to verify which treatment is preferred.*

- 1 _____
- 2 _____
- 3 _____

B8.4 Strategy: Use a usage guide

Example: *While editing a newspaper article, you come across the following sentence and wonder about the use of "disbarred" versus "debarred": "The barrister was found guilty of professional misconduct by a disciplinary council and was subsequently disbarred". To solve this, you consult a usage guide to determine which term to use.*

- 1 _____
- 2 _____
- 3 _____

B8.5 Strategy: Rely on your intuitive knowledge as a proficient speaker

Example: *While editing a text, you spot an error and are able to apply a solution without having to or previously having had to consult a source. For example, you may intuitively correct "He sung that song beautifully" to "He sang that song beautifully".*

- 1 _____
- 2 _____
- 3 _____

B8.6 Strategy: Rely on your experience, in which you have encountered the same problem many times and developed a solution

Example: *In your work as an editor, you may have frequently come across issues of gender-specific language. For example, you may have felt previously that "chairman" is an unnecessarily gendered compound, and decided to make use of the gender inclusive "chairperson" or an alternative gender-free term appropriate to context, such as "convener" or "coordinator". Since making this decision, you have applied this solution whenever you have encountered this word.*

- 1 _____
- 2 _____
- 3 _____

B8.7 Strategy: Consult a corpus

Example: *While editing a text, you come across the modal verb "must" in the following sentence, "Researchers must ensure that their applications are submitted by the deadline", and wonder whether it is necessary to replace "must" with "should". In order to determine which is used more often in this type of construction, you consult a corpus, such as the British National Corpus.*

- 1 _____
- 2 _____
- 3 _____

B8.8 Strategy: Ask editors or language experts in a web forum or electronic chat group

Example: *You are unsure if "neither" can be used in the following sentence: "The results indicate that neither a), b) nor c) is associated with d)." Although you know that "either" can be used with more than two elements, you are unsure if this can be extrapolated to "neither". To solve this*

problem, you submit a query in a web forum or to a chat group hoping to receive advice from fellow editors or language experts.

- 1 _____
- 2 _____
- 3 _____

B8.9 Strategy: Ask another person for advice

Example: You are editing a text and encounter the following, "His brother was owing him fifty dollars". You are unsure if this is correct, and so ask another person for advice.

- 1 _____
- 2 _____
- 3 _____

B8.10 Strategy: Use a search engine to do a frequency count

Example: You are editing a news report and encounter "sit for an exam". Unsure whether this should be changed to "sit an exam", you make use of a search engine, such as Google, to do a frequency count of the two in order to determine which is used most often.

- 1 _____
- 2 _____
- 3 _____

B8.11 Please list any other problem-solving strategies that you use in the course of your work in the text boxes below. Please include up to three examples of decisions you have solved using the strategy(-ies).

- 1 _____
- 2 _____
- 3 _____

B9 If you would like to comment on any of the questions in the survey or if you would like to include additional information on features you encounter very frequently, please use the text box to do so.

*We thank you for your time spent taking this survey.
Your response has been recorded.*

Appendix B: Final questionnaire

Survey: Editors' norm-providing sources

Chief Investigator's / Supervisor's Name & Title: Dr Haidee Kruger and Prof. Bertus van Rooy

Participant Information and Consent Form

Name of Project: The role of editorial intervention in ongoing language variation and change in South African and Australian English

You are invited to participate in a study of editorial intervention in published English texts. The purpose of the study is to determine what role editors play in the processes of language variation and change, and how their editorial decisions are guided by norm-providing sources. We would like to examine this in two varieties of English: Australian English and South African English.

The study is being conducted by Melanie Ann Law of Macquarie University, Sydney, Australia (email: melanie.law@students.mq.edu.au) and North-West University, Vaal Triangle Campus, South Africa (email: melanie.law@nwu.ac.za). The study will be conducted to meet the requirements of the degrees Doctor of Philosophy in Linguistics at Macquarie University, Sydney, Australia and Doctor of Philosophy in Linguistics and Literary Studies at North-West University, Vaal Triangle Campus, South Africa, under a Joint Doctoral Supervision Agreement between Macquarie University and North-West University, and under the supervision of:

- Dr Haidee Kruger (haidee.kruger@mq.edu.au) of the Department of Linguistics, Macquarie University
- Prof. Bertus van Rooy (bertus.vanrooy@nwu.ac.za) of the School of Languages, UPSET Research Focus Area, North-West University

If you decide to participate, you will be asked to complete the online questionnaire that follows this information letter. This questionnaire has three sections:

- Background information
- Sources used during editing
- Strategies for solving editing problems

The questionnaire should take approximately 30 minutes to complete.

There is no risk or discomfort involved in the completion of the questionnaire.

Any information or personal details gathered in the course of the study are confidential, except as required by law. No individual will be identified in any publication of the results. Only the investigators named above will have access to the data. A summary of the results of the data can be made available to you on request by contacting Melanie Law on one of the email addresses above.

Participation in this study is entirely voluntary: you are not obliged to participate and if you decide to participate, you are free to withdraw at any time without having to give a reason and without consequence.

The ethical aspects of this study have been approved by the Macquarie University Human Research Ethics Committee (ethics clearance reference number: 5201600843D) and the North-West University Research Ethics Committee for Language Matters (ethics clearance reference number: NWU-00558-16-A8). If you have any complaints or reservations about any ethical aspect of your participation in this research, you may contact the Macquarie University Human Research Ethics Committee through the Director, Research Ethics & Integrity (telephone +61 298 507 854; email ethics@mq.edu.au), or the North-West University Research Ethics Committee for Language Matters through the Chairperson, Prof. Bertus van Rooy, (telephone +27 16 910 3482; email bertus.vanrooy@nwu.ac.za). Any complaint you make will be treated in confidence and investigated, and you will be informed of the outcome.

Please indicate whether you give your consent to participate in this study:

- I consent to participate in this study (1)
 - I do not consent to participate in this study (2)
-

Instructions for completing the questionnaire

4. Please answer all questions.
 5. Please read all questions carefully.
 6. Complete the questionnaire using the selection options or text boxes provided.
-

Section 1: Background information

This section comprises several questions related to your demographic information (including your age, language profile, qualification, working languages, areas of experience and years' editing experience).

Please complete all questions in this section.

A1 Please select your age from the options below.

- 18-29 years old
- 30-39 years old
- 40-49 years old
- 50-59 years old
- 60 years and older

A2 What language do you regard as your first or home language?

- Afrikaans
- An indigenous Australian language
- An indigenous South African language
- English
- Other

Display the following question:

A2.1 Please use the text box below to indicate your first or home language.

If *What language do you regard as your first or home language?* “Other” is selected

Or *What language do you regard as your first or home language?* “An indigenous Australian Language” is selected

Or *What language do you regard as your first or home language?* “One of the nine indigenous South African languages” is selected

A3 To which gender identity do you most identify?

- Prefer not to answer
- Female
- Male
- Other (please specify) _____

A4 In the text boxes below, please list all the languages you know, up to a maximum of five. List the languages in order, from your strongest (1) to your weakest (5) language.

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

A5 In the text boxes below, please list all the languages in which you do editing work, up to a maximum of five. List the languages in order, from the language you edit in most often (1) to the language you edit in least often (5).

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

A6 From the options available, please select the country in which you currently reside.

- Australia
- South Africa
- Other

Display the following question:

A7 Please use the text box to indicate the country in which you currently reside.

If *From the options available, please select the country you currently reside in.* “Other” is selected

A8 Have you always lived in the country in which you currently reside?

- Yes
- No

Display the following question:

A8.1 How many years have you lived in the country in which you currently reside?

- 0-10 years (1)
- 11-20 years (2)
- 21-30 years (3)
- 31-40 years (4)
- 41-50 years (5)
- 51 years or more (6)

If *Have you always lived in the country in which you currently reside?* “No” is selected

Display the following question:

A9 Please use the first text box to indicate the country in which you were born and the number of years in which you lived in this country. If you have lived in countries other than the one(s) in which you were born and currently reside, please also list these countries and the number of years spent in each, in the second text box below.

I was born in: _____

I have lived in: _____

If *Have you always lived in the country in which you currently reside?* “No” is selected

A10 How many years’ editing experience do you have?

- 1-10 years (1)
- 11-20 years (2)
- 21-30 years (3)
- 31-40 years (4)
- 41 years and more (5)

A11 Do you work as a full-time editor or part-time editor?

- Full-time (1)
- Part-time (2)

Display the following question:

A11.1 In addition to editing, please list the other kind(s) of work you do in the text box below.

If *Do you work as a full-time editor or part-time editor?* “Part-time” is selected

A12 Do you work as a freelance or self-employed editor?

- Yes (1)
- No (2)

Display the following question:

A12.1 Please provide a short description of the organisation or company where you are employed. (For example, a community newspaper.)

If *Do you work as a freelance or self-employed editor?* “No” is selected

A13 What is your highest level of qualification?

- Year 12 / Matric
- First degree or equivalent
- Honours degree
- Masters degree
- Doctoral degree
- Other

Display the following question:

A13.1 Please use the text box to indicate your highest level of qualification.

If *What is your highest qualification?* “Other” is selected

Display the following question:

A14 In what discipline is this qualification? (For example, Accounting, Languages, Engineering, etc.)

If *What is your highest qualification?* “First degree or equivalent” is selected

Or *What is your highest qualification?* “Honours degree” is selected

Or *What is your highest qualification?* “Masters degree” is selected

Or *What is your highest qualification?* “Doctoral” is selected

Or *What is your highest qualification?* “Other” is selected

A15 From the options below, please select all text types in which you have editing experience.

- Academic articles appearing in scholarly journals
- Academic books
- Administrative texts (including minutes, institutional communication, etc.)
- Contracts and legal documents
- Dissertations and theses
- Instruction manuals
- Magazine articles
- Medical and pharmaceutical texts
- Newsletters
- Newspaper reports: online
- Newspaper reports: print
- Novels, novellas and short stories for leisure reading
- Plays
- Poetry
- Popular non-fiction books
- Schoolbooks
- Study guides

- Teacher's guides
- Website material
- Other

Display the following question:

A15.1 Please indicate any other text types in which you have editing experience in the text boxes below.

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____
- 6 _____
- 7 _____
- 8 _____
- 9 _____
- 10 _____

If *From the options below, please select all text types in which you have editing experience. "Other" is selected*

Carry Forward Entered Choices from A15 and Entered Text from A15.1

A16 In which text type do you have the most editing experience?

- Academic articles appearing in scholarly journals
- Academic books
- Administrative texts (including minutes, institutional communication, etc.)
- Contracts and legal documents
- Dissertations and theses
- Instruction manuals
- Magazine articles
- Medical and pharmaceutical texts
- Newsletters
- Newspaper reports: online
- Newspaper reports: print
- Novels, novellas and short stories for leisure reading
- Plays
- Poetry
- Popular non-fiction books
- Schoolbooks
- Study guides
- Teacher's guides
- Website material
- Other
 - 1 _____
 - 2 _____
 - 3 _____

- 4 _____
- 5 _____
- 6 _____
- 7 _____
- 8 _____
- 9 _____
- 10 _____

Carry Forward Entered Choices from A15 and Entered Text from A15.1

A17 In which text type is your most recent or current editing experience?

- Academic articles appearing in scholarly journals
- Academic books
- Administrative texts (including minutes, institutional communication, etc.)
- Contracts and legal documents
- Dissertations and theses
- Instruction manuals
- Magazine articles
- Medical and pharmaceutical texts
- Newsletters
- Newspaper report: online
- Newspaper reports: print
- Novels, novellas and short stories for leisure reading
- Plays
- Poetry
- Popular non-fiction books
- Schoolbooks
- Study guides
- Teacher's guides
- Website material
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____
 - 6 _____
 - 7 _____
 - 8 _____
 - 9 _____
 - 10 _____

For the next question, you will be asked to select the option that most accurately describes how you view the function or aim of your editorial role. The description for each function is as follows:

Normative: Normative work aims to bring a text into conformance with predefined rules of usage as set out in dictionaries, house styles, style manuals, usage guides and other prescriptive sources.

Communicative: Communicative work aims to ensure that the text communicates optimally with its intended reader, and may make use of language that may not (yet) be set out in dictionaries, house styles, style manuals usage guides and other prescriptive sources.

A18 I view my editorial role as fulfilling

- an exclusively normative function. (1)
- an exclusively communicative function. (2)
- an equally normative and communicative function. (3)
- a primarily normative function, with communicative elements. (4)
- a primarily communicative function, with normative elements. (5)

A19 Please use the text box below to explain or motivate your answer.

Section 2: Sources used during editing

Editors rely on many **different sources** when editing. The questions that follow seek to gather information on the kinds of sources you use in the course of your editorial work.

If you do not make use of a particular source, please select the "I do not make use of (type of source) in the course of my editorial work" option.

For each source type you may select up to five sources.

B1 From the options provided below, please select up to five online dictionaries you use most often in the course of your editorial work.

- I do not use online dictionaries in the course of my editorial work.
- Australian National Dictionary
<http://australiannationaldictionary.com.au/index.php>
- Cambridge Free English Dictionary and Thesaurus
<http://dictionary.cambridge.org/dictionary/english/>
- Collins English Dictionary
<http://www.collinsdictionary.com/dictionary/english>
- Dictionary of South African English on Historical Principles
<http://dsae.co.za/>
- The Free Dictionary by Farlex
www.thefreedictionary.com
- Longman Dictionary of Contemporary English
<http://www.ldoceonline.com/>
- Macmillan Dictionary

- <http://www.macmillandictionary.com/>
- The Macquarie Dictionary
<https://www.macquariedictionary.com.au/>
- Merriam-Webster Dictionary and Thesaurus
<http://www.merriam-webster.com/>
- Oxford English Dictionary (OED)
<http://www.oed.com/>
- Oxford Dictionaries
<https://en.oxforddictionaries.com/definition/>
- Oxford Reference Online
www.oxfordreference.com
- Wiktionary
https://en.wiktionary.org/wiki/Wiktionary:Main_Page
- Other

Display the following question:

B1.1 Please list any other online dictionaries that you use in the course of your work:

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

If From the options provided below, please select the online dictionaries you use most often in the course of your editorial work. "Other" is selected

Carry Forward Entered Choices from B1. and Entered Text B1.1

B1.2 Of the online dictionaries you selected or listed, which one do you use most often in the course of your work?

- Australian National Dictionary
<http://australiannationaldictionary.com.au/index.php>
- Cambridge Free English Dictionary and Thesaurus
<http://dictionary.cambridge.org/dictionary/english/>
- Collins English Dictionary
<http://www.collinsdictionary.com/dictionary/english>
- Dictionary of South African English on Historical Principles
<http://dsae.co.za/>
- The Free Dictionary by Farlex
www.thefreedictionary.com
- Longman Dictionary of Contemporary English
<http://www.ldoceonline.com/>
- Macmillan Dictionary
<http://www.macmillandictionary.com/>
- The Macquarie Dictionary
<https://www.macquariedictionary.com.au/>
- Merriam-Webster Dictionary and Thesaurus

- <http://www.merriam-webster.com/>
- Oxford English Dictionary (OED)
<http://www.oed.com/>
- Oxford Dictionaries
<https://en.oxforddictionaries.com/definition/>
- Oxford Reference Online
www.oxfordreference.com
- Wiktionary
https://en.wiktionary.org/wiki/Wiktionary:Main_Page
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

B2 From the options below, please select up to five print dictionaries you use most often in the course of your editorial work. Please indicate the edition number(s) of your selections in the text boxes.

- I do not use print dictionaries in the course of my editorial work.
- Australian National Dictionary, Oxford University Press. Edition number(s):

- Australian Oxford Dictionary, Oxford University Press. Edition number(s):

- Cambridge Advanced Learner's Dictionary, Cambridge University Press. Edition number(s):

- Collins English Dictionary – Complete and Unabridged, HarperCollins Publishers. Edition number(s): _____
- Collins Dictionary for Writers and Editors, HarperCollins Publishers. Edition number(s):

- Collins South African English Dictionary, HarperCollins Publishers. Edition number(s):

- English Dictionary for South Africa, Pharos. Edition number(s): _____
- Longman Dictionary of Contemporary English, Longman/Pearson. Edition number(s):

- The Macquarie Dictionary, Macquarie Library Pty Ltd, Macquarie University. Edition number(s):

- Merriam-Webster's Collegiate Dictionary, Merriam Webster Inc. Edition number(s):

- New Oxford Dictionary for Writers and Editors, Oxford University Press. Edition number(s):

- Oxford South African Concise Dictionary, Oxford University Press. Edition number(s):

- Other

Display the following question:

B2.1 Please list any other dictionaries that you use in the course of your work (including information about editions) in the text boxes:

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

If From the options below, please select the print dictionaries you use most often in the course of your editorial work. "Other" is selected

Carry Forward Entered Choices from B2 and Entered Text from B2.1

B2.2 Of the print dictionaries you selected or listed, which one do you use most often in the course of your work?

- Australian National Dictionary, Oxford University Press. Edition number(s):

- Australian Oxford Dictionary, Oxford University Press. Edition number(s):

- Cambridge Advanced Learner's Dictionary, Cambridge University Press. Edition number(s):

- Collins English Dictionary – Complete and Unabridged, HarperCollins Publishers. Edition number(s): _____
- Collins Dictionary for Writers and Editors, HarperCollins Publishers. Edition number(s):

- Collins South African English Dictionary, HarperCollins Publishers. Edition number(s):

- English Dictionary for South Africa, Pharos. Edition number(s): _____
- Longman Dictionary of Contemporary English, Longman/Pearson. Edition number(s):

- The Macquarie Dictionary, Macquarie Library Pty Ltd, Macquarie University. Edition number(s):

- Merriam-Webster's Collegiate Dictionary, Merriam Webster Inc. Edition number(s):

- New Oxford Dictionary for Writers and Editors, Oxford University Press. Edition number(s):

- Oxford South African Concise Dictionary, Oxford University Press. Edition number(s):

- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

B3 From the options provided below, please select up to five usage guides you use most often in the course of your editorial work. Please indicate the edition number(s) of your selections in the text boxes.

- I do not use usage guides in the course of my editorial work.
- The Cambridge Guide to Australian English Usage, Peters, P, Cambridge University Press. Edition number(s): _____
- The Cambridge Guide to English Usage, (2004), Peters, P, Cambridge University Press.
- Collins Gem English Grammar, Hardie, RG, HarperCollins Publishers. Edition number(s):

- English, our English: Common Errors in South African English and How to Solve Them, (2009), Linnegar, J, Pharos.
- Fowler's Modern English Usage, Oxford University Press. Edition number(s):

- Garner's Modern English Usage, Oxford University Press. Edition number(s):

- The Little Green Grammar Book, (2008), Tredinnick, M, UNSW Press.
- Modern Australian Usage, Oxford University Press/Allen & Unwin. Edition number(s):

- The Oxford A–Z of Grammar & Punctuation, Seely, J, Oxford University Press. Edition number(s):

- Right Words: A Guide to English Usage in Australia, Murray-Smith, S, Viking/Penguin. Edition number(s): _____
- Usage and Abusage: A Guide to Good English, Partridge, E, Penguin Australia. Edition number(s):

- Other

Display the following question:

B3.1 Please list any other usage guides that you use in the course of your work (including information about editions) in the text boxes:

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

If From the options provided below, please select the usage guides you use most often in the course of your editorial work. "Other" is selected

Carry Forward Entered Choices from B3 and Entered Text from B3.1

B3.2 Of the usage guides you selected or listed, which one do you use most often in the course of your work?

- I do not use usage guides in the course of my editorial work.
- The Cambridge Guide to Australian English Usage, Peters, P, Cambridge University Press. Edition number(s): _____
- The Cambridge Guide to English Usage, (2004), Peters, P, Cambridge University Press.
- Collins Gem English Grammar, Hardie, RG, HarperCollins Publishers. Edition number(s): _____
- English, our English: Common Errors in South African English and How to Solve Them, (2009), Linnegar, J, Pharos.
- Fowler's Modern English Usage, Oxford University Press. Edition number(s): _____
- Garner's Modern English Usage, Oxford University Press. Edition number(s): _____
- The Little Green Grammar Book, (2008), Tredinnick, M, UNSW Press.
- Modern Australian Usage, Oxford University Press/Allen & Unwin. Edition number(s): _____
- The Oxford A–Z of Grammar & Punctuation, Seely, J, Oxford University Press. Edition number(s): _____
- Right Words: A Guide to English Usage in Australia, Murray-Smith, S, Viking/Penguin. Edition number(s): _____
- Usage and Abusage: A Guide to Good English, Partridge, E, Penguin Australia. Edition number(s): _____
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

B4 From the options provided below, please select up to five style manuals you use most often in the course of your editorial work. Please indicate the edition number(s) of your selections in the text boxes.

- I do not use style manuals in the course of my editorial work.
- AMA Manual of Style: A Guide for Authors and Editors, Oxford University Press. Edition number(s): _____
- The Australian Government Style Manual: For Authors, Editors and Printers, John Wiley. Edition number(s): _____
- Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors and Proofreaders, Cambridge University Press. Edition number(s): _____
- Chicago Manual of Style, University of Chicago Press. Edition number(s): _____

- Collins Good Writing Guide, King, G, HarperCollins Publishers. Edition number(s):

- Effective Writing: Plain English at Work, Murphy, EM, Lacuna. Edition number(s):

- The Elements of Style, Strunk, W and White, E, Allyn & Bacon. Edition number(s):

- MLA Handbook for Writers of Research Papers, Modern Language Association. Edition number(s): _____
- The New Oxford Style Manual, Ritter, RM, Oxford University Press. Edition number(s):

- Publication Manual of the American Psychological Association, American Psychological Association. Edition number(s): _____
- Working words, Murphy, EM, Canberra Society of Editors. Edition number(s):

- Other

Display the following question:

B4.1 Please list any other style manuals that you use in the course of your work (including information about editions) in the text boxes:

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

If From the options provided below, please select the style manuals you use most often in the course of your editorial work. "Other" is selected

Carry Forward Entered Choices from B4 and Entered Text from B4.1

B4.2 Of the style manuals you selected or listed, which one do you use most often in the course of your work?

- AMA Manual of Style: A Guide for Authors and Editors, Oxford University Press. Edition number(s): _____
- The Australian Government Style Manual: For Authors, Editors and Printers, John Wiley. Edition number(s): _____
- Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors and Proofreaders, Cambridge University Press. Edition number(s): _____
- Chicago Manual of Style, University of Chicago Press. Edition number(s):

- Collins Good Writing Guide, King, G, HarperCollins Publishers. Edition number(s):

- Effective Writing: Plain English at Work, Murphy, EM, Lacuna. Edition number(s):

- The Elements of Style, Strunk, W and White, E, Allyn & Bacon. Edition number(s):

- MLA Handbook for Writers of Research Papers, Modern Language Association. Edition number(s): _____
- The New Oxford Style Manual, Ritter, RM, Oxford University Press. Edition number(s): _____
- Publication Manual of the American Psychological Association, American Psychological Association. Edition number(s): _____
- Working words, Murphy, EM, Canberra Society of Editors. Edition number(s): _____
- Other
 - 1 _____
 - 2 _____
 - 3 _____
 - 4 _____
 - 5 _____

- B5 Do you make use of any other sources in the course of your editorial work, not covered by the categories mentioned in the questions in this section (online dictionaries, print dictionaries, usage guides, style manuals), such as organisations, people, house style sheets, corpora?
- Yes (1)
 - No (2)

Display the following question:

B5.1 Please provide the names and/or bibliographic details of these sources in order, from the one you use most often (1) to the one you use least often (5), in the text boxes below.

- 1 _____
- 2 _____
- 3 _____
- 4 _____
- 5 _____

If *Do you make use of any other sources in the course of your editorial work, not covered by the categories in questions B3, B5, B7, and B9 (such as organisations, people, house style sheets)? “Yes” is selected*

- B6 In the course of your editorial work, do you encounter particular linguistic features for which you cannot find sufficient advice in the language sources you use?
- Yes (1)
 - No (2)

Display the following question:

B7 Please use the text box to provide more information or examples.

If *In the course of your editorial work, do you frequently encounter particular linguistic features for which you cannot find sufficient advice in the language sources you use? “Yes” is selected*

B7 If you would like to comment on any of the questions in the survey or if you would like to include additional information on features you encounter very frequently, please use the text box to do so.

*We thank you for your time spent taking this survey.
Your response has been recorded.*

Appendix C: Source type categorisation

SOURCE TYPE	TYPE	ORIENTATION	No. OF AUS RES.	No. OF SA RES.
Dictionary				
New Oxford Dictionary for Writers and Editors, 1981, 2000, 2005, 2014, nd, Oxford: Oxford University Press.	Pr. Dict.	BrE	6	16
Concise Oxford English Dictionary, 1964, , 1976, 1999, 2002, 2008, nd, Oxford: Oxford University Press.	Pr. Dict.	BrE	2	7
Pharos Afrikaans–English English–Afrikaans Dictionary, 2005, 2010, nd, Cape Town: Pharos.	Bi. Pr. Dic	SAE-SAAfr	0	6
Collins English Dictionary: Complete and Unabridged, 1991, 2003, 2005, 2010, 2014, nd, Glasgow: HarperCollins.	Pr. Dict.	BrE	3	5
Oxford Advanced Learner's Dictionary, 1995, 2005, 2010, Oxford: Oxford University Press.	Pr. Dict.	BrE	0	5
Concise South African Oxford Dictionary, 2010, nd, Cape Town: Oxford University Press.	Pr. Dict.	SAE	0	5
South African Concise Oxford Dictionary, 2002, Cape Town: Oxford University Press.	Pr. Dict.	SAE	0	5
Subject Dictionaries: Publisher unspecified.	Pr. Dict.	Ind.	3	4
Collins Dictionary for Writers and Editors: Essential Reference for Writers, Editors and Proofreaders, 2007, Glasgow: HarperCollins.	Pr. Dict.	BrE	0	4
Pharos English Dictionary for South Africa, 2011, Cape Town: Pharos.	Pr. Dict.	SAE	0	4
Oxford Dictionary of English, 1998, 2003, 2010, Oxford: Oxford University Press.	Pr. Dict.	BrE	1	3
Longman Dictionary of Contemporary English, 2014, Harlow: Pearson Education Ltd.	Pr. Dict.	AmE	0	2
Handwoordeboek van die Afrikaanse Taal, 2009, 2015, Cape Town: Pearson.	Pr. Dict.	SAAfr	0	2
A Dictionary of South African English on Historical Principles, 2000, Cape Town: Oxford University Press	Pr. Dict.	SAE	0	2
Pharos Dictionary for the Health Sciences, 2011, Cape Town: Pharos.	Pr. Dict.	SAE	0	2
Merriam-Webster's Collegiate Dictionary, 2003, nd, Springfield: Merriam-Webster.	Pr. Dict.	AmE	3	1
New Oxford Spelling Dictionary, 2005, 2014, Oxford: Oxford University Press.	Pr. Dict.	BrE	1	1
Webster's Third New International Dictionary, 1993, Springfield: Merriam-Webster.	Pr. Dict.	AmE	0	1
Collins COBUILD Advanced Learner's English Dictionary, 2008, London: HarperCollins.	Pr. Dict.	AmE/BrE	0	1
Collins Shorter English Dictionary, nd, Glasgow: HarperCollins.	Pr. Dict.	AmE/BrE	0	1
Encarta World English Dictionary, 1999, London: Bloomsbury.	Pr. Dict.	AmE/BrE	0	1
Oxford Dictionary of Foreign Words and Phrases, 2010, Oxford: Oxford University Press.	Pr. Dict.	AmE/BrE	0	1
Blackwell's Nursing Dictionary, 2004, Oxford: Wiley Blackwell.	Pr. Dict.	BrE	0	1
Dictionary of Agriculture, 2006, London: A & C Black.	Pr. Dict.	BrE	0	1
Dictionary of Economics Terms, 1981, London: Butterworth & Co Publishers Ltd.	Pr. Dict.	BrE	0	1
Macmillan English Dictionary for Advanced Learners, 2007, Oxford: Macmillan Education.	Pr. Dict.	BrE	0	1
Oxford Compact Dictionary and Thesaurus, nd, Oxford: Oxford University Press.	Pr. Dict.	BrE	0	1
Oxford Dictionary of Economics, 2009, Oxford: Oxford University Press.	Pr. Dict.	BrE	0	1
Subject Dictionaries: Oxford University Press.	Pr. Dict.	BrE	0	1
The Penguin Dictionary for Writers and Editors, nd, London: Penguin.	Pr. Dict.	BrE	0	1

Appendix C: Source type categorisation

Concise Dictionary of English Etymology, 1993, Hertfordshire, Wordsworth Editions Ltd.	Pr. Dict.	Ind.	0	1
Pharos Woordeboek van Regs- en Handelsterme, 1992, Cape Town: Pharos.	Pr. Dict.	SA Afr	0	1
Verklarende Afrikaanse Woordeboek, 2010, Cape Town: Pharos.	Pr. Dict.	SA Afr	0	1
A Dictionary of South African English, 1992, Cape Town: Oxford University Press.	Pr. Dict.	SAE	0	1
Collins South African English Dictionary, 2016, HarperCollins.	Pr. Dict.	SAE	0	1
Collins GEM Latin–English English–Latin Dictionary, 1962, Glasgow: Collins.	Bi. Pr. Dic	Eng-For	0	1
Kuberwoordeboek / Cyber Dictionary, 2006, Pretoria: Protea Boekhuis.	Bi. Pr. Dic	SAE-SAAfr	0	1
Pharos Bilingual Phrase Dictionary: English-Afrikaans, 1997, Cape Town: Pharos.	Bi. Pr. Dic	SAE-SAAfr	0	1
Pharos Business Dictionary/ Sakewoordeboek, 2009, Cape Town: Pharos.	Bi. Pr. Dic	SAE-SAAfr	0	1
Pharos Ekonomiese en Bedryfswoordeboek / Dictionary of Economics and Business, 2004, Cape Town: Pharos.	Bi. Pr. Dic	SAE-SAAfr	0	1
Pharos Kern-woordeboek/Concise Dictionary, 2007, Cape Town: Pharos.	Bi. Pr. Dic	SAE-SAAfr	0	1
Pharos New Words / Nuwe Woorde, 1999, Cape Town: Pharos.	Bi. Pr. Dic	SAE-SAAfr	0	1
Reader's Digest Afrikaans–Engelse English–Afrikaans Dictionary, 1984, Reader's Digest.	Bi. Pr. Dic	SAE-SAAfr	0	1
The Eskom Dictionary for the Electricity Supply Industry / Eskom Woordeboek vir die Elektrisiteitsvoorsieningsbedryf, nd, Eskom.	Bi. Pr. Dic	SAE-SAAfr	0	1
The Macquarie Dictionary, 1981, 1991, 1997, 2001, 2005, 2009, 2011, 2013, 2017, nd, Sydney: Macquarie Dictionary Publishers Pty. Ltd.	Pr. Dict.	AusE	42	0
Australian Oxford Dictionary, 2004, , 2011, nd, Melbourne: Oxford University Press.	Pr. Dict.	AusE	6	0
Australian National Dictionary, 1988, 2016, Melbourne: Oxford University Press.	Pr. Dict.	AusE	3	0
Shorter Oxford English Dictionary, 1972, 1993, 2007, New York: Oxford University Press.	Pr. Dict.	BrE	3	0
The American Heritage Dictionary of the English Language, 2000, Boston: Houghton Mifflin.	Pr. Dict.	AmE	1	0
Bryson's Dictionary of Troublesome Words, 2002, Portland: Broadway Books.	Pr. Dict.	AmE/BrE	1	0
Authors' & Printers' Dictionary, 1956, London: Oxford University Press.	Pr. Dict.	AusE	1	0
The Australian Modern Oxford Dictionary, nd, Melbourne: Oxford University Press.	Pr. Dict.	AusE	1	0
The Oxford Australian Writer's Dictionary, 1997, Melbourne: Oxford University Press.	Pr. Dict.	AusE	1	0
Bryson's Dictionary for Writers and Editors, nd, Portland: Broadway Books.	Pr. Dict.	BrE	1	0
Cambridge Advanced Learner's Dictionary, nd, Cambridge: Cambridge University Press.	Pr. Dict.	BrE	1	0
Faber Medical Dictionary, nd, London: Faber & Faber.	Pr. Dict.	BrE	1	0
The Oxford Dictionary of Quotations, 1992, Oxford: Oxford University Press.	Pr. Dict.	BrE	1	0
Larousse Dictionary of Science and Technology, 1995, Paris: Larousse.	Pr. Dict.	Ind.	1	0
The New Fontana Dictionary of Modern Thought, nd, London: HarperCollins.	Pr. Dict.	Ind.	1	0
Kamus Indonesia Inggris: An Indonesian-English Dictionary, 1992, Jakarta: Gramedia.	Bi. Pr. Dic	Eng-For	1	0
Foreign Language Dictionaries: Publisher unspecified	For. Pr. D	Ind.	1	0
Online dictionaries				
The Macquarie Dictionary, online, https://www.macquariedictionary.com.au/	On. Dict.	AusE	62	35
Merriam-Webster Dictionary and Thesaurus, online, http://www.merriam-webster.com/	On. Dict.	AmE	27	28

Oxford English Dictionary (OED), online, http://www.oed.com/	On. Dict.	BrE	21	22
Oxford Dictionaries, online, https://en.oxforddictionaries.com/definition/	On. Dict.	BrE	19	20
Collins English Dictionary, online, http://www.collinsdictionary.com/dictionary/english	On. Dict.	BrE	10	12
Cambridge Free English Dictionary and Thesaurus, online, http://dictionary.cambridge.org/dictionary/english/	On. Dict.	BrE	7	9
Wiktionary, online, https://en.wiktionary.org/wiki/Wiktionary:Main_Page	On. Dict.	Ind.	7	8
The Free Dictionary by Farlex, online, www.thefreedictionary.com	On. Dict.	AmE	5	7
Australian National Dictionary, online, http://australiannationaldictionary.com.au/index.php	On. Dict.	AusE	5	4
Macmillan Dictionary, online, http://www.macmillandictionary.com/	On. Dict.	BrE	2	3
Urban Dictionary, online, https://www.urbandictionary.com/	On. Dict.	Ind.	2	2
MedlinePlus, online, https://medlineplus.gov/encyclopedia.html	On. Dict.	AmE	1	1
Merriam-Webster's Learners Dictionary, online, http://www.learnersdictionary.com	On. Dict.	AmE	1	1
OneLook Dictionary Search, online, www.onelook.com	On. Dict.	AmE/BrE	1	1
New Oxford American Dictionary, OS-based dictionary.	OS Dict.	AmE	1	1
Oxford English Dictionary, OS-based dictionary	OS Dict.	BrE	1	0
JEdict, OS-based dictionary.	OS Dict.	Eng-For	1	0
Longman Dictionary of Contemporary English, online, http://www.ldoceonline.com/	On. Dict.	AmE	0	0
Dictionary of South African English on Historical Principles, online, http://dsae.co.za/	On. Dict.	SAE	0	0
WordReference.com, online, http://www.wordreference.com/	Bi. On. Dict.	Eng-For	0	0
Linguee bilingual dictionary, online, https://www.linguee.com	Bi. On. Dict.	Eng-For	0	0
InterActiveCorp (Random House), http://www.dictionary.com	On. Dict.	AmE	0	0
The Online Etymology Dictionary, online, https://www.etymonline.com/	On. Dict.	AmE	0	0
Handwoordeboek van die Afrikaanse Taal aanlyn, online, https://longman-hat.co.za/	On. Dict.	SAAfr	0	0
Glossary				
A Glossary of Literary Terms, 2011, Boston: Cengage Learning.	Glos.	AmE	0	1
Australian Department of Defence Glossary, online, http://www.defence.gov.au/Budget/14-15/2014-2015_Defence_PAES_05_Glossary.pdf	Glos.	AusE	1	0
English–Afrikaans Glossary of Statistical Terms, 2009, https://sastat.org.za/sites/default/files/documents/files/Eng_Afr_fin.pdf	Bi. Glos.	SAE-SAAfr	0	1
Multilingual Statistical Terminology, 2013, Pretoria: Statistics South Africa.	Bi. Glos.	SAE-SALang	0	1
Thesaurus				
InterActiveCorp (Random House), http://www.thesaurus.com	On. Thes.	AmE	1	1
Oxford America Writer's Thesaurus, OS-based thesaurus.	OS Thes.	AmE	1	0
Collins Dictionary of Internet-Linked Synonyms and Antonyms, 2005, Glasgow: HarperCollins.	Pr. Thes.	AmE/BrE	0	1
Macquarie Essential Thesaurus, 2000, Sydney: Macquarie Dictionary Publishers Pty Ltd.	Pr. Thes.	AusE	1	0
Roget's Thesaurus of English Words and Phrases, 1962, 1966, 1987, nd, London: Longman.	Pr. Thes.	BrE	3	1
Chambers Paperback Thesaurus, 1995, Edinburgh: Chambers.	Pr. Thes.	BrE	0	1
Oxford Paperback Thesaurus, 2006, Oxford: Oxford University Press.	Pr. Thes.	BrE	0	1
Webster's English Thesaurus, 2010, New Lanark, Scotland: Geddes & Grosset.	Pr. Thes.	BrE	0	1
Online thesaurus: Publisher unspecified.	On. Thes.	Ind.	1	0
Pharos Groot Tesourus van Afrikaans, 1994, Cape Town: Pharos.	Pr. Thes.	SAAfr	0	1

Appendix C: Source type categorisation

Style guide				
In-house Style Guide, unspecified.	Sty. Gd.	Ind.	44	19
New Hart's Rules: The Oxford Style Guide, 2002, 2005, 2014, nd, Oxford: Oxford University Press.	Sty. Gd.	BrE	3	15
The New Oxford Style Manual, 2003, 2012, 2016, nd, Oxford: Oxford University Press.	Sty. Gd.	BrE	2	14
The Chicago Manual of Style, 1911, , 1917, 1925, 1927, 1993, 2003, 2010, 2017, nd, Chicago: University of Chicago Press.	Sty. Gd.	AmE	37	13
The Elements of Style, 1959, 1979, 2000, 2011, nd, New York: Harcourt/Macmillan/Pearson Longman.	Sty. Gd.	AmE	16	9
Publication Manual of the American Psychological Association, 1994, 2009, nd, New York: American Psychological Association.	Sty. Gd.	AmE	22	7
Pharos Afrikaanse Woordelys en Spelreels, 2007, 2009, 2017, nd, Cape Town: Pharos.	Sty. Gd.	SAAfr	0	5
University of South Africa In-house Style Guide, nd.	Sty. Gd.	SAE	0	4
American Psychological Association Style Blog, online, blog.apastyle.org	On. Sty. Gd.	AmE	2	3
Pharos Skryf Afrikaans van A tot Z, 2011, nd, Cape Town: Pharos.	Sty. Gd.	SAAfr	0	3
NWU Referencing Guide, 2012, Potchefstroom: NWU Library Services.	Sty. Gd.	SAE	0	3
Style Manual for Authors, Editors and Printers, 1966, 1994, 2002, nd, Milton Qld: John Wiley.	Sty. Gd.	AusE	70	2
AMA Manual of Style: A Guide for Authors and Editors, 1976, 2007, New York: Oxford University Press.	Sty. Gd.	AmE	5	2
The Economist Style Guide: The Bestselling guide to English Usage, 2005, 2013, nd, London: Profile Books Ltd.	Sty. Gd.	BrE	4	2
Purdue OWL APA Style, online, https://owl.english.purdue.edu/owl/	On. Sty. Gd.	AmE	0	2
The Chicago Manual of Style online, online, http://www.chicagomanualofstyle.org/home.html	On. Sty. Gd.	AmE	2	1
Online style guide: APA	On. Sty. Gd.	AmE	1	1
Conscious Style Guide, online, https://consciousstyleguide.com/	On. Sty. Gd.	AmE	0	1
Online style guide: Chicago	On. Sty. Gd.	AmE	0	1
The SBL Handbook of Style, 2014, Atlanta, GA: SBL Press.	Sty. Gd.	AmE	0	1
Online style guide: Oxford	On. Sty. Gd.	BrE	0	1
Oxford University Standard for the Citation of Legal Authorities, 2012, Oxford: Faculty of Law, Oxford University.	Sty. Gd.	BrE	0	1
Routledge In-house Style Guide, nd.	Sty. Gd.	BrE	0	1
SAGE Publications Submission Guidelines, online, https://uk.sagepub.com/en-gb/afr/manuscript-submission-guidelines	On. Sty. Gd.	BrE	0	1
The Guardian and Observed Style Guide, online, Guardian Newspapers Ltd: https://www.theguardian.com/info/series/guardian-and-observer-style-guide	Sty. Gd.	BrE	0	1
Natural Resource Governance Institute In-house Style Guide, nd.	Sty. Gd.	Ind.	0	1
Harvard Referencing Style Guide, 2009, online, http://openjournals.net/files/Ref/HARVARD2009%20Reference%20guide.pdf	On. Sty. Gd.	Ind.	0	1
Springer In-House Style Guide, nd.	Sty. Gd.	Ind.	0	1
Afro-Middle East Centre In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
Bibliographic Style and Reference Techniques, 2011, Pretoria: Unisa Press.	Sty. Gd.	SAE	0	1
Department of Higher Education and Training In-House Style Guide, nd.	Sty. Gd.	SAE	0	1
Editorial Style Guide, 2013, Pretoria: Government Communication and Information Systems.	Sty. Gd.	SAE	0	1
International Organisation for Migration In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
Juta Publications In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
Language Centre Style Guide, 2014, Stellenbosch: Language Centre Stellenbosch University.	Sty. Gd.	SAE	0	1

Making Sense of Referencing: The Harvard, APA and Vancouver Methods and Footnote System, 2010, Stellenbosch: Stellenbosch University Language Centre.	Sty. Gd.	SAE	0	1
Momentum In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
Referencing Guide: DUT Harvard Referencing Style, 2017, Durban: Durban University of Technology Library.	Sty. Gd.	SAE	0	1
Solidarity In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
South African Cities Network In-House Style Guide, nd.	Sty. Gd.	SAE	0	1
Southern Business School Harvard Referencing Guide, 2014, Krugersdorp: Southern Business School.	Sty. Gd.	SAE	0	1
Statistics South Africa In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
Troupant Publisher's In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
Van Schaik In-house Style Guide, nd.	Sty. Gd.	SAE	0	1
Australian Guide to Legal Citation, 2010, nd, Melbourne: Melbourne University Law Review Association Inc.	Sty. Gd.	AusE	5	0
Australian Manual of Scientific Style, online, Biotext: https://www.sciencestyle.com.au/	On. Sty. Gd.	AusE	4	0
MLA Handbook, 2016, New York: Modern Languages Association.	Sty. Gd.	AmE	2	0
Scientific Style and Format: The CSE Manual for Authors, Editors and Publishers, 2014, Chicago: University of Chicago Press.	Sty. Gd.	AmE	2	0
Digital Transformation Agency Content Guide, online, https://guides.service.gov.au/content-guide/how-to-use-guide/	On. Sty. Gd.	AusE	2	0
Own Style Sheet.	Own Sty. Gd.	Ind.	2	0
Modern Language Association Style Center, online, https://style.mla.org/	On. Sty. Gd.	AmE	1	0
Purdue OWL, online, https://owl.english.purdue.edu/owl/	On. Sty. Gd.	AmE	1	0
The ACS Style Guide: Effective Communication of Scientific Style, 2006, New York: Oxford University Press.	Sty. Gd.	AmE	1	0
The Associated Press Style Book and Briefing on Media Law, 2015, New York: Basic Books.	Sty. Gd.	AmE	1	0
The Chicago Manual of Style Q&A, online, http://www.chicagomanualofstyle.org/qanda/topicList.html	On. Sty. Gd.	AmE	1	0
The Christian Writer's Manual of Style, 2004, Grand Rapids, Mich: Zondervan.	Sty. Gd.	AmE	1	0
The New York Times Manual of Style and Usage, 2015, New York: Three Rivers Press.	Sty. Gd.	AmE	1	0
Aboriginal Studies Press Publishing Style Guide for Authors and Editors, 2013, Canberra ACT: Aboriginal Studies Press.	Sty. Gd.	AusE	1	0
Animal Health Australia Style Guide, nd, Turner ACT: AHA.	Sty. Gd.	AusE	1	0
Australia National University Press In-house Style Guide, nd.	Sty. Gd.	AusE	1	0
Australian Defence Force: Defence Writing Standards, nd, Canberra ACT: Defence Publishing Service.	Sty. Gd.	AusE	1	0
Australian Department of Agriculture and Water Resources Style Guide, nd.	Sty. Gd.	AusE	1	0
Capstone Editing In-house Style Guide, nd.	Sty. Gd.	AusE	1	0
Cengage Learning Australia Mathematics In-house Spelling guide, nd.	Sty. Gd.	AusE	1	0
Cengage Learning Australia Science In-house Style guide, nd.	Sty. Gd.	AusE	1	0
Finch Publishing In-house Style Guide, 2017.	Sty. Gd.	AusE	1	0
Harvard Referencing Guide UniSA, 2015, online, https://lo.unisa.edu.au/pluginfile.php/438774/mod_book/chapter/39173/HRG%20January%202015%20final%20version.pdf	On. Sty. Gd.	AusE	1	0
La Trobe University Academic Referencing Tool, online, http://www.lib.latrobe.edu.au/referencing-tool/#content	On. Sty. Gd.	AusE	1	0
Queensland University of Technology In-house Style Guide, nd.	Sty. Gd.	AusE	1	0
Queensland University of Technology YuMi Deadly Centre In-House Style Guide, nd.	Sty. Gd.	AusE	1	0

Appendix C: Source type categorisation

South Australian Department for Education and Child Development In-house Style Guide, 2017.	Sty. Gd.	AusE	1	0
Style: The Essential Guide for Journalists and Professional Writers, nd, South Bank, Vic: News Custom Publishing.	Sty. Gd.	AusE	1	0
University of New England APA Referencing, online, http://www.une.edu.au/__data/assets/pdf_file/0007/12958/REF_APA-Referencing.pdf	On. Sty. Gd.	AusE	1	0
University of Queensland Written Style Guide, online, https://omc.uq.edu.au/written-style-guide	On. Sty. Gd.	AusE	1	0
NewSouth Publishing In-house Style Guide, 2017.	Sty. Gd.	AusE	1	0
Oxford University Press Australia and New Zealand In-house Style Guide, nd.	Sty. Gd.	AusE/NZE	1	0
The Guardian and Observer Style Guide, online, https://www.theguardian.com/info/series/guardian-and-observer-style-guide	On. Sty. Gd.	BrE	1	0
HarperCollins Publishers In-house Style Guide, 2017.	Sty. Gd.	Ind.	1	0
Wikipedia Manual of Style, online, https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style	On. Sty. Gd.	Ind.	1	0
Write, Edit, Print: Style Manual for Aotearoa New Zealand, nd, Lincoln, Canterbury: Lincoln University Press.	Sty. Gd.	NZE	1	0
Usage guide				
Cambridge Guide to Australian English Usage, 1995, 2007, nd, Melbourne: Cambridge University Press.	Usg. Gd.	AusE	21	0
Fowler's Modern English Usage, 1957, 1965, 1996, 2004, 2015, nd, Oxford: Oxford University Press.	Usg. Gd.	BrE	18	10
Cambridge Guide to English Usage, 2004, Cambridge: Cambridge University Press.	Usg. Gd.	BrE	11	3
Modern Australian Usage: A Practical Guide for Writers and Editors, 1997, 2015, Melbourne: Oxford University Press.	Usg. Gd.	AusE	7	0
Usage and Abusage: A Guide to Good English, 1969, 1973, 1994, nd, London: Hamish Hamilton.	Usg. Gd.	BrE	4	1
Right Words: A Guide to English Usage in Australia, 1987, 1989, Ringwood, Vic: Viking.	Usg. Gd.	AusE	3	0
The Penguin Working Words: An Australian Guide to Modern English Usages, 1993, Ringwood, Vic: Viking.	Usg. Gd.	AusE	2	0
Engleish, our Engleish: Common Errors in South African English and How to Solve Them, 2009, Cape Town: Pharos.	Usg. Gd.	SAE	1	14
Garner's Modern English Usage, 2016, New York: Oxford University Press.	Usg. Gd.	AmE	1	5
The Complete Guide to English Usage for Australian Students, 2011, Melbourne: Cengage Australia.	Usg. Gd.	AusE	1	0
The King's English, nd, Oxford: Clarendon Press.	Usg. Gd.	BrE	1	0
Online usage guide: Publisher unspecified.	On. Usg. Gd.	Ind.	1	0
Oxford A–Z of English Usage, 2013, nd, Oxford: Oxford University Press.	Usg. Gd.	BrE	0	2
ABC of English Usage, 1970, New York: Oxford University Press.	Usg. Gd.	BrE	0	1
The Right Word at the Right Time: A guide to the English Language and How to Use it, 1985, Reader's Digest.	Usg. Gd.	BrE	0	1
Editing and Writing sources				
Butcher's Copy-editing: The Cambridge Handbook for Editors, Copy-editors, and Proofreaders, 1981, 1992, 2002, 2006, nd, Cambridge: Cambridge University Press.	E&W Hbk	BrE	6	16
Oxford Guide to Plain English, 2009, nd, Oxford: Oxford University Press.	E&W Hbk	BrE	1	7
Pharos The Write Stuff: The Style Guide with a Difference, 2008, 2011, 2014, 2015, nd, Cape Town: Pharos.	E&W Hbk	SAE	0	7
Working Words, 2011, Manuka, ACT: Canberra Society of Editors.	E&W Hbk	AusE	8	3
Effective Writing: Plain English at Work, 1989, 2014, Melbourne: Pitman.	E&W Hbk	AusE	7	3
Collins Good Writing Guide: The Essential Guide to Good Writing, 2003, nd, Harlow: HarperCollins.	E&W Hbk	BrE	1	2

Course materials: Indeterminate origin	E&W Gd.	Ind.	0	2
Pharos The (Essential) Companion for Writers and Editors, 2011, Cape Town: Pharos.	E&W Hbk	SAE	0	2
Scientific Writing Skills: Guidelines for Writing Theses and Dissertations, 2007, Stellenbosch: SUN PRess.	E&W Hbk	SAE	0	2
Text Editing: A Handbook for Students and Practitioners, 2012, Brussels: ASP Editions.	E&W Hbk	SAE	2	1
Random House Guide to Good Writing, 1991, New York: Ballantine Books.	E&W Hbk	AmE	0	1
The Writer's Harbrace Handbook, 2021, Boston: Wadsworth Cengage Learning.	E&W Hbk	AmE	0	1
Good English: How to Write It, nd, London: Pan Books.	E&W Hbk	BrE	0	1
Consistency, Consistency, Consistency: The PEG Guide to Style Guides, 2009, Johannesburg: PEG.	E&W Gd.	SAE	0	1
Marketing Your Freelance Services: The PEG Guide for Practitioners, 2015, Johannesburg: PEG.	E&W Gd.	SAE	0	1
The Business of Editing, 2001, Johannesburg: PEG.	E&W Gd.	SAE	0	1
The Master's Dissertation and Doctoral Thesis: A Guide to Research and the Organization of Material, nd, Pretoria: University of South Africa Press.	E&W Hbk	SAE	0	1
The Writer's Voice: A Workbook for Writers in Africa, 1998, Cape Town: Zebra Press.	E&W Hbk	SAE	0	1
The Australian Editing Handbook, 1994, 2014, nd, Milton, Qld: John Wiley & Sons.	E&W Hbk	AusE	5	0
The Editor's Companion, 2004, 2011, nd, Port Melbourne: Cambridge University Press.	E&W Hbk	AusE	5	0
The Little Red Writing Book, 2011, nd, Sydney: UNSW Press.	E&W Hbk	AusE	3	0
The Sense of Style: The Thinking Person's Guide to Writing in the 21st Century, 2014, New York: Penguin Books.	E&W Hbk	AmE	2	0
Copyediting Newsletter, online, https://www.copyediting.com	E&W Gd.	AmE	1	0
Letting Go of the Words: Writing Web Content that Works, 2007, San Francisco: Elsevier.	E&W Hbk	AmE	1	0
Nielsen Norma Group Be Succinct: Writing for the Web, online, https://www.nngroup.com/articles/be-succinct-writing-for-the-web/#contentBody	E&W Web.	AmE	1	0
On Writing Well: The Classic Guide to Writing Nonfiction, 2006, New York: HarperCollins.	E&W Hbk	AmE	1	0
Poynter Institute, online, https://www.poynter.org/person/rclark	E&W Web.	AmE	1	0
The Copyeditor's Handbook: A Guide for Book Publishing and Corporate Communications, 2006, Berkeley: University of California Press.	E&W Hbk	AmE	1	0
The Gettysburg Approach to Writing and Speaking Like a Professional, 2009, INDI Publishing Group.	E&W Hbk	AmE	1	0
The McGraw-Hill Desk Reference for Editors, Writers and Proofreaders, 2006, New York: McGraw-Hill.	E&W Hbk	AmE	1	0
Australian Handbook for Writers and Editors, nd, Sydney: Woodslane Press.	E&W Hbk	AusE	1	0
Editing Research Theses, online, http://iped-editors.org/About_editing/Editing_theses.aspx : IPEd	E&W Gd.	AusE	1	0
Guidelines for the Ethical Publishing of Aboriginal and Torres Strait Islander Authors and Research from those Communities, 2015, Canberra, Aboriginal Studies Press.	E&W Gd.	AusE	1	0
Speechwriting for Government in Australia, 2010, Kippax: James Groves & Associates.	E&W Hbk	AusE	1	0
Writing: Protocols for Producing Indigenous Australian Writing, nd, Surry Hills, NSW: Australian Council of the Arts.	E&W Hbk	AusE	1	0
Complete Plain Words, 1987, Harmondsworth: Penguin.	E&W Hbk	BrE	1	0
Grammar Rules: Writing with Military Precision, 2011, London: Kyle Books.	E&W Hbk	BrE	1	0
Penguin Writer's Manual, 2002, London: Penguin.	E&W Hbk	BrE	1	0

Appendix C: Source type categorisation

Modern Rhetoric, 1974, New York: Harcourt Brace Jovanovich.	E&W Hbk	Ind.	1	0
Language Reference				
Pharos Online, https://www.pharosonline.co.za/home	On. Lang. Ref.	SAE & SAAfr	0	14
The Oxford A-Z of Grammar & Punctuation, 2004, 2007, 2009, 2013, nd, Oxford: Oxford University Press.	Pr. Lang. Bk	BrE	0	11
Oxford Reference Online, online, www.oxfordreference.com	On. Lang. Ref.	AmE/BrE/ CaE/AusE /NZE	5	9
Grammar Girl: Quick and Dirty Tips, online, https://www.quickanddirtytips.com/grammar-girl	Lang. Web.	AmE	8	3
Grammarly, online, https://www.grammarly.com	On. Gram. Chkr	Ind.	0	2
Eats, Shoots and Leaves: The Zero Tolerance Approach to Punctuation, 2003, London: Profile Books.	Pr. Lang. Bk	BrE	1	1
Online sources: Grammar blogs, unspecified.	Lang. Web.	Ind.	1	1
Merriam-Webster's Pocket Guide to Punctuation, 2001, Springfield: Merriam-Webster.	Pr. Lang. Bk	AmE	0	1
Random House Webster's Grammar, Usage & Punctuation, 1998, New York: Random House.	Pr. Lang. Bk	AmE	0	1
Collins COBUILD English Grammar, nd, Glasgow: HarperCollins.	Eng. Gram.	AmE/BrE	0	1
Collins COBUILD Intermediate English Grammar, 2011, Glasgow: HarperCollins.	Eng. Gram.	AmE/BrE	0	1
Longman Student Grammar of Spoken and Written English, 2002, Harlow: Pearson Education Ltd.	Eng. Gram.	AmE/BrE	0	1
ABC of Common Grammatical Errors, 1995, Oxford: Macmillan Heinemann.	Pr. Lang. Bk	BrE	0	1
British National Corpus, online, https://corpus.byu.edu/bnc	Corp.	BrE	0	1
Everyday Grammar, 2001, Oxford: Oxford University Press.	Eng. Gram.	BrE	0	1
Grammar for Grown-ups: A Straightforward Guide to Good English, 2012, London: Square Peg.	Pr. Lang. Bk	BrE	0	1
The Oxford Companion to the English Language, 1992, New York: Oxford University Press.	Pr. Lang. Bk	BrE	0	1
The Little Green Grammar Book, 2008, Sydney: UNSW Press.	Pr. Lang. Bk	AusE	9	0
Google Books Ngram Viewer	Corp.	AmE/BrE	2	0
Collins Gem English Grammar, 1990, nd, Harlow: HarperCollins.	Eng. Gram.	BrE	2	0
The Chicago Guide to Grammar, Usage and Punctuation, 2016, Chicago: University of Chicago Press.	Pr. Lang. Bk	AmE	1	0
Between You and I: A Little Book of Bad English, 2003, Cambridge: Icon Books.	Pr. Lang. Bk	BrE	1	0
Cambridge Grammar of English: A Comprehensive Guide, 2006, Cambridge: Cambridge University Press.	Eng. Gram.	BrE	1	0
Manual of English Grammar and Composition, 1964, London: Macmillan.	Eng. Gram.	BrE	1	0
My Grammar and I (Or Should That Be Me?), 2011, London: Michael O'Mara Books.	Pr. Lang. Bk	BrE	1	0
Swearing: A Social History of Foul Languages, Oaths and Profanity in English, nd, London: Penguin Books.	Pr. Lang. Bk	BrE	1	0
The Cambridge Grammar of the English Language, 2002, Cambridge: Cambridge University Press.	Eng. Gram.	BrE	1	0
The Right Word: Making Sense of Words that Confuse, 2012, Wollombi, NSW: Exisle publishing.	Pr. Lang. Bk	BrE	1	0
The Use of English, 1968, London: Longman.	Pr. Lang. Bk	BrE	1	0
World Wide Words, online, http://www.worldwidewords.org/	Lang. Web.	BrE	1	0
Online sources: Grammar forum, unspecified.	Lang. Web.	Ind.	1	0
Online sources: Grammar websites, unspecified.	Lang. Web.	Ind.	1	0
People				

Social Media: Editors Victoria	Ppl via Web.	AusE	1	0
Colleagues	Ppl	Ind.	4	4
People	Ppl	Ind.	2	1
Online sources: Editors' groups	Ppl via Web	Ind.	1	0
Social Media: Linguistic Groups, Facebook	Ppl via Web.	Ind.	0	1
Subject experts	Ppl	Ind.	0	1
Professional Editors' Guild Chat-Group	Ppl via Web.	SAE & SAAfr	0	5
ZaLang - South African Language Workers, online chat group, https://groups.yahoo.com/neo/groups/ZaLang/info	Ppl via Web.	SAE & SAAfr	0	2
ProLingua, South African Association of Terminology Experts, https://www.prolingua.org.za/index.php	Ppl via Web.	SAE & SAAfr	0	1
South African Translators' Institute, www.translators.org.za	Ppl via Web.	SAE & SAAfr	0	1
Various online and print sources				
Butterick's Practical Typography, online, https://practicaltypography.com	Sub. Spec. Wbs.	AmE	1	0
Australian Curriculum, Assessment and Reporting Authority, https://www.acara.edu.au	Sub. Spec. Wbs.	AusE	1	0
Australian Securities and Investments Commission, online, asic.gov.au	Sub. Spec. Wbs.	AusE	1	0
Engaging – A Guide to Interacting Respectfully and Reciprocally with Aboriginal and Torres Strait Islander People, and their Arts Practices and Intellectual Property, 2015, online, http://www.zuckermann.org/guide.html	Sub. Spec. Srcs	AusE	1	0
National Library of Australia, online, https://www.nla.gov.au/app/eresources/	On. Srch TI	AusE	1	0
The Design Manual, 2009, Sydney: UNSW Press.	Sub. Spec. Srcs	AusE	1	0
Meter and Meaning: An Introduction to Rhythm in Poetry, 2003, Oxon: Routledge.	Sub. Spec. Srcs	BrE	1	0
Mastering English Literature, 2006, Hampshire: Palgrave Macmillan.	Sub. Spec. Srcs	BrE	0	1
National Health Services Choices, online, https://www.nhs.uk/pages/home.aspx	Sub. Spec. Wbs.	BrE	0	1
Google Search Engine	On. Srch TI	Ind.	8	6
Wikipedia, online, https://en.wikipedia.org/wiki/Main_Page	On. Enc.	Ind.	4	2
Academic Journal Databases: Unspecified	On. Srch TI	Ind.	1	0
Online sources: Organisational websites	Sub. Spec. Wbs.	Ind.	1	0
Subject documents: Unspecified.	Sub. Spec. Srcs	Ind.	1	0
Online sources: Unspecified	On. Srcs	Ind.	0	2
Google Scholar	On. Srch TI	Ind.	0	1

Appendix D: Calculation of total estimated proportions and normalised frequencies

S-genitives

STEP 1

A random sample of 1000 instances of ‘s and s’ were extracted from each register from the unedited texts, differentiated for sub-corpus (AusE and SAfE). Where there were fewer than 1000 instances, all instances were extracted.

- Samples of 1 000 instances were drawn for: SAfE Academic, Creative, Instructional and Popular; AusE Creative, and Popular.
- The whole set of instances was drawn for: SAfE Reportage (336 instances); and AusE Academic (307), Instructional (157) and Reportage (32).

Each instance was manually sorted in a step-wise fashion as follows:

- First, interchangeable instances were sorted from non-interchangeable instances using the criteria set out in Section 4.3.4.
- Using this set of aligned interchangeable instances extracted from the unedited files, each instance was coded as follows:
 - a. If the token was unchanged in the edited, it was coded **Retained** (thus is occurs in the unedited and edited file, unchanged).
 - b. If it was replaced by the other genitive construction, then it was coded **Replaced** (thus is occurs in the unedited files as one of the two variants, and in the edited file and the other).
 - c. If the genitive construction was completely removed from the unedited, then it was coded as **Removed**.

STEP 2:

The edited files were searched for instances where the ‘s or s’ occurred in the edited file but not in the unedited file (thus reflecting an **added** instance). This was differentiated for sub-corpus (AusE and SAfE). There were fewer than 1 000 instances per register, and so all instances were extracted.

- SAfE: Academic (432), Creative (67), Instructional (392), Popular (275), Reportage (44).
- AusE: Academic (7), Creative (540), Instructional (26), Popular (108), Reportage (5).

Each instance was manually sorted in order to identify interchangeable and non-interchangeable contexts using the criteria set out in Section 4.3.4. All interchangeable contexts identified were coded as **Added**. The following total interchangeable instances were obtained.

- SAfE: Academic (194), Creative (64), Instructional (167), Popular (101), Reportage (30).
- AusE: Academic (0), Creative (11), Instructional (0), Popular (0), Reportage (0).

OF-GENITIVES

STEP 3

A random sample of 1000 instances of *of* were extracted from each register from the unedited texts, differentiated for sub-corpus (AusE and SAfE). Where there weren’t 1 000 instances, all instances were extracted.

- Samples of 1 000 instances were drawn for: SAfE Academic, Creative, Instructional and Popular; AusE Academic, Creative, Instructional and Popular
- The whole set of instances was drawn for: SAfE Reportage (897) and AusE Reportage (90).

- Each interchangeable instance was then manually identified using the procedure set out in Section 4.3.4 and was coded as follows:
 - a. If the token was unchanged in the edited, it was coded **Retained** (thus is occurs in the unedited and edited file, unchanged).
 - b. If it was replaced by the other genitive construction, then it was coded **Replaced** (thus is occurs in the unedited files as one of the two variants, and in the edited file and the other).
 - c. If the genitive construction was completely removed from the unedited, then it was coded as **Removed**.

STEP 4:

The edited files were searched for instances where the *of* occurred in the edited file but not in the unedited file (thus reflecting an **added** instance). This was differentiated for sub-corpus (AusE and SAfE). All instances were extracted where there were less than 1 000 instances in the data.

- Samples of 1 000 instances were drawn for: SAfE Academic, Creative, Instructional, Popular
- The whole set of instances was drawn for: SAfE Reportage (58); AusE Academic (441), Creative (781), Instructional (568), Popular (87), Reportage (28).

Each instance was manually sorted in order to identify interchangeable and non-interchangeable contexts using the criteria set out in Section 4.2.4. All interchangeable contexts identified were coded as **Added**. The following total interchangeable instances were obtained.

- SAfE: Academic (336), Creative (415), Instructional (842), Popular (382), Reportage (44)
- AusE: Academic (184), Creative (363), Instructional (31), Popular (11), Reportage (26).

Steps 1–4 yielded the following raw data:

Raw data obtained from the academic register, per sub-variety

	<i>s-genitive</i>					<i>of-genitive</i>				
	AfrE	BSAfe	WSAfe	AusE	Other (SA)	AfrE	BSAfe	WSAfe	AusE	Other (SA)
Non-alternating	68	22	3	42	12	326	112	18	571	97
Retained	659	31	41	250	112	249	84	9	392	73
Removed	22	1	3	10	4	12	1	0	35	5
Replaced	9	6	0	42	11	1	1	0	0	1
Added	116	65	13	0	56	149	554	17	3	155

Raw data obtained from the creative register, per sub-variety

	<i>s-genitive</i>		<i>of-genitive</i>	
	WSAfe	AusE	WSAfe	AusE
Non-alternating	477	543	667	676
Retained	409	361	280	271
Removed	114	92	47	52
Replaced	0	1	0	0
Added	64	11	124	248

Raw data obtained from the instructional register, per sub-variety

	<i>s-genitive</i>					<i>of-genitive</i>				
	AfrE	BSAfe	WSAfe	AusE	Other (SA)	AfrE	BSAfe	WSAfe	AusE	Other (SA)
Non-alternating	98	13	152	71	65	189	52	229	647	138
Retained	225	7	281	79	85	114	31	130	332	73
Removed	6	3	40	7	6	5	3	12	19	7
Replaced	0	0	3	0	3	0	0	1	0	0
Added	74	37	56	0	55	118	664	350	6	412

Raw data obtained from the popular register, per sub-variety

	<i>s-genitive</i>				<i>of-genitive</i>			
	AfrE	BSAfE	WSAfE	AusE	AfrE	BSAfE	WSAfE	AusE
Non-alternating	115	43	327	641	190	105	401	631
Retained	174	64	233	350	101	66	114	354
Removed	6	1	17	3	3	3	12	14
Replaced	1	5	2	1	0	0	0	0
Added	4	7	90	1	17	2	102	2

Raw data obtained from the reportage register, per sub-variety

	<i>s-genitive</i>				<i>of-genitive</i>			
	AfrE	BSAfE	AusE	Other (SA)	AfrE	BSAfE	AusE	Other (SA)
Non-alternating	42	86	12	27	123	221	59	160
Retained	37	46	17	49	83	106	28	142
Removed	5	5	3	1	8	44	3	4
Replaced	0	4	0	0	0	0	0	0
Added	3	27	4	3	2	23	1	0

STEP 5:

The proportional contributions of Non-alternating, Retained, Removed and Replaced for each register, per sub-variety were calculated.

Proportional contribution of Non-alternating, Retained, Removed and Replaced from the academic register, per sub-variety

	<i>s-genitive</i>					<i>of-genitive</i>				
	AfrE	BSAfE	WSAfE	AusE	Other (SA)	AfrE	BSAfE	WSAfE	AusE	Other (SA)
Non-alternating	0,09	0,37	0,06	0,14	0,09	0,55	0,57	0,66	0,57	0,55
Retained	0,87	0,52	0,88	0,83	0,80	0,42	0,42	0,34	0,39	0,41
Removed	0,03	0,02	0,06	0,03	0,03	0,02	0,01	0,00	0,04	0,03
Replaced	0,01	0,10	0,00	0,00	0,08	0,00	0,01	0,00	0,00	0,01

Proportional contribution of Non-alternating, Retained, Removed and Replaced from the creative register, per sub-variety

	<i>s-genitive</i>		<i>of-genitive</i>	
	WSAfE	AusE	WSAfE	AusE
Non-alternating	0,48	0,54	0,67	0,68
Retained	0,41	0,36	0,28	0,27
Removed	0,11	0,09	0,05	0,05
Replaced	0,00	0,00	0,00	0,00

Proportional contribution of Non-alternating, Retained, Removed and Replaced from the instructional register, per sub-variety

	<i>s-genitive</i>					<i>of-genitive</i>				
	AfrE	BSAfE	WSAfE	AusE	Other (SA)	AfrE	BSAfE	WSAfE	AusE	Other (SA)
Non-alternating	0,30	0,57	0,32	0,45	0,41	0,61	0,61	0,62	0,65	0,63
Retained	0,68	0,30	0,59	0,50	0,53	0,37	0,36	0,35	0,33	0,34
Removed	0,02	0,13	0,08	0,04	0,04	0,02	0,03	0,03	0,02	0,03
Replaced	0,00	0,00	0,01	0,00	0,02	0,00	0,00	0,00	0,00	0,00

Proportional contribution of Non-alternating, Retained, Removed and Replaced from the popular register, per sub-variety

	<i>s-genitive</i>				<i>of-genitive</i>			
	AfrE	BSAfE	WSAfE	AusE	AfrE	BSAfE	WSAfE	AusE
Non-alternating	0,39	0,38	0,57	0,64	0,65	0,60	0,76	0,63
Retained	0,59	0,56	0,40	0,35	0,34	0,38	0,22	0,35
Removed	0,02	0,01	0,03	0,00	0,01	0,02	0,02	0,01
Replaced	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00

Proportional contribution of Non-alternating, Retained, Removed and Replaced from the reportage register, per sub-variety

	<i>s-genitive</i>				<i>of-genitive</i>			
	AfrE	BSAfE	AusE	Other (SA)	AfrE	BSAfE	AusE	Other (SA)
Non-alternating	0,50	0,61	0,38	0,35	0,57	0,60	0,66	0,52
Retained	0,44	0,33	0,53	0,63	0,39	0,29	0,31	0,46
Removed	0,06	0,04	0,09	0,01	0,04	0,12	0,03	0,01
Replaced	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00

STEP 6:

The total number of 's and s' and *of* were then counted per unedited and edited file, per register and sub-variety. Using this total count, the estimated number of alternating *s-* and *of-genitives* were reconstructed for each sub-corpus using the proportion contributions calculated in STEP 5. This yielded the estimated total number of alternating *s-* and *of-genitives* for unedited writing per register and sub-variety and edited writing per register and sub-variety, which were used in the analyses in Sections 5.3.1 and 5.3.2.

Reconstructed total counts per alternating construction, per variety, per unedited register

	AfrE		BSAfE		WSAfE		AusE	
	<i>s-genitive</i>	<i>of-genitive</i>	<i>s-genitive</i>	<i>of-genitive</i>	<i>s-genitive</i>	<i>of-genitive</i>	<i>s-genitive</i>	<i>of-genitive</i>
Academic	2483	15846	739	5314	137	682	283	736
Creative					1395	886	1683	2126
Instructional	695	3966	58	1042	1044	1055	97	575
Popular	202	557	76	340	406	643	1092	3505
Reportage	53	199	59	328			21	28
Total count	3433	20568	932	7024	2982	3266	3176	6970
Total Proportion (%)	14	86	12	88	48	52	31	69

Reconstructed proportions of *s-* and *of-genitives*, per variety, per unedited register

	AfrE		BSAfE		WSAfE		AusE	
	<i>s-genitive</i>	<i>of-genitive</i>	<i>s-genitive</i>	<i>of-genitive</i>	<i>s-genitive</i>	<i>of-genitive</i>	<i>s-genitive</i>	<i>of-genitive</i>
Academic	14	86	12	88	17	83	28	72
Creative					61	39	44	56
Instructional	15	85	5	95	50	50	14	86
Popular	27	73	18	82	39	61	24	76
Reportage	21	79	15	85			43	57

Reconstructed normalised frequencies (per 1 000 words) for *s-* and *of-*genitives, per variety, per unedited register

	AfrE		BSaFE		WSaFE		AusE	
	<i>s-</i> genitive	<i>of-</i> genitive						
Academic	2,68	17,11	2,44	17,56	3,00	14,95	7,65	19,90
Creative					3,44	2,18	5,33	6,73
Instructional	2,53	14,42	1,00	17,96	2,84	2,87	2,20	13,01
Popular	3,66	10,10	3,54	15,84	2,78	4,40	3,98	12,32
Reportage	3,74	14,06	1,90	10,57			4,84	6,45

Reconstructed total counts per alternating construction, per variety, per edited register

	AfrE		BSaFE		WSaFE		AusE	
	<i>s-</i> genitive	<i>of-</i> genitive						
Academic	3359	15293	784	5605	140	699	250	596
Creative					1557	2294	1212	1800
Instructional	758	3850	78	1607	957	4492	78	427
Popular	196	559	74	325	467	811	936	2919
Reportage	40	187	78	261			21	28
Total count	4353	19889	1014	7799	3121	8296	2498	5770
Total Proportion (%)	18	82	12	88	27	73	30	70

Reconstructed proportions of *s-* and *of-*genitives, per variety, per edited register

	AfrE		BSaFE		WSaFE		AusE	
	<i>s-</i> genitive	<i>of-</i> genitive						
Academic	18	82	12	88	17	83	29	71
Creative					40	60	40	60
Instructional	16	84	5	95	18	82	16	84
Popular	26	74	19	81	37	63	24	76
Reportage	18	82	23	77			43	57

Reconstructed normalised frequencies (per 1 000 words) for *s-* and *of-*genitives, per variety, per edited register

	AfrE		BSaFE		WSaFE		AusE	
	<i>s-</i> genitive	<i>of-</i> genitive						
Academic	3,67	16,72	2,57	18,37	3,01	15,03	6,53	15,50
Creative					4,25	6,27	3,64	5,40
Instructional	2,69	13,66	1,23	25,31	2,69	12,62	1,70	9,28
Popular	3,43	9,81	3,47	15,22	2,66	4,62	3,33	10,38
Reportage	2,88	13,29	2,78	9,34			4,69	6,25

Appendix E: Participant Information and Consent Forms

PICF for textual material

Supervisor's Name & Title: Prof. Bertus van Rooy and Dr Haidee Kruger

Participant Information and Consent Form

Name of Project: The role of editorial intervention in ongoing language variation and change in South African and Australian English

You are invited to participate in a study of editorial intervention in published English texts. The purpose of the study is to investigate the role that editors (sub-editors, copyeditors, text editors) of different written, published text types play in ongoing language variation and change. Specifically, we are investigating what editors do when they encounter particular linguistic features in the texts that they edit. We are comparing this in two varieties of English: South African English and Australian English.

In order to do this, we are following a corpus-based approach in which we work with large collections of authentic texts. We are building two parallel, aligned corpora of English unedited and edited texts, one for each of the two varieties under investigation. These two corpora each comprise texts from five published text types:

- Academic texts (academic articles appearing in scholarly journals, academic books, and dissertations and theses)
- Instructional texts (schoolbooks, teacher's guides, instruction documents and manuals, "hobby texts" such as DIY books, and administrative documents)
- Popular texts (non-fiction, popular books, and magazines)
- Fiction (narratives including short stories, novellas and novels)
- Reportage (news reports and newsletters)

The study is being conducted by Melanie Ann Law of North-West University, Vaal Triangle campus, South Africa (email: melanie.law@nwu.ac.za) and Macquarie University, Sydney, Australia (email: melanie.law@students.mq.edu.au). The study will be conducted to meet the requirements of the degrees Doctor of Philosophy in Linguistics and Literary Studies at North-West University, Vaal Triangle campus, South Africa and Doctor of Philosophy in Linguistics at Macquarie University, Sydney, Australia, under a Joint Doctoral Supervision Agreement between North-West University and Macquarie University under the supervision of:

- Prof. Bertus van Rooy (bertus.vanrooy@nwu.ac.za) of the School of Languages, UPSET Research Area, North-West University
- Dr Haidee Kruger (haidee.kruger@mq.edu.au) of the Department of Linguistics, Macquarie University

If you decide to participate, we ask you to contribute unedited texts and their edited counterparts for inclusion in the corpus. Please send the texts via e-mail to Melanie Law (melanie.law@nwu.ac.za; or melanie.law@students.mq.edu.au) in one of the following formats: an edited text with track changes

(from which we will be able to recover the original document and the edited document), or two separate documents (the unedited version and the edited version).

We would like to bring the following important points to your attention:

- All texts in the corpus will be anonymised and allocated a code and will not be linked to a particular editor, author, or publishing institution. We are only interested in the raw linguistic material, which we will use to extract and compare the relative frequency of certain linguistic features in the unedited and edited texts.
- Because different varieties of English are spoken in Australia and South Africa, we do need to know the author's mother tongue, as well as that of the editor. However, that is all the additional information we need for the analyses.
- Although all texts will be anonymised, we will sometimes cite from the texts for exemplification purposes. If authors, editors or publishing institutions are strongly against being cited in the research, then we will note this for the particular text and will never cite extracts from that text in the research. Where texts are cited, care will be taken to ensure that the citation is not traceable using online sources (such as search engines).
- The construction of this corpus is not for commercial use. It will only be used for research purposes by the investigators mentioned above. While the corpus is collected in the first instance for the study outlined above, the corpus may be used for future studies by the investigators mentioned above.
- We understand that there may be confidentiality agreements between editors, authors and publishing institutions, and are therefore willing to seek the necessary permission from the relevant person if required.

These principles are followed to ensure that no individual author's or editor's work will be readily identifiable in any published results of the research.

Any information or personal details gathered in the course of the study are confidential, except as required by law. No individual will be identified in any publication of the results. Only the investigators named above will have access to the data. A summary of the results of the data can be made available to you on request by contacting Melanie Law on one of the emails given above.

Participation in this study is entirely voluntary: you are not obliged to participate and if you decide to participate, you are free to withdraw at any time without having to give a reason and without consequence.

Submission of texts as requested will be regarded as consent to use them for research purposes.

The ethical aspects of this study have been approved by the North-West University Research Ethics Committee for Language Matters (ethics clearance reference number: NWU-00558-16-A8) and the Macquarie University Human Research Ethics Committee (ethics clearance reference number: 5201600843D). If you have any complaints or reservations about any ethical aspect of your participation in this research, you may contact the North-West University Research Ethics Committee for Language Matters through the Chairperson, Prof. Bertus van Rooy, (telephone 0027 16910 3482; email bertus.vanrooy@nwu.ac.za) or the Macquarie University Human Research Ethics Committee through the Director, Research Ethics & Integrity (telephone (02) 9850 7854; email ethics@mq.edu.au). Any complaint you make will be treated in confidence and investigated, and you will be informed of the outcome.

(INVESTIGATOR'S [OR PARTICIPANT'S] COPY)

PICF for survey

Chief Investigator's / Supervisor's Name & Title: Dr Haidee Kruger and Prof. Bertus van Rooy

Participant Information and Consent Form

Name of Project: The role of editorial intervention in ongoing language variation and change in South African and Australian English

You are invited to participate in a study of editorial intervention in published English texts. The purpose of the study is to determine what role editors play in the processes of language variation and change, and how their editorial decisions are guided by norm-providing sources. We would like to examine this in two varieties of English: Australian English and South African English.

The study is being conducted by Melanie Ann Law of Macquarie University, Sydney Australia (melanie.law@hdr.mq.edu.au) and North-West University, Vaal Triangle campus, South Africa (melanie.law@nwu.ac.za). The study will be conducted to meet the requirements of the degrees Doctor of Philosophy in Linguistics at Macquarie University, Sydney, Australia and Doctor of Philosophy in Linguistics and Literary Studies at North-West University, Vaal Triangle campus, South Africa, under a Joint Doctoral Supervision Agreement between Macquarie University and North-West University under the supervision of:

- Dr Haidee Kruger (haidee.kruger@mq.edu.au) of the Department of Linguistics, Macquarie University, Sydney Australia
- Prof. Bertus van Rooy (bertus.vanrooy@nwu.ac.za) of the School of Languages, UPSET Research Focus Area, North-West University, Vanderbijlpark, South Africa

If you decide to participate, you will be asked to complete an online questionnaire. This questionnaire has two sections:

1. Background information
2. Sources used during editing

The questionnaire should not take longer than 30 minutes to complete.

There is no risk or discomfort involved in the completion of the questionnaire.

Any information or personal details gathered in the course of the study are confidential, except as required by law. No individual will be identified in any publication of the results. Only the investigators named above will have access to the data. A summary of the results of the data can be made available to you on request by contacting Melanie Law on one of the emails given above.

Participation in this study is entirely voluntary: you are not obliged to participate and if you decide to participate, you are free to withdraw at any time without having to give a reason and without consequence.

Completion of the questionnaire will be regarded as consent to use the information for research purposes.

The ethical aspects of this study have been approved by the Macquarie University Human Research Ethics Committee (ethics clearance reference number: 5201600843D) and the North-West University Research Ethics Committee for Language Matters (ethics clearance reference number: NWU-00558-16-A8). If you have any complaints or reservations about any ethical aspect of your participation in this research, you may contact the Macquarie University Human Research Ethics Committee through the Director, Research Ethics & Integrity (telephone +61 298 507 854; email ethics@mq.edu.au), or

the North-West University Research Ethics Committee for Language Matters through the Chairperson, Prof. Bertus van Rooy, (telephone +27 16 910 3482; email bertus.vanrooy@nwu.ac.za). Any complaint you make will be treated in confidence and investigated, and you will be informed of the outcome.

(INVESTIGATOR'S [OR PARTICIPANT'S] COPY)

Appendix F: Ethics approval letter

Macquarie University Faculty of Human Sciences Ethics committee approval letter (double click on icon to view the letter).



RE_ HS Ethics
Application - Approv