WRITTEN AFRIKAANS

SINCE STANDARDIZATION

A Century of Change

JOHANITA KIRSTEN

Written Afrikaans since Standardization

Written Afrikaans since Standardization

A Century of Change

Johanita Kirsten

For B

Contents

Figures

Tables

Acknowledgments

I have to acknowledge certain important contributions to my corpus collection efforts and thank several people and institutions for assisting me in compiling my *Historical Corpus of Written Afrikaans: CTexT* (Centre for Text Technology, North-West University) for the use of the *PUK-Protea Corpus* and the *Taalkommissiekorpus*; Ana Deumert for the use of the *Corpus of Cape Dutch Correspondence*; the Pretoria repository of the National Archives of South Africa, for access to letters from 1911 to 1920, 1941 to 1950, and 1971 to 1980; the North-West University Archives for access to letters from 1941 to 1950; the Epog Archives for access to letters from 2001 to 2010; all the anonymous individuals for access to letters and diaries from private collections; and my temporary assistant, Chantelle Kruger, for hours of scanning, typing, and quality control.

The North-West University, specifically the School of Languages, as well as the UPSET Research Focus Area, for many years of supporting my research, funding my corpus collection efforts, and giving me a semester of research leave to work on studies and research publications.

Anette Rosenbach and Letizia Vezzosi, for collaborating with me on genitive research, reading my work and giving valuable feedback, and welcoming me into your homes when the opportunity arose.

Roger Lass—the first time I came into contact with you was just after I finished reading your book on historical linguistics. I was in awe and inspired by your work. And as our friendship developed, I've come to realize how much more I value the person *behind* the work. I have turned to you in difficult times, and you've always been there, ready with words of understanding and wisdom, willing to engage on a deep emotional and intellectual level. You have contributed not only to my development as a scholar but also to

xiii

the person I am today, in ways for which I do not have the words to express my gratitude.

Bertus van Rooy—my first real memory of your teaching is a subtle joke you made in Morphology class that made a few of us smile with anticipation for the rest of the semester. Your way with language and linguistics intrigued me, and your advice and support during my undergrad studies were encouraging and inspiring. You are the reason I pursued further studies and have opened up so many opportunities for me since. You are still my mentor and my inspiration. If I think of all the hours you've taken out of such a busy life to read my work, comment on it thoughtfully, and talk extensively about linguistics and research and so many other enriching things; if I think of the patience and understanding with which you handled episodes of despondency and anxiety, if I think of the ways you continue to support my career, I can only say one completely inadequate thing: *thank you*.

I have a long list of colleagues and friends I want to thank for the meaningful things you contribute to my life and work: Jacques Heyns and Gordon Matthew for hundreds of mugs of coffee, hours of discussion, dry runs of conference papers, and IT support; Koos de Villiers for always finding humor in things; Emile Hoffmann for the quiet comments; Rickus Ströh for your friendship and support; Ian Kotzé for making me look good as a supervisor and collaborating on articles; Ansie Maritz for taking care of my teaching while I was on leave and being such a wonderful colleague; Henri Laurie for so many discussions and never hesitating to challenge me; Mashuda Ebrahim for the meals we shared and all your supportive messages; Mandy Strauss for being such an amazing longtime friend and always keeping things interesting; Khululani Sizwe Mehlwana and all the others from the Bophelong Tutoring Programme for understanding my absence to work on the book and continuing to perform your important tasks; Selina Mohapi for always helping me to keep things together; Caroline Piotrowska for not hating me after my questionable driving on our data collection trip, for letting me use your camera on that trip, and for your pleasant company; Gustav and Anneke Butler for occasionally reminding me to enjoy things outside of work; and Susan Coetzee-Van Rooy for your leadership and example at the university and the joy you always bring.

My parents in law, Lotte and Christo van der Merwe, for your interest in and excitement about my work, that your hearts and your home are always open for me, and that you bear with my absences while working and writing.

My parents, Rika and Koos Kirsten, for raising me with love and patience, for exposing me to so many books growing up, for always encouraging and supporting my intellectual endeavors, for all the phone calls and visits and just being such loving, caring parents.

My love, CW. For bringing love and peace, respect and humility, joy and trust, balance and good things. For the journey.

Abbreviations

AUX	auxiliary
CONTR	contraction
DIM	diminutive
GEN	genitive
HON	honorific
INF	infinitive
INFORM	informal
NEG	negative
PASS	passive
PL	plural
PRS	present tense
PST	past tense
PTCL	particle
REFL	reflexive
SG	singular

1

Introduction

CONTEXTUALIZATION: INVESTIGATING LANGUAGE CHANGE

Language change is a common phenomenon that's already been recognized by early linguists—as Sapir (1921, 160) mentions, "All grammars leak." In spite of De Saussure's ([1916] 1966, 54) opinion that the diachronic study of language change is of lesser importance for linguistics, other early linguists attended to it more closely. Bloomfield (1914, 21) refers to the "unceasing process" of language change in spite of its apparent stability. Sapir (1921, 39) states that "[l]anguage . . . has drift" (Sapir 1921, 160). However, they also claim that language change progresses very slowly (Bloomfield 1914, 21; Sapir 1921, 128) and that speakers themselves cannot perceive these changes (Bloomfield 1914, 198). In this regard, Bloomfield (1933, 347) states clearly, "The process of linguistic change has never been directly observed; we shall see that such observation, with our present facilities, is inconceivable."

Bloomfield said this over 80 years ago, however, and in the meantime "our present facilities" have changed quite dramatically. Mair (2006, 15) claims that "[a]ll things considered, we are much better placed now than a century ago for the study of ongoing language change." What Mair refers to here is the use of corpus linguistic methods to investigate the course of ongoing language change.

In spite of the recent developments in tools to investigate language change, most of the general comments on recent and ongoing language change are still based on anecdotal evidence of a few changing phenomena (Mair 2006, 15). These comments often come from a prescriptivist angle (Leech et al. 2009, 18) and are sometimes repeated over and over, "gaining a life of their

1

own and solidifying into a body of folk-linguistic knowledge whose truth is taken for granted and no longer challenged even in scholarly publications" (Mair 2006, 18). A further shortcoming of such anecdotal observations of language change is that those who make these comments often focus on the "visible" tip of the iceberg of change, without taking the underlying changes or shifts into account (Mair 2006, 85).

It is exactly for the sake of accuracy and reliability that Mair (2006, 34) and Leech et al. (2009, 19) suggest corpus methods to investigate current assumptions about ongoing language change. Leech et al. (2009, 18) point out that many of the current hypotheses and assumptions about ongoing change can often be good starting points for further investigation, as they are often not completely unfounded (even if they are sometimes exaggerated or out of context). The advantage of using corpora for studying language change is that both the changing phenomena and the stable phenomena can be observed and documented (Leech et al. 2009, 19). Mair (2006, 3) says in this regard,

> Corpora make it possible to describe the spread of individual innovations against the background of the always far greater and more comprehensive continuity in usage, and corpus-based studies of linguistic change in progress are therefore likely to correct more alarmist perceptions based on the unsystematic collection of examples or impressionistic observation, which are inevitably biased toward the strange, bizarre, and unusual.

Corpus linguistic methods can be used not only to document surface changes (phonological and lexical) but also to investigate deeper underlying grammatical changes. Grammatical change differs from lexical and phonological change in at least two regards: first, it proceeds at a much slower pace and often take several centuries to develop to completion; second, it happens below the level of consciousness of speakers (Mair 2006, 82). Leech et al. (2009, 19) also point out that grammatical changes proceed at different paces in different varieties and text types. To investigate these complex grammatical changes that are often difficult to observe, the linguist needs a more solid foundation to work from than subjective and nonsystematic observations can provide (Leech et al. 2009, 8)—and corpus linguistics offer a suitable alternative. Using corpus linguistic methods, a linguist can investigate changes in the frequency of constructions as well as other quantitative patterns in language use.

Mair (2006) uses corpora and comparative corpus linguistic methods to investigate and describe changes in twentieth-century written English, and Leech et al. (2009) is a similar but expanded study. Mair (2006, 83) points out that while the course of any one change is unlikely to come to completion in one century, we can observe salient grammatical changes through the course

of a century. While he refers to English in that statement, there is no reason that it would not be true for another language, like Afrikaans.

The formation of Afrikaans has been investigated and described from a variety of angles. On the one hand, there is the philological school, introduced by Scholtz (1950, 1963, 1980) and continued by, among others, Raidt (1989, 1991) and Conradie (1986). These scholars focus especially on the origin and early development of Afrikaans and base their research on archival sources and collections of written documents from the seventeenth to the nineteenth century. An important viewpoint of the philological school is that modern Afrikaans already existed in essence in the vernacular at the end of the eighteenth century (Raidt 1991, 171) and that speakers' written language diverged from the vernacular to quite some extent in attempts to write "acceptable" Dutch (Raidt 1991, 135). A different perspective on the origin and development of Afrikaans comes from abroad from scholars like Den Besten (2012c), Roberge (1994, 2002), and Deumert (2002, 2004). These scholars question the tendency of the philological school to interpret written language as scribal artifacts (Deumert 2004, 19)—they interpret variation in the written language as a more direct representation of the spoken language and consequently also attend to later variation more closely.

There has not been published a comprehensive study of linguistic change in twentieth-century Afrikaans. The dissertation of Stoops (1972) investigates language use in a small, specialized corpus over a short period of time: 50 years of language use in *Die Huisgenoot* (The House Mate / Home Companion), a popular Afrikaans magazine, using data from ten-year intervals between 1920 and 1970. Other investigations include specific changes during the twentieth century; for example, Conradie (1992) looks at change in the use of tense in subsequent Afrikaans Bible translations and later at tense in Afrikaans narratives (Conradie 1998), Deumert (2004) investigates changes in Cape Dutch letters from the late nineteenth to the early twentieth century, Kirsten (2013) studies the standardization of verbal forms in the early twentieth century, and Uys (1983) focuses on trends of Dutchification in Afrikaans during initial standardization. The MA dissertation of Roux (2001), "'n Kritiese bespreking van taalverandering met verwysing na Afrikaans en sy variëteite" (A critical discussion of language change with reference to Afrikaans and its varieties), discusses general processes of language change and how it can be applied to Afrikaans but does not look at any specific instances of change. We also have a number of studies focusing on lexical change: Coetzee (1990) investigates *r*-deletion as a type of lexical change, Swart (1995) looks at changes in the Afrikaans lexicon of news media, Van der Merwe (2003) investigates semantic changes in the lexicon, and Van Helden and Coetzee (2006) focus on lexical changes in the 1936 and 2001 Afrikaans hymnbooks. Otherwise,

there are a number of sociolinguistic studies available on more recent or current language variation, such as the chapters on Afrikaans and its varieties in *Language and Social History* (Mesthrie 1995), the collection *Taallandskap* (Linguistic landscape) (Carstens and Grebe 2001), and many other articles about variation on different grammatical levels.

It becomes clear, then, that all the available studies on recent changes in Afrikaans focus on one or a few aspects of change and/or focus on only one text type or text and/or describe changes with one or two intervals. In this study, I attempt to give a more comprehensive view of linguistic changes in written Standard Afrikaans over the course of a century. I investigate different aspects of change in four categories during the past century, using corpora compiled from different text types that represent four intervals. The aim of this study is to fill the gaps in the literature that I just mentioned and serve as a background against which further descriptions of ongoing change can be contextualized. I also attend to the role of language contact in Afrikaans language change as well as how certain changes in Afrikaans fit into the bigger West Germanic language family.

However, this is not only a descriptive endeavor aimed at linguistic changes in Standard Afrikaans. Because of certain unique aspects of the history of Afrikaans, it enables reflection on certain topics within the theory of language change.

BUT WHAT IS LANGUAGE CHANGE?

In the description of language change, it is necessary to clearly describe what language in a historical context is and how language change works. There are many ways to define and describe language, particularly in a historical context. Lass (1997), Croft (2000), and Mufwene (2008) all see similarities between biological evolution and language change and develop paradigms within which language change can be described using the same concepts as biological evolution.

For Afrikaans, the distinction between internally motivated and externally motivated language change is a particularly problematic one. Afrikaans, with its creole background, originated in a language contact situation and has been in contact with different languages ever since, particularly English. Donaldson (1988, 1995) claims that English has had and still has a profound influence on Afrikaans because of ongoing contact. While Donaldson (1988, 10) intends to distinguish between *interference* and language change, he tends to focus on Anglicisms and the influence of English on the Afrikaans idiom. He furthermore points out the interesting possibilities for investigating language

change that flows from the contact situation in South Africa (Donaldson 1988, 138; 1995, 226).

Donaldson is by far not the only scholar arguing for English influence on Afrikaans. From a prescriptive perspective, Carstens (2018) regularly refers to the role of English in so-called language errors—for example, many apparently mistaken ways of using certain pronouns are attributed to English influence (Carstens 2018, 75–79) as well as some issues in the past-tense passive construction (Carstens 2018, 401–3). This is not a recent development in prescriptivism, as Malherbe pointed to this issue in past-tense passives in 1920 already and attributed it to English influence (De Villiers 1971, 34). Also from a more descriptive perspective, variation and change in Afrikaans are sometimes linked to English, for example, Ponelis in his description of a number of syntactic constructions, including pronoun use (Ponelis 1979, 88) and the passive (Ponelis 1979, 267).

However, English influence on Afrikaans language use on a deeper grammatical level is not as easy to determine as Donaldson and others apparently assume. Poplack, Zentz, and Dion (2012b, 250) point out the assumption of both linguists and laypeople that language contact necessarily causes changes in the smaller language. Lass (1997, 197–98) says about this assumption and its consequences,

> This is certainly in principle a valid approach for extracting histories from character correspondences; but it is unreliable in itself, and requires considerable background research before it can be trusted. Cases like this are difficult because they are apparently so "obvious," but may turn out to be garden paths.

He points out that apparent similarities between languages in contact do not *necessarily* indicate influence and that thorough comparative work remains necessary, as "the mere fact of extensive contact does not merit being assigned a causal role" (Lass 1997, 202). Poplack, Zentz, and Dion (2012b, 253) similarly do not reject the concept of external language change but emphasize that it should be handled with care, without assuming that apparent similarities necessarily come from language contact. They further develop a number of criteria to determine, first, whether change has occurred in the contact situation and, if so, if it was brought about by language contact.

Unfortunately, it is not as simple to apply the criteria to Afrikaans as it is with Canadian French in Poplack, Zentz, and Dion's (2012a, 2012b) study—there are limited data of Afrikaans prior to contact with English, and there are no varieties of Afrikaans without prolonged contact with English at least to some extent. That contact with English could enhance or speed up a change is a different question, even though Lass (1997, 200) points out how difficult it is to pinpoint (and measure) such influence because nobody knows how the

change would have proceeded without language contact. Lass (1997, 209) then proposes a conservative approach:

> Therefore, in the absence of evidence, an endogenous explanation of a phenomenon is more parsimonious, because endogenous change *must* occur in any case, whereas borrowing is never necessary. If the (informal) probability weightings of both source-types converge for a given character, then the choice goes to endogeny.

Mufwene (2003, 276), on the other hand, emphasizes that "contact has played an important part in the histories of all languages" and it is exactly because it has played such an important role in the formation of Afrikaans that we cannot just ignore it. While issues around language contact and influence in the twentieth century should be handled with caution according to some, the ongoing contact in the history of Afrikaans also provides an interesting opportunity: in Afrikaans, we can observe language change in a recently standardized variety of a creole that has never really existed without ongoing language contact.

There is one particular theoretical point of order that I want to engage with in this book—the issue of the relationship between internal and external factors of language change: what the role of language contact is in change and what the implications are not only for Afrikaans specifically but also for any language in contact. On the one hand, Lass (1997, 197–98) claims that influence from another language should not get priority if internal factors could explain the change, while, for example, Mufwene (2003, 276) emphasizes the importance of language contact in the history of all languages. Donaldson's (1988, 1995) publications are good examples of the tradition of attributing changes in Afrikaans to contact with English, and it becomes important to explore this issue more deeply. While influence from another language is not "undesirable" and to be explained away, I have to be responsible and rigorous in explaining specific instances of language change. This also circles back to the literature on the theory of language change.

WHERE TO LOOK FOR AFRIKAANS LANGUAGE CHANGE

There are a number of grammatical categories that invite further investigation in terms of historical changes. These grammatical categories are temporal reference, pronouns, the genitive, and connectives. Apart from existing linguistic research and preliminary data analysis, I considered comments in prescriptive sources, especially when they refer to alleged changes, as starting points for investigation.

Time and Tense

Ponelis (1979, 261) describes *tense* as systematic temporal differences as it is expressed on the verb. He further describes the Afrikaans tense system as the distinction between the present tense (*presens* in Afrikaans) and the preterit (*preteritum* in Afrikaans) (Ponelis 1979, 263); the preterit, according to him, is the old perfect that was adapted as the new past tense (Ponelis 1979, 264). According to Conradie (1992, 61), the present tense refers to simultaneity with the time of the matrix clause, while the marked preterit (or, historically, the perfect) expresses completeness in relation to the matrix clause, or preceding events. De Villiers (1971, 23) refers to the present tense, the perfect and imperfect; he also claims that the present tense is the unmarked category and the perfect the marked category in Afrikaans verb phrases and that the present tense is used more often and in more contexts (De Villiers 1971, 36). Furthermore, he claims that the traditional distinction between the perfect and imperfect has largely disappeared in Afrikaans (De Villiers 1971, 47). Ponelis (1979, 264) and Conradie (1992, 61; 1998, 37) agree by also referring to the present tense as the unmarked form.

In Afrikaans language use, tense is indicated primarily analytically, that is, with the contrast between a single verb (like *gee* 'give') and a combination of verbs (like *het gegee* 'have given') (Ponelis 1979, 264). However, temporal reference in Afrikaans is not limited to this contrast, as reference to time can also be made with adverbs and phrases (De Villiers 1971, 47; Ponelis 1979, 265). Van der Merwe (1996), for example, wrote an article about *nou* 'now,' *toe* 'then,' and *dan* 'then' as temporal lexical elements in Afrikaans.

There are a number of studies on changes in Afrikaans tenses, among them that of Conradie (1992), where he compares tense usage in the consecutive Afrikaans Bible translations. He points out that a general decline in the use of the unmarked form and increase in the use of the marked preterit can be observed (Conradie, 1992, 60). He also puts forward an important question (Conradie 1992, 66) that is pertinent to this study:

> 'n Vraag wat nou gestel moet word, is in watter mate die ontwikkeling in die gebruik van tempusvorme oor die afgelope eeu 'n weerspieëling is van 'n ontwikkeling in gesproke of geskrewe Afrikaans oor dié tydperk.
>
>   (A question that should be posed now, is to what extent the development in the use of tense forms in the past century reflects that of developments in spoken or written Afrikaans during this time.)

Conradie continues his investigation of Afrikaans tense in an article in 1998 on tense in Afrikaans narratives (Conradie 1998). He focuses especially on the decline of the preterit as preferred narrative tense and the rise, in its

place, of the unmarked present tense and marked past tense, comprised of the perfect and remnants of the preterit in the verb *wees (was)* 'be' (+past tense) and a few modal auxiliaries (Conradie 1998, 37).

While these studies provide a framework within which to start working on more general changes in temporal reference, it has not been done more comprehensively yet. The above-mentioned linguists, as well as Carstens (2018) as a prescriptivist, point out a number of phenomena related to temporal reference in Afrikaans that invites further investigation. I already mentioned Conradie's (1992, 66) question, and the disappearance of the preterit (Conradie 1998) can also be further investigated, especially when used in other contexts than narratives.

A phenomenon that receives attention in the literature is the form (preterit or otherwise) of modal auxiliaries. De Villiers (1971, 31) refers to the use of *kon* 'could' as a variant of *kan* 'can' in infinitive constructions in spoken language but does not encourage this usage because of its supposed low frequency. Ponelis (1979, 271) describes the preterit forms of all deontic modal auxiliaries after *het* 'have' and shows that certain epistemic modal auxiliaries can remain in the present tense in similar contexts (also see De Villiers 1971, 29), particularly in spoken language. De Villiers (1971, 24) furthermore points to the forms *dag/dog* 'thought,' *wis* 'knew,' and *had* 'had' as preterits—*dag/dog* is supposedly frequently used in spoken language, *wis* is becoming rare, and *had* is rare already. Instead, speakers use *gedink* 'thought' (and *gedag/gedog*), *het geweet* 'have known,' and *het gehad* 'have had' (De Villiers 1971, 24).

Furthermore, there are questions relating to the use of one tense to refer to another, for example, the historical present (De Villiers 1971, 48), as well as the matter of changes in the use of adverbs for temporal reference (Kirsten 2016).

It is clear, then, that there are some unresolved issues regarding temporal reference in Afrikaans, justifying further investigation regarding Afrikaans language use in the past century. Contextualizing these phenomena within the wider West Germanic language family might also shed light on some of the issues.

Pronouns

There is a range of questions regarding pronoun use in Afrikaans that require further investigation, drawing comments from linguists and prescriptivists, pointing toward possible variation and change.

One prominent issue in the literature is the use of *-self* '-self' with reflexive pronouns (Carstens 2018, 75; Ponelis 1979, 88). Carstens (2018, 75–76)

identifies certain contexts where the use of -*self* is justified, but both he and Ponelis (1979, 88) claim that there is a rise in the use of reflexive pronouns with -*self* even when it is technically not necessary. This is then attributed to English influence. Carstens (2018, 76) further refers to the use of *geniet* 'enjoy' with especially *jouself* 'yourself' and other similar constructions and proceeds to explain that it is unacceptable in Afrikaans.

Another issue receiving a lot of attention is the independent use of the demonstrative pronouns *hierdie* 'this/these' and *daardie* 'that/those' (Carstens 2018, 78–79; Ponelis 1979, 89). Ponelis (1979, 90) and Carstens (2018, 78) attribute this to English influence, and Carstens explicitly labels it as incorrect in Afrikaans.

The use of a preposition with the relative pronoun *wat* 'who/that' instead of the forms *waarop* 'there.up,' *waarmee* 'there.with,' and so on also draws some attention (Carstens 2018, 73). He also points out the pronoun *wie* 'who,' which is apparently increasingly being used, "incorrectly," independently as a relative pronoun (Carstens 2018, 72).

Other categories of pronoun use that might show some changes are generic pronouns (Ponelis 1979, 105–6; Van Rooy 1996) and specifically the use of masculine forms as "gender neutral" as well as the declining use of the honorific second person pronoun *u* 'you.HON.'

Genitive

The label *genitive* is not uncontroversial in Afrikaans linguistics. According to the very prominent Afrikaans syntactician Ponelis (1979, 126; 1989, 268), the Afrikaans genitive is expressed with (only) possessive pronouns and *se*, which is similar in some ways to the English and Dutch *s*-genitive. Ponelis (1979, 126) does not attempt to clearly define the genitive but emphasizes that the basic meaning of the construction cannot be more specific than that it expresses a relation between referents.

Putting the terminological issues aside for the moment, there are indications of some changes in what we can broadly refer to as possessive constructions in Afrikaans. In the early twentieth century, the genitive was still very variable in written language, and both the Dutch *s*-genitive as well as some Dutch possessive pronouns were still being used sporadically. This loss of variability can be further investigated. Regarding later developments, Vink (1981, 432) claims that the possessive particle *se* has been expanding in its contexts of use at the expense of *van* 'of' but does not present any quantified data. Furthermore, there are many studies on changes in genitive alternation in related West Germanic languages, such as Dutch (Weerman and de Wit 1999) and especially English (Hinrichs and Szmrecsanyi 2007). Changes

in Afrikaans genitive alternation would not be unexpected in this broader context.

Connectives

In the normative Afrikaans literature, the use of subordinators sometimes receives attention regarding the word order changes that "should" happen in the subordinate clause (Carstens 2018, 54–55; Müller 2003, 664), and possible English influence is mentioned here when independent word order is maintained. Müller also makes several comments regarding the use of certain conjunctions, like when to use *as* 'if' and when to use *wanneer* 'when' (2003, 32) or *(al)hoewel* 'however/although,' *ofskoon* 'although,' and *terwyl* 'while' (2003, 102), and certain overlap and other distinctions between *eerder* 'rather/before' and *liewer* 'rather' (2003, 75), and she comments on some connectives and adverbials that are often used unnecessarily, according to her, like *dan* 'then' (2003, 63) and *verder* 'further' and *voorts* 'furthermore' (2003, 225). Furthermore, Bosman and Otto (2012, 70) lament the inadequacy of dictionary entries on many Afrikaans subordinators and the lack of guidance the dictionaries give in spite of errors (especially regarding word order) that many speakers apparently make.

From the more descriptive side, Messerschmidt and Messerschmidt (2006, 16) describe the category of subordinators that form through *-dat* '-that' compounds and mention a recent addition that has not entered into dictionaries. They also offer in-depth descriptive research on specific conjunctions (Messerschmidt and Messerschmidt 2011; 2012), and there are other articles that focus on usage patterns in student and research writing (Van Rooy and Esterhuizen 2011) and general factors that cause the use or the omission of the complementizer *dat* 'that' (Van Rooy and Kruger 2016). None of these descriptive studies focus on changes, however, leaving room for further exploration in this regard.

SPECIFIC QUESTIONS

The specific questions that I aim to answer in this book are twofold, one mainly empirical and the other more theoretical in focus:

1. What is the nature and extent of linguistic changes in written Afrikaans from 1911 to 2010, with a focus on temporal reference, pronouns, the genitive, and connectives?

2. What are the differences (if any) between internal and external language change, and, based on that, what can we learn from the data regarding language contact and its influence on language use?

I will work toward answering these questions through exploring the relevant literature and discussing the relevant theoretical matters as well as with corpus linguistic methods.

OVERVIEW OF THE BOOK

In chapter 2, I report on and discuss the relevant language theoretical literature pertaining to especially language change to properly inform the analyses and interpretation in the later chapters. Following that, in chapter 3, I focus a bit on language standardization, with some consideration of the role of ideology and an account of the history of Afrikaans standardization and language attitudes from the past century. In chapter 4, I describe and discuss the data gathering, extraction, and analysis methods I use in some detail. The following chapters each attend to different types of language change that emerge from the categories I investigate: chapter 5 focuses on paradigmatic changes; in chapter 6, I explore developments in grammaticalization; and in chapter 7, I focus on broader language change regarding connectives, formalization, and colloquialization and a decline in linguistic sexism. In chapter 8, I return to the research questions above and conclude in drawing together the empirical findings, the theoretical literature, and the relevant ideological matters.

2

The Theory of Language Change

LANGUAGE THEORY

In current linguistics, the omnipresent, ongoing change of language is a given. Bloomfield (1933, 281) already claims that the apparent stability of a language is only an illusion and that every language is always changing. Sapir (1921, 160) refers to the *drift* of language as well. A responsible linguistic theory should, then, be able to account for such a common phenomenon, and there is indeed a lot about language change in linguistic theory.

In this chapter, I strongly rely on usage-based theories of linguistics, with concepts such as *emergent grammar* and *exemplar models*, to lay the groundwork of further investigations into linguistic change. In addition, I engage with the theory of language change, investigating certain themes within this domain: language change as evolution, the distinction between the origin and the spread of a linguistic change, and the possibility of "unnatural" change. Finally, I look at the explanation of linguistic change, attending to good argumentation, the role of speakers, internally and externally motivated change, and forces of linguistic change.

LINGUISTIC-THEORETICAL PERSPECTIVE

A Theory of Language and Linguistics

Traditional formal linguistics describes language as rule based, and Sapir's leaking of grammars often cause theoretical and descriptive problems. This is, however, not the only theoretical possibility, as MacWhinney (2001, 449) puts it:

13

> Grammars keep on leaking, language keeps on changing, and humans keep on varying their behavior. Frustrated by these facts, linguists have begun to question the methodology that commits them to the task of stipulating a fixed set of rules or filters to match a specific set of data. Searching for more dynamic approaches, they have begun to think of language as an emergent behavior.

Hallan (2001, 91) points to the discrepancy between linguistic knowledge gained through introspection and linguistic knowledge gained through usage data. She further warns against the limitations of knowledge through introspection, among which is that it cannot predict what will happen in real language use and that speakers are more easily influenced by prescriptive perceptions during introspection than during real language use (Hallan 2001, 91–92). Furthermore, scholars offer strong arguments against the inevitable prejudice involved when observations about language are impressionistic and anecdotal rather than systematic and based on quantitative data (see Leech et al. 2009, 19, and Mair 2006, 3, among others).

Because introspection then does not supply a reliable source of linguistic knowledge, some linguists have turned to language usage data as the *basis* for linguistic theory, not merely an area of application, giving rise to a number of trends.

One important trend is the rejection of the Chomskyan interpretation of *Universal Grammar* (Bybee 2010b, 201), even if linguistic knowledge still resides in the cognition of individuals (Bybee 2010b, 9). This view leaves room for influence from extralinguistic factors (Bybee and Hopper 2001, 19) and leads to viewing language structure not as a holistic, autonomous system but as something rather more fluid and variable (Bybee and Hopper 2001, 2).

Another trend that flows from this change in perspective is that the heterogeneity of and variation in language are seen as fundamentally part of language (Croft 2001, 7) and not just as an almost irrelevant manifestation of *performance* in the Chomskyan sense. Croft (2001, 364) views language as fundamentally dynamic and interactional, not only accommodating variation but also elevating it to the status quo, moving away from the traditional focus on *competence*.

The last important trend I will point to here is that language change becomes central in language conceptualization, and the distinction between synchronic and diachronic linguistics is much fuzzier. Bybee (2010a, 945) suggests that linguists seek explanations for current linguistic structures in how it came to be because all synchronic states in languages are the result of a long chain of diachronic developments (Bybee 2010a, 945). She further claims that linguistic theory cannot be complete without embracing the contribution of change to how we understand language structure (Bybee 2010b, 119). She states explicitly,

> Language change is not just a peripheral phenomenon that can be tacked on to a synchronic theory; synchrony and diachrony have to be viewed as an integrated whole. Change is both a window into cognitive representations and a creator of linguistic patterns. (Bybee 2010b, 105)

Consequently, a theory of language should focus on the dynamic processes that continually create and re-create language (Bybee 2010b, 1), which will enhance our understanding of the processes and directions of language change as well as the individual's cognitive system for language (Bybee 2010b, 10).

In the description of grammar, a new concept has been developed in the past few decades: *emergent grammar*. Emergent grammar breaks with standard ideas about grammar—it relativizes language structure to speakers' real language experiences (Van Rooy and Kruger 2015, 49) and sees structure as an ongoing reaction on communicative pressures rather than a preexisting matrix (Bybee and Hopper 2001, 3). Grammar has no autonomous existence outside of mental representation and processing, and it is continuously adapted for use (Bybee and Hopper 2001, 2–3). Thompson and Hopper (2001, 48) explain this particularly well:

> We could say, then, that what we think of as grammar is a complex of memories we have of how our speech community has resolved communicative problems. "Grammar" is a name for the adaptive, complex, highly interrelated, and multiple categorized sets of recurrent regularities that arise from doing the communicative work humans do.

Recent studies have shown that the human brain's capacity for long-term memory storage and retrieval is much greater than previously claimed (Bybee 2010b, 15; Pierrehumbert 2001, 140), solving the previously anticipated problem of storing such vast amounts of linguistic knowledge. Bybee (2001, 355) claims that speakers have a great capacity not only for the storage of lexicon but also for a great number of precomposed expressions and combinations. Human cognition also has the capacity for a lot of redundant information (Hare, Ford, and Marslen-Wilson 2001, 196), further reducing the need for abstract structures and systems to explain linguistic knowledge.

Another concept that developed with the help of this new knowledge about the human brain is that of the *exemplar model*. Bybee (2010b, 19) explains that an exemplar, in this context, is built from a set of signs that has been experienced by the speaker and regarded as similar in some way. In the exemplar model, categories in the speaker's memory are represented by a great cloud of recalled signs from a particular category, organized in a cognitive map where similar examples are close to each other and dissimilar examples are farther from each other (Pierrehumbert 2001, 140).

Even though usage frequency is not overtly encoded in the model, it still plays an important role in the cognitive representation of categories (Pierrehumbert 2001, 143). The importance of usage and frequency for the origin and change of a grammar is confirmed by several linguists (e.g., Langacker 2010; Bybee 2010a). Apparently, human cognition is sensitive to frequency, and without instruction or request, it sorts and represents elements based on context-relevant relative frequency (Fenk-Oczlon 2001, 433). High frequency leads to familiarity, enabling a speaker to recall constructions more quickly and identify, recognize, anticipate, and predict it more accurately (Fenk-Oczlon 2001, 432).

For such a usage-based theory, quantitative studies becomes particularly important to understand the breadth of linguistic experience (Bybee 2010b, 12). Every example of usage influences the cognitive representation of language in strengthening the representation of an item, and this strength accumulates through time with repetition (Bybee 2010b, 18). Fenk-Oczlon (2001, 432) points out that frequency does not *directly* affect linguistic structure; rather, it influences cognitive processes of speakers, which are in turn involved in processing language use, and so influences linguistic structure. On these grounds, frequency influences linguistic behavior in a variety of ways (Bybee and Hopper 2001, 10; Deutscher 2005, 261; Hare, Ford, and Marslen-Wilson 2001, 181; Jurafsky et al. 2001, 230; Van Rooy and Kruger 2015, 53).

Convention plays an important role in language usage, and Bybee (2010b, 90) indicates that linguistic convention and grammaticality are established and strengthened through frequency. Furthermore, because high frequency strengthens mental representation and familiarity makes recalling forms easier and quicker, more frequent forms often resist structural change, while less frequent ones change more easily because they are more difficult to recognize and recall (Bybee 2010a, 962; Phillips 2001, 134). On the other hand, sound changes progress faster in forms with high frequency because they proceed incrementally (Phillips 2001, 134).

In this cognitive usage-based framework, language cannot be seen as a monolithic system anymore but rather as a massive collection of heterogeneous *constructions* (Bybee and Hopper 2001, 3). Constructions are the direct link between form and meaning (Croft 2010, 472) and can manifest in terms of syntax, morphology, and lexicon (Croft 2010, 463).

Croft (2010, 463) claims that constructions form the basis of syntactic structure or, rather, the basic units of syntactic representation, based on the representation of grammatical knowledge in the speaker's cognition (Croft 2010, 471). This links with Langacker's (2010, 425) view that any facet of language use or sequence of events in discourse can be abstracted and conventionalized as a unit.

In concluding the theoretical exposition of language as emergent from repeated usage, a language can also be described as a *complex adaptive system* (Bybee 2010b, 2). The same cognitive processes that are responsible for processing language are also the processes that lead to linguistic change (Bybee 2010b, 10). On this point, I move on to engage shortly with diachronic linguistics in order to give the necessary background to discuss theories of language change.

Diachronic Linguistics

Diachronic linguistics is the subdiscipline within linguistics that focuses specifically on the historical development of language and languages. Lass (1997, 17) describes diachronic linguistics as making stories or interpretations of "what happened to language over time." There are two important components to diachronic investigations: first, the description of language over time and, second, the explanation of what happened over time.

Lass (1997, 10) and Joseph (2004, 60) indicate that responsible investigations into the linguistic past should be based on a thorough understanding of the linguistic present; it is based on the assumption that all the conditions in the past were in principle the same as current conditions. Deutscher (2005, 9) puts it as "the present is the key to the past," based on the principle of *uniformitarianism*. Synchronic and diachronic linguistics cannot be sharply distinguished (Lass 1997, 12), and it does not have to be since it forms two complementary dimensions of linguistics (Lass 1997, 14).

As far back as we can currently determine, the principle of "nothing comes from nothing" is relevant (Deutscher 2005, 10). There are extensive theories on how human communication and language originally developed, referring, for example, to concepts such as *joint intentionality* and *collective intentionality* (Tomasello, 2014). We unfortunately don't have any data from that time, and for all the data that we do have, the principle of uniformitarianism holds. Explanations of any current state are, at least partially, found in states that preceded it. Every construction has its own history, even though common trends of change can be identified (Fischer, Norde, and Perridon 2004, 13). In this regard, Lass (1997, 327) reminds us that questions and answers about linguistic change are necessarily highly complex, as they are about historical systems that are the products of long processes of evolution. The types of explanations we would seek for historical questions will, of course, depend on the specific questions.

The diachronic linguist makes as much sense of historical data as possible and uses theoretical insights and appropriate methods of investigation to describe and explain whichever type of linguistic change as thoroughly as possible.

LINGUISTIC CHANGE

The theory of linguistic change as we know it today has its origins in the nineteenth century, and while it brought a great deal of important insights, it also carries a lot of baggage. For exactly this reason, it is important to take Mufwene's (2008, 102) recommendation to heart: "It is not unwarranted to reexamine our working assumptions every now and then, before we stray too far on mistaken paths."

Metalanguage is an obvious part of any field of study, but users of that metalanguage can become so used to it that its importance is underestimated (Lass 1997, 41). On the one hand, the metalanguage of diachronic linguistics and linguistic change is often more metaphorical than linguists realize, and, on the other hand, linguists often do not realize how important these metaphors are as frameworks for thought and investigation in the field (Lass 1997, 41–42). Haspelmath (2004, 24) warns us to be careful with such metaphors—while it can be very useful, we should maintain the distinction between the metaphor (and its boundaries) and the real phenomena it is describing. The distinction should also act as a counter force against linguists getting stuck in the limitations of the metaphor and not looking for explanations beyond it (Lass 1997, 280).

A theory is often a formalization of such a metaphor (Lass 1997, 32), and a theoretical framework is necessary for scientific practice. Vincent and Börjars (2010, 279) further confirm the mutual relationship between theory and data. A theory and its relevant metaphors are formed within a scientific paradigm. Such a paradigm is "an object for further articulation and specification under new or more stringent conditions" (Kuhn 1996, 23) that determines not only the questions that scientists ask but also the instruments they use and where they search for answers (Kuhn 1996, 111). So, to identify something as a historical fact or a linguistic fact or to see it as relevant already entails some interpretation that can happen only with the mediation of a theory (Lass 1997, 17). Any theory, however, has potential shortcomings, such as the theory of creolization, which still carries baggage from an era of racial separation and discrimination (Mufwene 2008, 109). Because of this, it is important to at least remain aware of different theoretical approaches, as this creates more appreciation for and awareness of different possible explanations.

It is not, however, possible to account for all theories of language change in this book, so I selected a number of theories that have enough in common for a reasonable comparison but also enough differences to give a critical perspective on each other.

Linguistic Change as Evolution

Language as Evolving Entity

There is a movement in diachronic linguistics that portrays linguistic change as evolution, with both similarities and differences to biological evolution. To see linguistic change as evolution, we need a description of *language* that fits into that framework. Bybee's (2010b, 196) description of language as a complex adaptive system works well in this regard, but there are other, more specific descriptions that are also useful.

There are different ways to describe language within an evolutionary framework. Lass (1997, xvii) describes languages as "populations of variants moving through time," "bundles of historical accidents, not perfect and predictable machines" (Lass 1990, 81), and, famously, "[l]anguages are imperfectly replicating systems" (Lass 1997, 354). He focuses on language as a system that replicates similarly to a population, and the ongoing imperfect replication of variants is the process that drives language evolution.

Croft, who also describes linguistic change as evolution, defines a language as "the population of utterances in a speech community" (Croft 2000, 26) and a "loosely coordinated set of linguemes" (Croft 2000, 230), where *linguemes* are the units of replication. He also compares a linguistic system to a plant rather than an animal (Croft 2000, 230), but he focuses on language more as a population than as an organism.

Mufwene (2008, 63) describes language as a species, rather than a population, of idiolects and rather than utterances or variants. He emphasizes the importance of viewing languages as products produced by people, continuously adapted for communicative purposes (Mufwene 2008, 211).

The function of language is discussed in several different ways. Labov (2010, 6) mentions the fundamental cognitive function of language to transfer information through temporal and spatial dimensions, and Deutscher (2005, 46) links with that by describing the basic function of language as providing a stable system of conventions that allows coherent communication. The matter is not necessarily as simple as just communication, though, and Mufwene (2008, 253) points to the possible instrumental value of language as a tool used by speakers to help them adapt to changing socioeconomic circumstances. These and other possible functions that language can fulfill should be kept in mind when seeking the causes of language evolution.

The view that, in spite of other possible functions, the basic function of language is communication leads to the assumption that the basic manifestation of language is face-to-face conversation (Mair 2006, 183). Still, everyday communication is littered with inefficiencies, miscommunications, and misunderstandings, showing that language does not always work as well as

we would like (Labov 2010, 21). Labov (2010, 47) specifically mentions that we can observe considerable distance between intention and what we actually accomplish in linguistic interaction. There are furthermore grammatical phenomena in all languages that can be regarded as afunctional (Trudgill 2002, 76). Deutscher (2005, 40) calls these imperfections, irregularities, redundancies, and idiosyncrasies "that mar the picture of a perfect design." So, not only does language not always fulfill its basic function efficiently, but the grammatical structure of any language is also far from what we would call *perfect*.

Exactly here can we search for possible mechanisms of change that are so common in all languages at all times.

How Language Evolution Works

Mufwene (2008, 16) asks directly whether it is useful to talk about language *evolution* and not just *change*. He then explains that the term *evolution* includes more than the traditional use of *change*—it deals not only with structural and pragmatic change but also with speciation and the birth and death of languages (Mufwene 2008, 16). But this is not the only reason to call linguistic change evolution—several linguists draw certain parallels between biological and linguistic evolution.

In his book on language evolution, Croft (2000, 10) claims that language change is similar to biological evolution in some ways. He defines language and speakers in terms of populations: a geographic variety is similar to a geographical race of a species, diverging slightly in terms of structure but not enough to prevent communication (or, in biological terms, crossbreeding) (Croft 2000, 19). For Croft (2000, 172), communicative isolation is more important for language speciation than communicative interaction, and in that way, a society is similar to a biological species.

Mufwene (2008, 11) compares languages to species rather than organisms, and he goes further to compare language species specifically to viruses rather than plant or animal species. He claims that languages are like biological populations where every organism's genotypical and phenotypical individuality is preserved, even if they share certain characteristics with other members of the species (Mufwene 2008, 14).

A further similarity between languages and populations of species that Mufwene (2008, 15) points out is the way in which languages are "born" and "die." Different from organisms, there is no moment of conception with languages, no pregnancy or birth—languages are identified as new languages *post factum*, after a particular variety is viewed as sufficiently different from an earlier population. A language also dies when it has no more speakers,

just as a species dies when there are no more members of it left (Mufwene 2008, 15).

Returning to Mufwene's insistence that languages resemble virus species more than anything else, one of the most important reasons for this is that contact plays a central role in language evolution. Language speciation happens through population movement and contact and is a consequence of competition and selection in different circumstances (Mufwene 2008, 128). These circumstances are called the *ecology of a language* and may include contact and multilingualism (Mufwene 2008, 181) as well as socioeconomic circumstances (Mufwene 2008, 232). Multilingualism can cause linguistic structures to be influenced by other languages used alongside one another, and the socioeconomic ecology can cause language shift or loss.

The comparison between biological and language evolution can be even more specific: both Croft (2000) and Mufwene (2008) identify replicators in the process of language evolution, although their concepts differ from each other.

Croft (2000, 3) claims that the true, existing entities in linguistics are utterances as they are produced in context as well as speakers' linguistic knowledge and mental grammars. The evolution of both these entities, especially utterances, proceeds through replication, not inherent change (Croft 2000, 3). He defines a *replicator* as an entity that passes its structure on mostly unchanged and an *utterance* as a specific, real occurrence of human behavior in communicative interaction as it is pronounced, grammatically structured, and semantically and pragmatically interpreted (Croft 2000, 26).

Croft (2000, 28) refines his theory further with the concept of a *lingueme*— the equivalent of a gene, with the equivalent of alleles being variants of a lingueme. The lingueme pool, similar to a gene pool, is the total set of linguemes in a population of utterances (the language) as well as the grammars of the speakers as a whole (Croft 2000, 28–29). Linguemes can function on different levels of inclusivity, and linguemes with higher levels of inclusivity are more independent as replicators than linguemes with lower levels of inclusivity (Croft 2000, 34).

The concept of a lingueme pool enables Croft's theory to handle language variation efficiently. If variation can be viewed as indicative of a lingueme pool in the population of utterances created by society, then both the heterogeneous nature of language and the structural nature of that heterogeneity are expected (Croft 2000, 172). The lingueme pool can even change quite drastically without causing the language to cease being a genetic descendant of its parent (Croft 2000, 197).

Based on the details of his theory, Croft (2000, 30) calls it the Theory of Utterance Selection for language change. He makes three statements about

this theory. First, it does not exclude the possibility of selection on other levels of language, such as individuals and society; it places the primary locus of change only in utterance selections. Second, it does not presuppose a particular set of causal mechanisms for replication and selection of linguemes. Finally, it places linguistic convention at its center.

Mufwene (2008, 116), on the other hand, identifies the idiolect as replicator in language evolution. He mentions that it does differ from organisms, especially animals, in that variation within one idiolect is possible. He emphasizes the importance of family resemblance in the coexistence of idiolects in a communal language (Mufwene 2008, 66), which is a heterogeneous population of idiolects (Mufwene 2008, 115).

Different from Croft (2000), Mufwene (2008) does not refer to a lingueme pool, which can be expected if utterances are not replicators. Instead, he refers to a *feature pool*—it includes the entirety of linguistic knowledge of speakers, including variants that compete with one another for similar structural or communicative functions (Mufwene 2008, 17). Different members of a speech community, especially those who interact with one another, contribute to each other's feature pool of language (Mufwene 2008, 117). Language evolution takes place on the basis of competition in the feature pool, and features are added to the pool through contact between idiolects (Mufwene 2008, 128). Any linguistic state is thus the consequence of contact between idiolects, and any language is in that sense a hybrid (Mufwene 2008, 132). Language contact and mutual influence, as well as language variation, are accommodated in this framework, and Mufwene (2008, 178) specifically states that the concept of a feature pool makes the boundaries between languages more osmotic than linguists made it out to be in the past.

Processes of Language Evolution

Deutscher (2005, 46) points out a truism in diachronic linguistics—in the history of languages, change is the rule, not the exception. It is, however, not that simple: what is language evolution really? As Traugott and Trousdale (2010, 21) put it, it depends on the theoretical framework within which the changes or evolution is described. For generative linguists, it is changes in the grammar that happen during language acquisition; for functional linguists, structural change is linked to usage patterns, and change occurs throughout the life of speakers (Traugott and Trousdale 2010, 21). The latter perspectives concur with Deutscher's (2005, 261) statement: "Language is a tool that has been worn into shape by continual use."

Language evolution is also not random and chaotic—Croft (2000, 42) and McMahon (1994, 6) point out the regularity in linguistic change, and McMa-

hon claims that it is exactly the arbitrariness of language that causes the non-arbitrariness of linguistic change. Not only is language evolution dependent on language-internal factors, but social, political, and environmental factors also contribute to the process (McMahon 1994, 228–29).

To answer more accurately the question "what is language evolution and change?," I now turn to different processes of change in more detail.

Different Processes of Language Evolution

The first process I'm describing is *erosion*. Erosion is the continuing shortening and loss of sounds in speech, described by Deutscher (2005, 169) as a very useful compacting mechanism that enables speakers to convey ideas faster and more efficiently. Erosion entails sound change (Croft 2000, 163), traditionally not included under grammatical change. Deutscher (2005, 208) points out, however, that it often leads to grammatical changes in conjunction with other language-evolutionary processes (such as grammaticalization, which I will return to later).

The second process of linguistic change is that of *analogy*. Analogical change traditionally referred to morphophonological change, specifically the loss of normalization of alternation within paradigms or the extension of alternations from one paradigm to another (Bybee 2010a, 958). The concept can be applied more broadly as well, where a linguistic phenomenon is restructured to be similar to an already existing phenomenon (Traugott and Trousdale 2010, 36). Bybee (2010b, 57) describes analogy as the process through which a speaker uses a new item in an existing construction and points out that analogy as a kind of change does not function apart from analogy as a cognitive processing mechanism (Bybee 2010b, 72). More frequent, conventionalized phrases typically serve as a template for the formation of new phrases (Bybee 2010b, 63).

A third process of linguistic change is *reanalysis*. Reanalysis is viewed not as a process that brings about particular changes in itself but rather as a mental process that causes grammatical changes together with other processes. Reanalysis is partially syntagmatic, while analogy, for example, is paradigmatic (Traugott and Trousdale 2010, 35). Some linguists claim that reanalysis is involved in all structural or grammatical change (Croft 2000, 140; Traugott and Trousdale 2010, 33), and the latter authors go so far as to claim that "there is no change without reanalysis" (Traugott and Trousdale 2010, 39).

Another process of linguistic change that links up with reanalysis is the handling of dysfunctional entities during language evolution, called *exaptation*. Lass (1990, 81) suggests the model of organic exaptation to deal with this phenomenon in language evolution—when linguistic units cease to serve

their purpose and are then used for a new, different purpose. Rosenbach (2010, 172) concurs that diachronic change exploits "weak points" in language. Linguistic systems are not elegantly symmetrical, and there are often various ways to say the same thing (Croft 2000, 70), pointing to the redundancies in language that can be adapted for new purposes (Lass 1997, 313).

Lass (1997, 315) refers to *linguistic junk*, originating from changes that cause redundancy. Three things may happen to this junk (Lass 1997, 317): (1) it is left as it is, even if not serving any purpose (such as the second negative particle *nie* 'not' in Afrikaans); (2) it disappears from language use; or (3) it gains a new purpose, that is, exaptation (e.g., the Afrikaans attributive *-e*, previously linked to the gender of nouns and the nature of determiners, which was adapted as an attributive suffix after the collapse of the gender system in the development from Dutch; for details, see Lass 1990, 88–91). All three of these strategies are used in language evolution, and the last one refers to exaptation. Lass (1990, 98) further shows that even phenomena that still serve a function can be exapted for other functions; however, not all adjustments made during language evolution are exaptations.

The underlying issue that exaptation can help to elucidate is that linguists have claimed in the past that linguistic change can be attributed to the human mind's aversion to "useless" variation; in truth, however, languages are littered with useless variation (Lass 1997, 344). Exaptation is but one of the ways that speakers deal with redundancy in language. Lass (1997, 350) also points out how ridiculous it is to claim that linguistic change can mend some broken parts in language but cannot prevent parts from breaking in the first place.

The last process of language evolution I attend to in this section is even more complex and comprehensive than the previous ones: *grammaticalization*. Grammaticalization is typically defined as the process through which a lexical item or sequence develops into a grammatical unit or where an already grammatical unit develops into an even more grammatical one.

Bybee (2010b, 106) places the locus of description of grammaticalization in the *construction*, as lexical items grammaticalize within particular constructions, leading to the formation of new constructions (Bybee 2010a, 965). The focus on constructions is also necessary because grammaticalization manifests in phonetics, morphosyntax, semantics, and pragmatics, and not just on *one* of the traditional grammatical levels. McMahon (1994, 161) calls it "the cross-componential change *par excellence*." The elements in constructions that grammaticalize fuses together increasingly, and the internal structure of the construction tends toward reduction (Bybee 2010a, 972). It supports the argument that grammar, specifically syntax, is not autonomous as structuralists and, particularly, generativists claim (Croft 1995, 517–18);

rather, linguistic units, or constructions, form on various levels and can move between different levels (Croft 1995, 521–22).

The process of reanalysis that I described earlier also links to grammaticalization. On the one hand, some scholars, such as Van Gelderen (2010, 129), claim that the micro steps within grammaticalization are really just reanalysis. On the other hand, Bybee (2010a, 974) warns against the reduction of grammaticalization to mere reanalysis because it denies the importance of usage-based factors and encourages viewing grammar as a discrete entity. Bybee (2010a, 973) does agree, though, that reanalysis is certainly a component of grammaticalization.

There are, then, a number of different processes involved in language evolution, but the distinction between them is not always simple or clear-cut. Language evolution is a complex phenomenon with many possible facets and contributing factors. One thing that often creates problems in the description of linguistic changes is the distinction between the *origin* and the *spreading* of a change, which I turn to now.

Distinguishing between the Origin and Spreading of a Linguistic Change

Croft (2000, 185) indicates that language evolution, like any evolutionary change, consists of two steps: (1) the origin of the change, or *altered replication*, or innovation, and (2) the spreading of the change, or *differential replication*, or propagation. In the description of specific linguistic changes, sometimes only one of these steps is taken into consideration. It is pertinent to determine what exactly the difference between the two steps is and how it relates to the description and explanation of linguistic changes.

McMahon (1994, 11) makes a clear distinction between what she calls the issue of inception and the issue of transmission or spreading. She indicates that linguists are often able to explain the spreading of a change but not the origin (McMahon 1994, 225). Lass (1997, 326), for example, claims that the classic actuation problem is not really a problem; he describes the origin of change as mere copying errors (Lass 1997, 112) and focuses on the course of the change.

Croft (2000, 8) says that the mechanism behind innovation is functional (phonetic and conceptual factors) and links innovation with the expression of meaning in grammatical form (Croft 2000, 115). Deutscher (2005, 266) also attributes innovation to local, small-scale concerns, such as reducing effort (as in erosion) or modeling one pattern on another (as in analogy). He claims that innovations are based on a principle of recycling, using existing entities for new functions (Deutscher 2005, 261). Mufwene (2008, 23) completes the picture with his description of innovation within his framework of language

evolution and includes the interaction of external ecology as an influence in causing linguistic change.

Innovation produces variability—a central concept in linguistic change. Poplack, Zents, and Dion (2012b, 251) mention that variation is always a precondition for change but that it should not be confused with change itself. On the contrary, variation *can* be the early stages of linguistic change or not (Trudgill 2002, 41).

Labov (2010, 369) claims that linguistic change does not only link with variation but parasitically depends on it. Poplack, Zents, and Dion (2012b, 251) describe the possibility of a long transition period after initial innovation during which the two variants are in competition—this manifests as variation synchronically. Deutscher (2005, 68) agrees in referring to linguistic change in terms of changes in the frequencies of competing variants.

The spreading (or lack thereof) of a linguistic variant resulting from innovation entails propagation and selection. According to Croft (2000, 166), the mechanism of propagation is social, showing similarities to the patterns of selection in biological populations. Convention is central to propagation, as it is in principle establishing a new convention—the first instances of using an innovation is not exactly conventional yet (Croft 2000, 100).

Mechanisms of selection should be distinguished from mechanisms of unaltered replication—unaltered replication does not include any mechanisms of change, only retaining convention, while selection mechanisms entail differential replication (Croft 2000, 73). Innovation is strengthened through usage to become part of a speaker's mental knowledge (Croft 2000, 186). According to Mufwene (2008, 17) the competition and selection of variants take place in the minds of speakers, although the variants can be obtained from different speakers or even other language systems. Selection often does not eliminate competition completely but determines it in terms of dominance (Mufwene 2008, 19–20). Consequently, different innovations do not spread at the same speed, and all innovations do not carry the same weight with speaker populations (Mufwene 2008, 62). Mufwene (2008, 66) further remarks that many innovations never even spread, as they are not used by any speaker other than the innovator.

Mufwene (2008, 119) returns to the idiolect as replicator to explain selection, although he admits that from a historical perspective, this can be extrapolated toward populations and systems. Interidiolectal variation forces speakers to select between different variants, and it manifests in how speakers accommodate one another in different communicative networks (Mufwene 2008, 126). The selections of individual speakers together form the collective community's selections, returning to the various internal and external ecological factors influencing competition and selection in a community's feature

pool (Mufwene 2008, 183). Mufwene (2008, 90) claims that ecology makes the final decisions on competition, even where variation is merely sustained.

This section focused on so-called natural processes of language evolution, holding for all languages at all times. However, there are what some call *unnatural* linguistic changes, caused primarily by standardization, which I turn to now.

Unnatural Linguistic Change?

Laitinen (2004, 248) claims that when standardized languages (varieties?) are studied, we are studying languages that are at least partially artificial. As Afrikaans has a fully standardized variety, established rather recently (about a century ago), this claim may have important consequences for investigating changes in Afrikaans language use.

Laitinen (2004, 247) urges linguists to take into account how the languages that we investigate have been influenced by standardization and how it affects the specific data under investigation in particular. Also, linguists should ask whose ideological trends or linguistic values are being investigated when analyzing standard languages (Laitinen 2004, 248; Mair 2006, 5). Prescriptive codification often removes emotional or subjective meanings from standard varieties (Laitinen 2004, 249) and, in principle, entails the suppression of optional variation (Mair 2006, 5). This suppression of variation is one of the most important differences between standardization (or planned linguistic change) and unplanned linguistic change, as variation is a natural consequence of the latter.

Mair (2006, 158) reminds us that language standardization has a functional element that is not controversial, even if it is sometimes taken much further than necessary and can lead to bitter social controversies. Standardization further has the consequence that dialect continuums, enabling communication between speakers of bordering dialects, are interrupted (Trudgill 2002, 30).

Regardless of the controversies surrounding standardization, it has an unquestionable influence on linguistic change. Curzan (2014, 8) points out, however, that standardization and prescriptivism are merely formalizations and institutionalizations of a tendency that is common in speech communities all over the world—some speakers, typically speakers of note and/or with authority, tell other speakers how they should and should not speak. She regards standardization and prescriptivism not as unnatural, like Laitinen, but rather as just another factor to take into account when describing and explaining linguistic changes (Curzan 2014, 8). She warns that a binary natural–unnatural distinction is often not useful when reflecting on the relationship between prescriptivism and language history (Curzan 2014, 10). The fact of

standardization—of deliberate, codified change or opposition to change and deliberate variation reduction—should be taken into account when investigating linguistic change (or stability) in a standard variety. It can be relevant for determining the origin of a variant or the spreading (or lack thereof) of that variant.

A related type of change that is sometimes regarded as even more artificial than more typical prescriptivism is guidelines for nonsexist language use (Curzan 2014, 122). In English, the use of singular "gender-neutral" pronouns has a long and controversial history in prescriptivism—the masculine form was prescribed since the eighteenth century; in the 1970s, combination forms, such as *he and she*, were recommended; and since then, several alternatives have been explored (Curzan 2014, 129). The prescription of the masculine form at first and the nonsexist options more recently is also a manifestation of the social embeddedness of language use and an illustration of the extent of success that prescriptions can acquire. Curzan (2014, 120) describes it as a two-way process:

> In other words, the feminist push for nonsexist terms—be those pronouns or nouns—eventually infiltrated the editorial boards that create these kinds of manuals and influenced the prescriptions. Change in prescriptions resulted in more nonsexist language in publications. More nonsexist language in publications changed what people read and saw as acceptable for formal usage, which has the power to influence standards of acceptability for spoken usage as well.

This illustrates the possible extent of prescriptive influence as well as the difference in the extent and/or tempo of acceptance in formal (typically written) and informal (typically spoken) language use. Speakers are more inclined to adhere to prescriptions in formal usage, while it is not necessarily important in informal speech (Laitinen 2004, 250). It can influence the spreading of linguistic change—changes are often established in informal spoken language, from where it spreads to formal written language (Mair 2006, 29). This means that, if a linguist wants to describe linguistic change accurately, this relationship between informal and formal usage should be part of the framework for investigating linguistic change.

Explaining Language Evolution and Specific Changes

The Role of Speakers

The role that speakers play in linguistic change causes some disagreement in linguistic circles. In the generative framework, speakers play little or no role; where grammar is linked to language use, it is linked to speakers, as

in functional and cognitive frameworks. Also, in descriptions of linguistic change as evolution, different scholars do not award speakers the same role.

Lass (1997, xviii) deliberately keeps the system (language) and the speakers separate, especially for the sake of methodological clarity. He explains that he regards speakers not as language builders but rather as end users of a system (Lass 1997, xviii). He claims that the social perspective should at most be complementary to the system perspective (Lass 1997, 286), as linguistic change is, according to him, not the immediate result of human actions (Lass 1997, 337). He further claims that even though communication and meaning are central in language usage, it is not central in extensive structural changes (Lass 1997, 324).

Returning to the linguistic-theoretical framework I explored earlier, the matter cannot be that simple, though. In usage-based theories, linguistic change happens *through* language *use* (Bybee 2010b, 114; Visconti 2004, 186), and in emergent grammar, the issue of frequency is of the utmost importance (Bybee 2010b, 12).

In that case, speakers should be viewed as agents of change. I engage with the more problematic details later, but a number of linguists agree with this in principle. Mufwene (2008, 11) emphasizes the agency of speakers and the communicative actions of speakers in linguistic change. McMahon (1994, 8) says that "speakers change languages," Croft (2000, 4) says that languages do not change themselves but are changed by speakers through their actions, and Deutscher (2005, 61) also places the locus of change in speakers.

Language use further implies interaction, so it is not necessarily only the speaker who is involved but also the listener (Croft 2000, 89; McMahon 1994, 17). McMahon (1994, 17) urges that the interaction between production and perception should not be neglected in explaining linguistic changes. Croft (2000, 94) expands on this by pointing out that personal relationships can have linguistic consequences, as they establish regular speaker–listener interaction between specific groups of people.

Even if speakers are not deliberately changing language (Deutscher 2005, 260; Mufwene 2008, 71), patterns of interaction play an important role in language evolution—speakers adapt (or not) to fellow speakers for a variety of reasons (Mufwene 2008, 64), often without particular regard for the prestige or lack of prestige of varieties or variants (Mufwene 2008, 213). Fellow speakers in a speech community can be viewed as the environment within which communicative interaction proceeds, possibly influencing the ways in which usage is adapted (Croft 2000, 27; Mufwene 2008, 250). The environment can also be more than the rest of the speech community, such as the socioeconomic ecology that may force speakers to adapt in many ways (Mufwene 2008, 231).

Deutscher (2005, 69) mentions an important question: if linguistic change happens in real language use and communicative interaction, how is it that speakers do not become confused by the ongoing changes? The answer lies in the ability of speakers to handle a great deal of variation at any point, and diachronic changes are managed with the same mechanisms by which synchronic variation is managed (Deutscher 2005, 69). Particular diachronic changes are strange for a contemporary speech community only because the variation is not part of their repertoire anymore; at the time of the change, speakers handled it in the same way contemporary speakers handle contemporary variation (Deutscher 2005, 70).

So far, most of the linguists agreed in broad terms on the role of speakers in language evolution. There is one particular point of disagreement, however: should the focus be on individual speakers or on speech communities?

Some scholars tend to favor focusing on speech communities rather than individuals, for example, Croft (2000, 90), Labov (2010, 7), and Poplack, Zents, and Dion (2012b, 248). Poplack, Zents, and Dion (2012b, 248) specifically claim that linguistic change does not happen in the minds of individuals but rather comes from regular interaction in the speech community. Croft (2000, 90) emphasizes the collectivity of language use and the importance of the speech community. This is in line with Croft's focus on linguistic convention I referred to earlier. Labov (2010, 7) also states that the individual is not the unit of linguistic analysis but rather the speech community.

Mufwene (2008, 16), on the other hand, focuses more attention on the individual speaker. While he admits that patterns of evolution can be identified only in the common language use of a speech community, he focuses on individuals because linguistic change is made up of the cumulative changes that individuals make (Mufwene 2008, 32). He also points out that innovations have to be initiated by individuals (Mufwene 2008, 67; see also Schøsler 2010, 212) and that innovations spread through adoption by individual speakers (Mufwene 2008, 61). Other factors, such as population structure, also influence how a change spreads, but it is the behavior of individuals that adds up to the drift in Sapir's terms (Mufwene 2008, 72).

The importance of the individual also lies in the selection of innovations (Mufwene 2008, 116) because linguistic change is an accumulation of similar individual choices—there can be no community-wide selections without individual selections (Mufwene 2008, 173). In tune with the exemplar model, Mufwene (2008, 170) places the representation of language in the mind of the individual speaker. Still, Mufwene (2008, 32) also states that language evolution emerges from cross-idiolectal patterns. So, while the mechanisms of change, especially innovations, are situated in individuals, *language* is conceptualized on a communal level, and changes spread on this level, leading to language evolution.

Another matter of disagreement is about whether linguistic change takes place during the transfer of language from one generation to the next or takes place throughout the lives of speakers. On the one hand, McMahon (1994, 270) finds it useful to draw parallels between children's language acquisition and linguistic change. On the other hand, most of the theoretical positions I have been discussing find important differences between the two processes—basically, language evolves not only or even primarily during acquisition but also through the continuous use of language throughout the lives of speakers. Bybee (2010b, 116) thinks that the similarities between child language acquisition and linguistic change flow from the fact that some of the same cognitive processes are involved. This does not mean that languages do not change during acquisition, just that they do not *primarily* change this way.

Haspelmath (2004, 35) points out an important principle in this regard: if linguistic change happens during language acquisition, it would be difficult to place restrictions on how language changes because it would be much more random than regular. We already know that this is not the case. Furthermore, the idea that languages change during acquisition is also based on the assumption that adults' grammars are stable and invariable (Croft 2000, 56), which has been shown to be problematic by emergent grammar and the exemplar model.

I now move on to another tricky topic: the distinction (or not) between internally motivated and externally motivated linguistic change.

Internally and Externally Motivated Change

Croft (2000, 6) states that a comprehensive framework for understanding language evolution should include both internal and external causes of change. Internally motivated change is currently not particularly problematic in the literature, and the mechanisms of change are generally accepted. Externally motivated change is a topic of more disagreement, though, even if its existence is not questioned as such. Lass (1997, 184–85) refers to the "scars" that all languages have because of external influences, and Fischer, Norde, and Perridon (2004, 10) mention standardization and language contact as examples of external factors in linguistic change.

To deal with all of the assumptions and ideas behind how language contact causes change, Poplack, Zents, and Dion (2012a, 204) compile a list of criteria to determine whether a particular phenomenon is caused by language contact:

1. The phenomenon should represent an actual change.
2. The phenomenon was not present in the variety before contact.

3. The phenomenon is still not present in other contemporary varieties without contact.
4. The phenomenon's behavior is similar to the borrowed equivalent in the source language.
5. The phenomenon differs in important ways from any apparently similar constructions already present in the borrowing language, if there are such constructions.

They mention that even if *one* of these criteria cannot be met, we cannot be certain that the relevant phenomenon is caused by language contact (Poplack, Zents, and Dion 2012b, 252). Because of the prevalence of claims attributing usage phenomena in Afrikaans to contact with English, the criteria may be useful to help determine the validity of these claims. However, it is important to keep in mind that the criteria are relevant only for determining the *origin* of something, not how and why an existing phenomenon is spreading.

Importantly, the criteria and their application are not readily accepted by all. Mufwene (2008, 29–31) finds the structural distinction between internally and externally motivated changes problematic and attributes it to the legacy of nineteenth-century ideologies of linguistic and racial purity—for him, the distinction is more a sociological than a linguistic one.

Mufwene (2008, 31) unifies the two proposed types of linguistic change by claiming that the only real type of language contact is interidiolectal contact. Furthermore, all causes of linguistic change are external to the structure of language and communicative actions of speakers (Mufwene 2008, 31–32). What does happen when a speaker comes into contact with different varieties or languages is that their feature pool gains more features and variation—a quantitative difference more than a qualitative one (Mufwene 2008, 122, 152, 183). The influence of contact can also be more subtle—apart from adding a feature from another language to a speaker's feature pool, covert transfer is also possible, especially between languages with similar constructions (Kruger and Van Rooy 2016, 120). Covert transfer is completely quantitative, where similar constructions from different languages are grouped together in bi- or multilingual speakers' exemplar models, where the cumulative frequency of all the constructions entrenches it to a greater extent than would be the case if the speaker knew only one of the languages (Kruger and Van Rooy 2016, 120).

Mufwene (2008, 161) does not deny some differences between so-called internal and external linguistic changes, but he insists that the differences come from the external ecology of the idiolect, including other speakers but also socioeconomic factors. In terms of mechanisms of change and how language evolution proceeds, the distinction between internally and externally

motivated changes does not add anything of value to our understanding of these phenomena (Mufwene 2008, 161). In summary, the point is not that the internal or external origin of a change is irrelevant but that a change typically *proceeds* in the same way whether its origin is internal or external in which-ever way. If the locus of change is seated in the individual, it makes sense, as the distinction is somewhat artificial on an individual level. Still, Mufwene does not shy away from the role of language contact in language evolution— on the contrary, he claims that contact plays an important role in the histories of all languages (Mufwene 2003, 276). He believes that it is possible to attri-bute the diversification of language across the world to population migration and contact (Mufwene 2008, 31) but still maintains that language contact is in principle always contact between idiolects that adds to the feature pools of the speakers involved (Mufwene 2008, 152).

Croft (2000, 8) argues in a similar manner that the mechanisms of inter-nally and externally motivated change operate similarly, as speakers typically know more than one variety or code. He refers to interference, which happens when speakers know more than one language well enough to make connec-tions between them (Croft 2000, 145). He also discusses *interlingual identi-fication* (Croft 2000, 145), where a speaker links elements from two different linguistic systems to one another, what I have called covert transfer earlier. Croft (2000, 146) further links it to an evolutionary framework, saying that interlingual identification is when speakers cognitively link corresponding linguemes from two languages. This mechanism also functions in language-internal innovations, which he calls *intraference* (Croft 2000, 148). Croft's stronger distinction between internal and external changes compared to that of Mufwene could be attributed to his focus on the speech community and communal language use rather than the individual speaker. While the mecha-nism being the same is useful on individual speaker level, the distinction be-tween internal and external *origins* of a lingueme might be more important on the community level and possible structural linguistic changes. Croft (2000) does not, however, deal with the spreading of a lingueme in terms of contact, only the origin. In this sense, Mufwene's (2008) approach is better equipped to explain not just the origin of an innovation but also the spreading of an innovation or an existing phenomenon.

There might be another matter involved in the differences between Croft's and Mufwene's views on internal and external change. Some elements of how Croft talks about language contact and influence suggest that he might conceptualize languages as self-contained systems of interdependent parts, at least to some extent. He calls influence from another language interference, which has connotations and implications of being unnatural and even undesir-able. Mufwene, in comparison, deliberately distances himself from this way of thinking—he avoids a naturalistic distinction between different languages

(acknowledging that the distinction is not simple or in all ways natural), taking the blending of languages in an individual's cognition into account without painting it as abnormal or undesirable in any way. While the distinction between languages or codes might be useful for practical reasons, especially if we take into account how speakers identify and delineate languages, it is important to keep in mind how this distinction translates (or not) into cognition, into processing and using language(s) by individual speakers.

Unfortunately, the quality of argumentation about the influence of language contact is not always particularly high. Poplack, Zents, and Dion (2012b, 250) point out the common assumption that minority languages adapt in the direction of majority languages. Also, it is often believed that surface similarities can easily be equated to structural similarities (Poplack, Zents, and Dion 2012a, 223), while this is not necessarily the case.

One way to prevent oversimplification of the effects of language contact is to not lose sight of the fact that a particular linguistic change may have more than one cause, so to speak (McMahon 1994, 210).

In an attempt to avoid too easily assuming contact influence, Lass suggests a conservative approach in confirming contact as the cause of linguistic change. He emphasizes that the unmarked form of linguistic parentage is parthenogenic, meaning that the burden of evidence falls on any alternative (Lass 1997, 109). While he readily admits to languages in contact influencing each other, he is careful to assume contact influence without sufficient evidence. Sufficient evidence, however, is not merely the fact of a close contact situation (Lass 1997, 197–98), as being in contact does not justify ascribing it a causal role (Lass 1997, 202). Because internally motivated changes would happen anyway, claiming contact influence should have a heavier burden of proof according to Lass (1997, 209).

Mufwene's more inclusive view of language contact enables him to approach the matter differently. He describes the common way of dealing with possible contact and influence as a consequence of the nineteenth-century myth of "pure" languages that still finds its way into many linguists' views (even if only subconsciously) (Mufwene 2008, 102). According to him, it is exactly this baggage that led to the assumption that there are significant differences between internally and externally motivated changes (Mufwene 2008, 110). To solve this problem, he returns to his point that all language contact comes down to contact between idiolects (Mufwene 2008, 111).

Mufwene's perspective links with my discussion about language and ideology in chapter 3. On the one hand, it is useful to heed the warnings of Poplack, Zents, and Dion and Lass and not compromise the quality of argumentation because something seems obvious. On the other hand, I am careful not to take contact influence as unnatural or even just less natural than

internally motivated change—especially taking Mufwene's framework into account, mutual influence between languages that share some speakers is a perfectly natural and expected consequence of contact.

Forces of Linguistic Change

It is widely accepted that languages are always changing. Among many nonlinguists, there is the additional perception that language is decaying and changing for the worst (Lass 1997, 341), a perception that is encouraged by prescriptivists and language authorities (Deutscher 2005, 73). If we can believe language authorities, it is a miracle that language has not degenerated into the grunts of apes by now (Deutscher 2005, 76). Even linguists held this view for some time, fearing the complete decay of language (Deutscher 2005, 7). While this view might seem misplaced from a linguistic point of view, it remains important to ask why this view has been held for so long and still receives some support. Lass (1997, 341), for example, dismisses this "romantic nonsense" as a result of "over-refined sensibility and lack of linguistic knowledge." Deutscher (2005, 76) has a more sympathetic view, conceding that certain processes of change might seem like structural decay at first glance but showing how it is different underneath the surface.

Deutscher (2005) explores a number of arguments regarding the origin of linguistic changes, or innovation, that have been proposed by different scholars. Time and again, it turns out that each of these arguments on their own—and even all of these arguments put together—don't and can't explain the extent to which languages change. He then concludes that there must be something in language itself, something inherently unstable and variable that causes its unceasing change (Deutscher 2005, 60–61). He then proceeds to investigate what he calls the "forces of creation" of language, driven by different motives, continuing to push and pull at language, causing it to never stay still (Deutscher 2005, 63).

Deutscher (2005, 62) identifies three basic motives and driving forces behind language evolution: *economy*, *expressiveness*, and *analogy*. These three principles, which do not apply only to language and linguistic change, can be described in the following way:

1. *Economy* refers to the tendency to save trouble where we can and is the driving force behind the shortcuts that speakers often take, especially in pronunciation.
2. *Expressiveness* refers to our attempts to strengthen the effect of utterances and extend its meanings.
3. *Analogy* refers to the mind's craving for order and our instinctive need to find regularity in language.

The first principle, economy, includes a whole range of linguistic trouble-saving methods that follow the principle of the least effort—at least as little effort as possible without compromising the transfer of meaning (Deutscher 2005, 88). Sounds can weaken with time and even disappear, which is part of the process of erosion that I discussed earlier. It makes sense in the usage-based theory of language, taking into account that sounds are more likely to weaken and disappear in items with a higher frequency (Phillips 2001, 134).

Moving on to Deutscher's second principle, apparently meaning erodes just as easily in language as sounds (Deutscher 2005, 97). This erosion typically takes place when speakers want to express themselves creatively or enhance meaning. Metaphors are typically associated with poetic language, but it is also enmeshed with everyday language use—speakers are just so used to it that they often do not notice it at all (Deutscher 2005, 120; Lakoff and Johnson 1999, 125). Lakoff and Johnson (1999) explore the terrain of conceptual metaphors and indicate that metaphors are not just one way of speaking but also essential to the way we conceptualize things. The meanings of metaphors change with time, through ongoing use, from concrete to abstract, and the functions of expressions are changed along with it (Deutscher 2005, 119; Lakoff and Johnson 1999, 124). What used to be a living, interesting metaphor, based on the concrete meaning of a word or expression, becomes the mundane, abstract, new meaning of a word or expression (Deutscher 2005, 120; Lakoff and Johnson 1999, 124).

Apparently, metaphors move from concrete to abstract because speakers *need* them to (Deutscher 2005, 128)—the only way we are able to expand expressive possibilities seems to be using concrete terms for abstract concepts (Lakoff and Johnson 1999, 122). This includes shifts such as from body parts to spatial relationships, from physical proximity to possession, from grabbing to understanding (Deutscher 2005, 142). This process does not only cause metaphors to change existing grammatical elements—the ability of metaphors to change content to structure is involved in the original creation of grammatical elements (Deutscher 2005, 143). The process through which this force often manifests is grammaticalization, during which the meaning of words or constructions bleaches and becomes more functional and grammatical in nature.

The third principle, analogy, is the principle that "saves" linguistic structure from the blind forces of erosion and expressiveness (Deutscher 2005, 177). The human mind is always looking for any repeated patterns, also in language, because regularity eases the pressures of dealing with masses of linguistic data. This human tendency provides the necessary regularity that we can observe in linguistic change, even if there are also random elements involved (Deutscher 2005, 72).

A last issue related to forces of linguistic change regards whether these forces are deliberate. Labov (2010, 244), for example, mentions that ideology can hasten or hinder linguistic changes, particularly during standardization. This would be deliberate and can continue for as long as prescriptivism is taken seriously by the speech community.

However, this is not typically the case in language evolution. Croft (2000, 70) says that speakers may have many goals in using language, but linguistic change is typically not one of them. Even if speakers deliberately innovate linguistically, the aim is usually not linguistic change in itself but rather to communicate something or another. There is no reason, then, to assume any teleological forces contributing to language evolution.

SUMMARY REMARKS

At the beginning of the chapter, I first established the appropriate theoretical framework for the analysis and interpretation of grammatical change in this book. I rejected the Chomskyan focus on competence rather than performance and adopted a usage-based model of language and grammar. Grammar is, in this model, *emergent*, with the exemplar model placing usage frequency at the center of grammar and linguistic change.

I then turned to the theory of linguistic change. Focusing on linguistic change as language evolution, I discussed conceptualizations by Croft and Mufwene in particular. Throughout, Mufwene's ideas about language evolution, of languages evolving similarly to viruses, of the idiolect as replicator and the locus of contact, remained useful and relevant. It links effortlessly with the concepts of emergent grammar and exemplar theory as well as the forces of linguistic change I discussed. It also shows internal consistency, dealing with the different aspects of linguistic change in a satisfactory way. Croft's theory, which focuses on languages as populations of utterances, offers a similar wealth of insights, balancing out Mufwene's focus on individual speakers by focusing on the population as the relevant level of description. Where Mufwene's insights are particularly relevant in describing and explaining the mechanisms of change and the forces behind language evolution, Croft's insights are more useful for the description of the course of specific linguistic changes. These insights provide a sound foundation for the description and explanation of particular changes, as I will attend to in chapters 5 to 7.

3

Afrikaans Standardization in Context

LANGUAGE STANDARDIZATION AND IDEOLOGY

There are many ways to approach the concept of *ideology*. It can be something practical, prereflective, structural, or behavioral instead of only the more traditional conceptualization as conscious, deliberate, systematically organized thought—it does not even have to be thought at all (Woolard 1998, 6). Ideology can be and often is conceptualized as derived from, rooted in, reflective of, or reactive to the experiences and interests of a specific social position (Woolard 1998, 6). Even so, it is typically presented as universal, especially by those who benefit from it. It can also be taken further as a direct link to positions of power, whether social, political, or economical; it is seen as ideas, discourses, or practices in service of the fight to attain and/or retain power (Woolard 1998, 7).

Language ideology, more specifically, may include ideas, cultural beliefs, processes of meaning making, implicit evaluations, and explicit commentary on social and linguistic relationships and how it is loaded with moral and political interest (Spitulnik 1998, 164).

Importantly, language ideology is typically not about *just* language (Spitulnik 1998, 164; Woolard 1998, 3). Language ideology supplies and confirms links between language and identity, aesthetics, morality, and epistemology and so involves not only linguistic form and use but also the idea of the person and the social group, of fundamental social institutions, such as religious rituals, the socialization of children, gender relations, the nation state, education, and the law (Woolard 1998, 3). Language ideology is about the construction and legitimization of power, the production of social relations of similarity and difference, and the creation of cultural stereotypes about

types of speakers and social groups (Spitulnik 1998, 164). Like other types of ideology, language ideology also tends to be naturalizing and universalizing. Language ideologies disguise the conditions of their own production as well as their intimate, strategic links to power, interests, and the creation of cultural value (Spitulnik 1998, 164). Even the existence of a language as a discreet entity is always a discursive project rather than an established fact (Woolard 1998, 20).

One important context where language ideology tends to be particularly visible is the standardization of languages, prescriptivism, and language purity. In spite of the often used justification of standardization as establishing a linguistic form that would facilitate effective communication over geographical and social boundaries (Milroy 2001, 534), the standard variety is often used as the symbol of a specific group or community (Langer and Nesse 2012, 611). This creates the opportunity and power to exclude anybody who has not mastered the standard. A standard variety, then, is not just a vehicle for supraregional communication but also a social norm that allows speakers to identify themselves as a member of the educated or, at least, non-lower social class (Langer and Nesse 2012, 612).

One salient consequence of standardization is the development of a consciousness of "correct" or "canonical" language forms among speakers (Curzan 2014, 43; Milroy 2001, 535). In what Milroy calls standard language cultures, most people subscribe to the standard language ideology, which includes, among other things, a strong view on so-called correctness (Milroy 2001, 535; see also Curzan 2014, 42). In this view, if there are two or more variants of a word or construction, only one of them can be correct (Milroy 2001, 353). It is taken for granted, as *common sense*, that certain forms are correct and others are not. For many people in a standard language culture, this is just how it is: we need no *justification* to reject a certain linguistic form other than its being *nonstandard* (Curzan 2014, 30; Milroy 2001, 535). Even when language-internal justifications are given for rejecting linguistic forms, that is typically not the "real" reason for the rejection—the reason is that it is simply common sense, everyone knows it, it is part of the culture to know it, and if you think differently, you are an outsider, *not a participant in the common culture*, and your views can simply be ignored (Milroy 2001, 536). Milroy (2001, 536) warns us that the power of the commonsense argument should not be underestimated—calling it "common sense" implies that further debate is unnecessary: everyone should certainly know that the viewpoint is the correct one (responsible, decent, moral), and those who disagree can surely not be taken seriously because they are eccentric, irresponsible, and possibly even dishonest.

The commonsense attitude is ideologically loaded, even if those who subscribe to it don't see it that way—they truly believe that their negative judgments about people who use language "incorrectly" are purely linguistic judgments, sanctioned by language authorities (Milroy 2001, 536). This ideology requires people to accept that language(s) is not the property of its native speakers—that which speakers learn informally prior to formal education is neither reliable nor fully correct (Milroy 2001, 537). The language authorities who have to provide the necessary guidelines for language use gain status similar to that of high priests; also, the "right" and "wrong" use of language is often presented in a moral light (Milroy 2001, 537) and linked to other values that are important in the culture, like transparency and honesty (Spitulnik 1998, 164; Woolard 1998, 21). A language is seen as a valuable heritage, a cultural possession similar to religion or a legal system, and, in reality, the idealized standard is equated to the language as a whole (Milroy 2001, 538–39). Milroy (2001, 537) summarizes this quite strikingly:

> The canonical form of the language is a precious inheritance that has been built up over the generations, not by the millions of native speakers, but by a select few who have lavished loving care upon it, polishing, refining and enriching it until it has become a fine instrument of expression. . . . It is believed that if the canonical variety is not universally supported and protected, the language will inevitably decline and decay.

This way of thinking directly relates to issues of language purism. The most fundamental assumption for any puristic activity is that languages can be damaged, whether through negligence of speakers or by external influences like language contact (Langer and Nesse 2012, 608). This attitude is maintained in spite of the fact that there does not exist anything like language purity or pure languages (Langer and Nesse 2012, 610). In truth, language purism is *not* about purity of languages in the first place; rather, it is a social commentary on the (alleged) decline in standards—linguistic, moral, educational—in modern society (Langer and Nesse 2012, 611). No purist is ever satisfied with the status quo; on the contrary, purists usually believe there to be a need, urgent even, to intercede, particularly in removing the corrupting and decaying influences from the language (Langer and Nesse 2012, 621). Purists are often speakers longing back to their "golden years" when their own social norms were formed; any deviation from these norms is seen as corruption of the norms, and it typically relates not only to language purism but also to social behavior, music, and fashion (Langer and Nesse 2012, 622). Curzan (2014, 36) points out that nostalgia is a characteristic of all kinds of different human behavior, so it is not strange that it is also involved in prescriptivism.

Apart from the personal side of purism, it also involves sociopolitical interests—influence from "nonthreatening" languages are often not deemed problematic (Langer and Nesse 2012, 612). Puristic initiatives typically focus on influences from *undesirable* languages and varieties (Langer and Nesse 2012, 621–22), often portrayed as cultural or political enemies that pose a "real" threat (Langer and Nesse 2012, 618; Woolard 1998, 21–22). This enables a rather arbitrary measurement of linguistic loans or other influences as *objectionable* on the one hand and *sufficiently integrated to be accepted* on the other (Langer and Nesse 2012, 617). Purists sometimes use scientific arguments to defend their views, but when science is not "on their side," they easily switch to intuitive and emotional arguments (Langer and Nesse 2012, 617). The reification of languages is essential to these types of arguments— languages are awarded organic and emotional characteristics, such as *health*, *illness*, *beauty*, or *warmth*, and the relevant issues are presented as defending someone close to you against foreign and corrupting influences (Langer and Nesse 2012, 617). Furthermore, metaphors such as *youthful*, *growing*, *blossoming*, and *dying* are used to describe languages and occasionally to demonstrate the superiority of a particular language (Langer and Nesse 2012, 622).

The role of purism in standardization has a powerful influence on people's perceptions of what a "proper language" is (Langer and Nesse 2012, 613). Concerns about the so-called *condition* of a language, especially regarding purity, are often really about sociological issues concerning the status of a prestigious language—languages do not disappear as long as they are used, but that is not enough (Langer and Nesse 2012, 616). Language purists want to protect a language that is *worth* protecting, in other words, a language with prestige (Langer and Nesse 2012, 622).

This conceptualization of language and views on language easily links with nationalism. The essence of nationalism since the nineteenth century is defining imagined communities based on conceptual boundaries that are not necessarily in touch with "objective" reality (Blommaert and Verschuren 1998, 189). Language is often very closely linked with national character (Crowley 1990, 40), and in identifying a nation, language is one of the markers of this identity, together with lineage, history, culture, and religion (Blommaert and Verschuren 1998, 192). It implies that nations can be identified on the basis of natural discontinuities, similar to the differentiation between animal species (Blommaert and Verschuren 1998, 192). In the absence of an "own" language, the legitimacy of a separate nation is easily questioned (Blommaert and Verschuren 1998, 192). While language is sometimes held as the only or the most important distinguishing characteristic of a group (Blommaert and Verschuren 1998, 193), it is not necessarily sufficient to gain group membership—if someone has the "wrong" lineage, history, or appearance, they can-

not become members of the group because they cause a break in the "natural" groupings of humanity into nations (Blommaert and Verschuren 1998, 196).

Ethnicity and nationhood are seen as stable and timeless and based on this the rise of nationalism during economic difficulty is seen as unproblematic and even natural (Blommaert and Verschuren 1998, 194). The tension between empire and nations or ethnic groups are portrayed as a natural consequence of the systematic denial of linguistic, cultural, and political rights of said nations/groups, and because language is seen as a natural characteristic of the group, these rights are viewed as holy and inalienable regardless of the uncompromising and radical ways it is claimed (Blommaert and Verschuren 1998, 202). It is simply *natural* that people protest when their own language and culture are threatened, which is why the nation-state, with one language, culture, religion, and history, is seen as the ideal system that will proceed without oppression (Blommaert and Verschuren 1998, 204). Alternative solutions to tension, such as individual multilingualism, cannot be the solution for this intrinsically unstable and dangerous society because conflict that is based on nationalistic feelings is fundamental conflict that cannot be resolved with surface accommodations (Blommaert and Verschuren 1998, 206). An important assumption of this argument is that people who differ too much do not like to live among each other and that successful community building requires as much uniformity as possible (Blommaert and Verschuren 1998, 205).

Language purism, standardization, and standard language cultures imprinted on linguistics in many ways. Because languages are often regarded as only their standard varieties, it is likely that the ideology of standard languages also influenced linguistics and linguistic analyses (Milroy 2001, 530). Linguists often claim that the standard variety is only one of many varieties, but in reality, the standard is not *just* a variety, and this has to be accounted for when investigating language use (Milroy 2001, 543). Linguists, especially theoretical linguists, often rely on but one variety, which implies that any other linguistic forms outside of this variety also fall outside of grammaticality (Milroy 2001, 545). Linguists typically live and work in standard language cultures and sometimes allow the standard variety to be equated to *the language* (Milroy 2001, 577). While it is of course acceptable to investigate specifically a standard variety in a linguistic study, the choice of variety should be motivated explicitly—unfortunately, this is not always done, and linguists just accept that a/the standard variety is appropriate for whatever the research aims and goals are (Milroy 2001, 544).

Linguists were/are not only influenced by ideology (I explore the case of Afrikaans sociolinguistics to some extent in Kirsten 2017) but have often assisted in establishing specific ideologies. This relates particularly to legitimization of the standard variety, during which time other linguistic forms

become increasingly illegitimate, at least in the popular view (Curzan 2014, 101; Milroy 2001, 547; 2012, 582). On the one hand, the codification of the standard is already a form of legitimization (Milroy 2001, 548), but another important aspect of legitimization, often provided by historical linguistics, is the historicization of the standard (Milroy 2012, 582). The historicization of a variety (or, as it is often viewed, a language) requires a long, continuous history, preferably a glorious history coinciding with that of a nation—a respected, legitimate descent and a long family tree (Langer and Nesse 2012, 617; Milroy 2001, 547; 2012, 582). The history of a variety is often presented as *the* history of a language, with the occasional nod in the direction of "the dialects" (Milroy 2001, 548; 2012, 579). It also helps when this history confirms the purity of that language (Langer and Nesse 2012, 617; Milroy 2001, 549), as is the case in the work of some scholars in the Afrikaans philological school (Kotzé and Kirsten 2016, 363). Even specific phenomena can be legitimized through historicization (Milroy 2001, 550).

However, the histories of languages do not always lend themselves to this ideology, in which case erasure is often practiced—the undesirable parts of the history of a language (meaning the parts that do not fit into the ideology) are rejected or simply ignored (Milroy 2012, 573). It causes some people and events to become invisible (Milroy 2012, 575). According to Milroy (2012, 580), the extent to which traditional linguists were willing to go to argue away that which is contrary to their views on the standard is quite astounding.

Another idea that stems from the so-called legitimacy of a language and the view that it is a valuable heritage is that it should be protected against corruption and decay (Milroy 2001, 550). Many influential nineteenth-century linguists were quite vocal on the corruption of language by ignorant speakers and claimed that they were capable of distinguishing between worthy language changes and corruptions (Milroy 2001, 550). This attitude confirmed the standard variety as the dominant, "legal" version of a language, to which the term "correctness" was applied rigidly and sometimes even morally (Milroy 2012, 582). According to Milroy (2001, 550), there is no way to objectively make this distinction, and he calls it "chiefly an elitist theory of social class and, sometimes, a discernable racist/nationalist element also. Purity is purity—whether it is purity of language or purity of race."

Even linguists who investigated specific contact phenomena still sometimes portrayed language mixing as something abnormal, even pathological (Milroy 2012, 578). This flows from the linguist's primary interest in a language as an internally coherently structured system rather than interest in how speakers successfully communicate using all possible linguistic (and nonlinguistic) resources at their disposal (Milroy 2012, 578). However, a linguist's view of language does not take away the fact of standardization

and the influence it might have on language use—it can most certainly be involved in language change and is one of the factors that linguists cannot afford to ignore (Milroy 2001, 535; see also Curzan 2014, 43).

Nationalism, particularly the link between one nation and one language, goes hand in hand with language purism (Milroy 2001, 549; Woolard 1998, 18). It relates to the authenticity and moral importance of the mother tongue as the first, *real* language of a speaker (Woolard 1998, 18). One language has often been linked to one nation in linguistics (Milroy 2001, 549), and the history of a language has typically been associated with the (often glorious) history of a nation (Milroy 2012, 572). In this entanglement of the history of a language with the history of a nation, the history of a language is often portrayed in terms of the inspirational tale of growth and development from obscure, humble origins to a glorious, powerful language of a glorious, powerful nation (Milroy 2012, 579). This is not absent from the writings of the Afrikaans philological school and its descriptions of the origin and development of Afrikaans, which show signs of nationalism and the history of Afrikaans being linked with the history of Afrikaners (Kotzé and Kirsten 2016, 363).

The period of language change I investigate in the later chapters of this book commences just as the standardization of Afrikaans gains some traction. In order to contextualize this, I briefly account the history of Afrikaans standardization in the next section.

SHORT HISTORY OF THE STANDARDIZATION OF AFRIKAANS

The first deliberate attempts to write Afrikaans instead of Dutch came from the nineteenth century (Ponelis 1992, 73). From the 1820s, Afrikaans vernacular texts started appearing in newspapers, mostly as humorous and satirical texts. They were typically published anonymously or under pseudonyms and typically took the form of epistolary commentaries and dialogues that dealt with local matters (Roberge 2003, 22). These texts were, however, typically written by people outside of the speech community portrayed in the texts— the authors were members of the local intelligentsia, and they used linguistic representations that did not necessarily accurately represent Cape Dutch Vernacular—it just had to be plausible to their audiences (Roberge 2003, 22). This tradition did not form part of any deliberate standardization attempts.

Another important Afrikaans writing tradition from the nineteenth century is called Arabic Afrikaans, the tradition of writing Afrikaans in Arabic script. The tradition originated from Muslim schools in the Cape, where Arabic script was taught and used. Ponelis (1992, 73) called this a considerable act

of cultivation, echoed by Davids (1987a, 35). In contrast to the image of Afrikaans cultivation as it has been upheld from an Afrikaner[1] perspective, Davids (1987a, 35) writes,

> What this study shows, is that Afrikaans, in a written form, was seriously and extensively used by the Cape Muslim community long before the white Afrikaners. In 1925 when Afrikaans became one of the official languages in this country, over 32 Arabic-Afrikaans publications were already in circulation. Similarly, when it was first introduced in non-Muslim schools in 1914, it was already the exclusive language of instruction in the Muslim religious schools.

However, this tradition never gained traction outside of the religious sphere where it developed, and it has not been practiced since the mid-twentieth century (Ponelis 1992, 73). More in-depth information about this practice can be found in Davids (1987b, 1990, 2011).

Another writing tradition is that of the GRA, the *Genootskap van Regte Afrikaanders* (Society of True Afrikaners). In their flagship publication, *Di Patriot* (The Patriot), they deliberately distanced themselves from the Dutch spelling system, attempting to establish a distinguishable Afrikaans way of writing (Ponelis 1992, 73). The GRA had a political agenda as well, and they marketed Afrikaans "as a God-given emblem of the Afrikaner people that could be stipulated *a priori* (as opposed to a segment along a continuum of lects)" (Roberge 2003, 26). Their main principle regarding spelling was simple: write as you speak. However, this tradition was also not established in more mainstream Afrikaans writing practice and was moreover arrested by the Second War of Freedom (also called the Anglo-Boer War), which lasted from 1899 to 1902.

During and after the war, Lord Alfred Milner, British high commissioner for South Africa at the time, instituted an anglicization policy (Roberge 2003, 29). Furthermore, in the aftermath of the war, there was large-scale urbanization, with (white) Afrikaners moving into the cities looking for work (Roberge 2003, 29). These factors led to fears of the anglicization of Afrikaans-speaking white people and the deliberate development of Afrikaans as a symbol of Afrikaner unity (Roberge 2003, 29–30).

In contrast to prewar attempts at standardization, the geographical focus of Afrikaans shifted from the Cape region to the north, mainly Johannesburg and Pretoria, which were developing into economic and cultural centers (Deumert 2004, 58). A number of new language societies emerged: *Afrikaanse Taalgenootskap* (Afrikaans Language Society), founded in Transvaal in 1905; *Onse Taal* (Our Language), founded in the Orange Free State in 1906; *Afrikaanse Taalvereniging* (Afrikaans Language Union), founded in the Cape Province in 1906; and, finally, *Zuidafrikaanse Akademie voor Taal, Letteren, en Kunst*

(South African Academy for Language, Literature, and Art), founded in 1909 (Deumert 2004, 58). In 1908, D. F. Malan, then chairman of the *Taalvereniging*, delivered a famous address in Stellenbosch called *Dit is ons ernst* (We Are in Earnest), and he tried to convince his fellow Afrikaners to elevate Afrikaans to a written language, arguing that elevation and elaboration of the vernacular "would allow its speakers to win a whole future for themselves as a nation and ensure the Afrikaner volk a place among the civilized cultures of the world" (Roberge 2003, 30).

After unification in 1910, Afrikaner nationalism developed considerable popular support, and a strong, new language movement emerged (Deumert 2004, 293). In 1911, Melt Brink suggested that the Academy compile standardized Afrikaans spelling rules, but at the time, the simplified Dutch spelling was already widely accepted, and the idea did not lead anywhere (Van Rensburg 2017, 262). A few years later, however, in 1914, the use of Afrikaans as the language of instruction was spreading in primary schools, and the Academy returned to Brink's suggestion (Van Rensburg 2017, 262). They emerged as the central institution of standardization and elaboration (Deumert 2004, 59). The first preliminary spelling guide of the Academy appeared in 1915, and the first official orthography, named *Afrikaanse Woordelijs en Spelreëls* (Afrikaans Wordlist and Spelling Rules), or *AWS* for short, appeared in 1917.

A number of Afrikaans periodicals started to appear from 1910, such as *Die Brandwag* (The Sentinel) in 1910, *Ons Moedertaal* (Our Mother Tongue) in 1914, and *Die Huisgenoot* (The Home Companion) in 1916 (Deumert 2004, 293). Afrikaans became accepted for publication of nonfiction texts and literary texts (especially poetry), and Afrikaans books were published increasingly (Deumert 2004, 293). Afrikaans was introduced in the civil service and some universities in 1918, and it was accepted by the Protestant churches from 1919 (Deumert 2004, 293). Schools and universities, as well as the publication of educational books, such as textbooks and grammars, played an important role in implementing the new norms of Standard Afrikaans (Conradie 2010, 68). Afrikaans was legally recognized as an official language of the Union of South Africa in 1925 and replaced Dutch in its official capacity (Roberge 2003, 31). One of the highest-regarded achievements of this period was the publication of a complete Afrikaans translation of the Christian Bible in 1933.

A less visible but very important trend during the standardization of Afrikaans was what was called *vernederlandsing* (Dutchification), which relied heavily on Standard Dutch to inform decisions regarding Afrikaans spelling and other matters. At the turn of the century, there was a strong movement to modernize and simplify Standard Dutch spelling based on the belief that the

new orthography, *Vereenvoudigde Nederlandse Spelling* (Simplified Dutch Spelling), would make Dutch more accessible for all the speakers of vernacular Afrikaans (Ponelis 1993, 49). Supporters of Dutch believed that English should be challenged with a strong, cultivated language like Standard Dutch rather than a young language like Afrikaans (Ponelis 1993, 49). During the first decade of the twentieth century, the Dutch camp was better organized than the Afrikaans camp and at the time had the support of the Transvaal and Orange Free State education departments as well as that of the University of the Cape of Good Hope (Van Rensburg 2017, 252). Later, as I described above, this situation changed, and eventually Afrikaans was accepted as an official, standard language. However, there was still far reaching influence from Dutch in the initial standardization of Afrikaans.

Conradie (2010, 68–69) indicates that importing Dutch forms into Afrikaans was an important part of the standardization process, and even obviously Dutch forms that differed from the Afrikaans vernacular were prominent in formal language use. Furthermore, there was the issue of specialized terminology, for which Afrikaans was not yet equipped, and the solution was adlexification from Dutch (Roberge 2003, 31). With regard to spelling, even though the principle of *verafrikaansing* (Afrikaansification) was implemented from the third edition of the *AWS*, the consequences of *vernederlandsing* can still be observed in the spelling of Standard Afrikaans (Van Rensburg 2017, 268).

In spite of the influence of Dutch, which was not always visible or apparent for those who lacked formal Dutch education, a strong sentiment of purism developed around Standard Afrikaans. The relationship between Afrikaners and English South Africans fundamentally changed during the Second War of Freedom, and avoiding English loans and Anglicisms was a matter of pride for many Afrikaners (Deumert 2004, 277). I return to these language attitudes and how they continued and developed through the twentieth century in the next section.

One of the most important driving forces behind the cultivation and standardization of Afrikaans was Afrikaner nationalism (Ponelis 1992, 72), and the new standard variety served as a marker of this particular political and social ideal and identity in its contrast with Dutch (Grebe 2012, 115). Consequently, acceptance of the new standard's norms was not an issue for white Afrikaans speaking South Africans (Roberge 2003, 32).

SPEAKER ATTITUDES TOWARD
STANDARD AFRIKAANS IN THE TWENTIETH CENTURY

The attitudes that speakers in general, especially prescriptivists, linguists, and other prominent Afrikaans figures, held about Standard Afrikaans was impor-

tant in creating a particular social climate that surrounded the standard variety in the course of the twentieth century. Nationalist and essentialist ideas about Afrikaans and a strive toward creating and maintaining a pure standard variety were important themes in work focusing on Afrikaans.

Malherbe writes in 1917 (13) already that it is obvious that those who want to learn the cultivated forms will go to those who present good morals and cultivation. According to him, the good author should teach the masses to write correctly, and even the usage of the majority is subject to deliberate cultivation (Malherbe 1917, 13). He further reflects on language purity, saying that people should take care to write Afrikaans *purely* and that foreign idiom should be avoided, as language purification is an important task awaiting speakers of Afrikaans (Malherbe 1917, 16). However, he warns against *overly* extreme attitudes, urging acceptance of words that have been embedded deeply in common usage (Malherbe 1917, 16).

In 1926, Botha and Burger (1926, 202) write about how some people are *guilty* of using Anglicisms and warn readers that if they are not careful, Afrikaans will decline into a sad jumble.

A few years later, Pienaar and Langenhoven (1932, 4) focus on the reverence that speakers should have for the Afrikaans language—their valuable, newfound possession. They emphasize that they live in the formative years of Afrikaans and encourage readers to take responsibility to keep this treasure not only *clean* but also *sparkling* for following generations (Pienaar and Langenhoven 1932, 4–5). Here, we can clearly see the picture of Afrikaans as a cultural possession that should be cherished and protected, and anyone who neglects their responsibilities toward the language is cast in a negative light. They further state explicitly that the extent to which the language is learned prior to formal education is inadequate, particularly because it is taught to children by the *kindermeid* 'children's maid' (a pejorative term reserved for women of color), and they lament that the *bad* is often so deeply ingrained by the time children start school that it is difficult to root out (Pienaar and Langenhoven 1932, 8–9). In a further attempt to justify the disapproval with which they regard language *errors* and those who make them, they directly link careful and precise thought with careful and precise expression (Pienaar and Langenhoven 1932, 25). This justifies discrimination on the grounds of *faulty* language use. The same authors then continue to sing the praises of the purity of the Afrikaans language treasure (Pienaar and Langenhoven 1932, 43) and warn against the use of Anglicisms that adulterate the language (Pienaar and Langenhoven 1932, 84). They also briefly refer to authors who make themselves *un-Afrikaans* through impure language use (Pienaar and Langenhoven 1932, 83).

Another author joins the discourse in accusing the use of Anglicisms of pushing aside so-called Afrikaans culture and leading to its demise (Le Roux

1939, 77). He pathologizes a bilingual speaker's inability to keep the different languages absolutely separate and pure (Le Roux 1939, 78). He attributes the use of Anglicisms partly to a lack of self-worth, which leads to laziness and sloppiness regarding language use, and remarks that mixing languages is typical of poor, underdeveloped, and uneducated types (Le Roux 1939, 79). He suggests that individual speakers not offend fellow Afrikaans speakers by using *uncivilized* language (Le Roux 1939, 31).

Le Roux (1946, 39) concurs with previous authors that it is not the language use of the majority that is the most important factor in linguistic norms but, rather, the language use of the majority of *competent* speakers. He says that he does not only include language people (as he calls them) in the category of competent speakers, but he would also not go to the *slums* for language advice (Le Roux 1946, 39). Here, he links Standard Afrikaans directly to the social elite and distances it explicitly from the lower social classes—this mutually enhances the legitimacy of both the variety and the elite. True to this type of discourse, he uses the image of a tree to explain the nature of language—he justifies the pruning of and care for the tree and claims that it would not violate the *nature* of the language (Le Roux 1946, 37). He talks about language in essentialist ways and warns against usage that goes against the nature of Afrikaans (this nature of Afrikaans can, he claims, be identified easily enough by a *good observer*) (Le Roux 1946, 208). With these claims, he puts the authority to judge good and bad language use squarely in the hands of the elite, a practice that is also prominent in English prescriptivism (Curzan 2014, 33). However, he does not provide any concrete measures to this end.

A decade later, Boshoff (1956, 58) also laments the use of Anglicisms and warns that if Afrikaans speakers use foreign patterns unnecessarily, they betray and assault the character of their own language (Boshoff 1956, 58–59). In principle, he accepts Anglicisms that do not go against the *nature* of Afrikaans and are not rejected by our linguistic intuitions, and he calls the linguistic nature of Afrikaans a fairly *objective* measurement to determine the acceptability of Anglicisms (Boshoff 1956, 60). He further mentions that only the linguistic intuitions of those who speak in a *civilized* manner and have a good feel for and practical mastery of their mother tongue are valid in the judgment of language use (Boshoff 1956, 60). Once again, the authority on good and bad language is placed solely in the hands of the elite, given the power to disapprove of basically anything according to their judgment.

This type of discourse continues throughout the century. Basson et al. (1968, 38) warn against *barbarisms*, particularly English influence, as other cultural languages do not pose a serious threat because of geographical distance. This warning is an excellent example of only *undesirable* languages being portrayed as real threats, as mentioned earlier.

De Villiers (1971, 1) writes in a linguistic book that his intentions with the book are not purely scholarly but also to guide those people who are interested in *good* language use and choosing the *right* forms—because, apparently, people can and do make mistakes in their mother tongue because of poor knowledge, absentmindedness, or confusion (De Villiers 1971, 10).

Van Schalkwyk and Viviers (1977, 3) bemoan *sloppy* pronunciation that leads to vulgarity and may offend more refined speakers. They urge speakers that their aim *must* remain to speak Afrikaans as purely as possible; otherwise, the continued existence of Afrikaans becomes threatened (Van Schalkwyk and Viviers 1977, 43). Their view of language includes some essentialist concepts, and they declare that every language has its own typical idiom and character that differentiate it from other languages (Van Schalkwyk and Viviers 1977, 43).

Van Schalkwyk and Kroes (1979, 160) warn that Afrikaans cannot follow English in the ease with which it accepts influences from other languages because English is a world language that has never really been threatened by other languages.

Botha and Van Aardt (1987, 29) explicitly condemn speakers who view language as *merely* a medium of communication and who do not hold attitudes of concern with and pride in their languages. They also mention that, *luckily*, society is critical of people who hold such a view and do not allow them to hold a high social standing (Botha and Van Aardt 1987, 30).

Raidt (1989, 116) writes in a linguistic textbook that bilingualism and careless language use lead to the use of Anglicisms, which remain part and parcel of Afrikaans in spite of the praiseworthy efforts of teachers.

Grobler (1993, 3) explains that most Afrikaans speakers (or, taking context into account, most *white* speakers) are uncertain about their own language use and are afraid that they might unknowingly use Anglicisms because the Afrikaans-speaking masses do not hold the same ideals or standards of linguistic purity as the establishment.

This establishment that Grobler refers to constitutes a direct link to Afrikaner nationalism and all of the related ideology. Similar to smaller nationalisms in Europe in the nineteenth century, Afrikaner nationalism started as concern of the members of a nondominant ethnic group about their inferior status in both politics and culture (Kriel 2013, 14). Here, as is often the case otherwise, the debates about and attitudes toward language are not about *just* language but also the nation it has come to symbolize (Kriel 2013, 171).

This attitude is made explicit in the foreword of the first edition of the *AWS*—they state proudly that *our nation* has awakened (Le Roux, Malherbe, and Smith 1917, iii). The irresistible urge for expressing the nation's *individuality*, possible only through the mother tongue, is of course a natural

consequence (Le Roux, Malherbe, and Smith 1917, iii). In publishing the *AWS*, they can now provide in the *national needs of the nation* and, in doing so, *heal* their education (Le Roux, Malherbe, and Smith 1917, iv). Everyone is called to help build their language, the only medium through which the nation's *individuality* can be preserved, the carrier of the spiritual independence of the nation (Le Roux, Malherbe, and Smith 1917, vii).

Others express similar sentiments. Malherbe (1917, 9) mentions national self-awareness as one of the two causes leading to a written language, and Bouman and Pienaar (1924, 17) point out that language plays an important role in strengthening national awareness. They refer to groups with a more developed racial and national awareness, speaking *languages* and not mere *dialects*, and they elevate Afrikaans to the same level as the European languages they hold in high regard (Bouman and Pienaar 1924, 18). To emphasize the unity of the nation and the role of Afrikaans in that unity, Bouman and Pienaar (1924, 20–21) discuss the (supposed) unusual uniformity of Afrikaans apart from the *less civilized Colored Afrikaans* (as they refer to it), called a racial dialect. (There are many varieties of Afrikaans, several spoken mainly by speakers of color, but it looks like it is not important to distinguish between these varieties.)

This theme is also carried through the century: Boshoff (1964, 39) claims that excessive linguistic loans cause the nation to become *spiritually indigent*, while Basson et al. (1968, 42) frame what they call *indiscreet borrowing* from other languages in terms of citizens' rights. Grobler (1993, 49) refers to the tight (unbreakable) bond between a nation and its language, even if some nations are somehow not as attached to their mother tongues as he would expect. Apparently, the language use of Afrikaans speakers of color is not actually Afrikaans (I translate directly): colored language has a character, humor, and history of its own, but it remains *family of Afrikaans* (Grobler 1993, 15, emphasis added).

The preceding pages serve as an illustration of the establishment of a tradition where Standard Afrikaans is equated to (the whole of) Afrikaans, where *correct* language use and purity is not negotiable, where Afrikaans and Afrikaner nationalism is deliberately and closely linked, and where varieties and speakers outside of traditional Afrikanerdom (who, as a matter of fact, constitute more than half of the speakers of Afrikaans and have done so since the early stages of its development) are simply ignored apart from the occasional, often patronizing nod. In this tradition, the *ownership* of Afrikaans is placed solely in (white) Afrikaner hands and authority over Afrikaans completely in the hands of the elite.

The original cultivation of Afrikaans was part of efforts to empower white Afrikaners economically and politically (Kriel 2013, 205). Cultivation led to

institutionalization, which led to Afrikaners gaining access to economic and political resources through Afrikaans (Kriel 2013, 205). In spite of more than half of the speakers of Afrikaans *not* being white (Webb and Kriel 2000, 20), the Afrikaans language movements of the nineteenth and twentieth centuries appropriated Afrikaans completely for white people—that which was originally a language of especially the nonelite, the working class, people of color, and white people without formal education became the exclusive domain of the white elite (Webb and Kriel 2000, 22). While language is a well-suited mechanism of exclusion in nationalist ideologies (Kriel 2006, 66–67), speaking Afrikaans was not sufficient for membership of the Afrikaner nation—one had to be *white* in the first place (Webb and Kriel 2000, 39–40).

In the past few decades and especially since full democracy of South Africa in 1994, attitudes toward Standard Afrikaans have started to change. The last few editions of the *AWS*, for example, became increasingly inclusive in its approach to codifying the standard orthography. In the most recent edition, it explicitly states that all the varieties and speakers of Afrikaans are taken into account and that the approach is not only prescriptive but also, to a large extent, descriptive (Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns 2017, v, xi). The word list includes many items from different spoken varieties of Afrikaans, and there is a special section with guidelines for representing vernacular Afrikaans in writing.

There are other indications as well that publishers and normative authorities are taking a more inclusive approach to written Afrikaans. First and second editions of an Afrikaans proverbs and expressions dictionary, *Spreekwoorde en waar hulle vandaan kom* (Proverbs and where they come from), were published recently (Prinsloo 2004, 2009b) and include proverbs and expressions from different Afrikaans varieties and communities. It is stated explicitly that racist proverbs and expressions were avoided to a large extent, and the handful of these that were included are specifically marked as such. For most of the twentieth century, many authors would not have given such matters explicit consideration. Furthermore, a dictionary of Afrikaans regional varieties was also recently published: *Annerlike Afrikaans: Woordeboek van Afrikaanse Kontreitaal* (Different Afrikaans: Dictionary of Afrikaans Regional Language). In the foreword, the author explains that there has been a gap in Afrikaans dictionaries for some time: while sociolinguists readily acknowledge that all varieties of a language (in this case Afrikaans) deserve equal recognition, there aren't many publications available about these varieties (Prinsloo 2009a, 7). He also acknowledges that the media have been propagating and spreading the standard variety of Afrikaans almost exclusively until recently, and the dictionary is an attempt make other varieties more visible and accessible to the general public (Prinsloo 2009a, 7).

Researchers have also been attending to the changing ideas about and approaches to Standard Afrikaans. Kotzé (2014) and Odendaal (2014) discuss matters regarding possible restandardization of Afrikaans, and Kotzé (2011) focuses specifically on democratization and standardization. Le Cordeur (2011) writes about identity and how it relates to the different varieties of Afrikaans as well as issues regarding Afrikaans teaching and speakers of varieties other than Standard Afrikaans (Le Cordeur 2015).

Literary texts in varieties other than Standard Afrikaans have been increasing in popularity as well. This further contributes to a broader profile of the Afrikaans used in published texts, which is very different from most of the twentieth century, which saw almost exclusively Standard Afrikaans in most publications.

IMPLICATIONS FOR STUDYING LANGUAGE CHANGE IN WRITTEN AFRIKAANS

Because of historical attitudes and the legacy of colonialism and Apartheid, the Afrikaans that was preserved in writing was mostly that of white people and, often, educated members of the elite. With a few exceptions, published Afrikaans texts used to be exclusively in the standard variety and, when possible, edited with some scrutiny. This provided ample opportunity for anti-English sentiments and opposition to what was perceived as Anglicisms to influence written Afrikaans as well as the prescriptions and proscriptions in normative grammars and other guides.

In focusing on written Afrikaans since standardization, we have the opportunity to see two opposing forces in action: on the one hand, the high levels of Afrikaans–English bilingualism and the contact-induced changes it can bring about and, on the other, the strong tradition of purism and prescriptive adherence to the standard norms as a possible conservative counterforce.

NOTE

1. The term *Afrikaner* came into use in the late nineteenth century and traditionally referred to white Afrikaans-speaking people in southern Africa. It is the term that the Apartheid government used for white Afrikaans speakers, and more recently, it is limited to those who actively identify with that particular ethnolinguistic identity. When I use it in historical contexts, it refers to the group (white Afrikaans-speaking people) in general regardless of social or political affiliation or identity.

4

Diachronic Corpus Linguistics

INTRODUCTION: METHODOLOGY

In scientific inquiry, it is a matter of course that data require theory and that theory requires data—there is an underlying mutual and self-sustaining relationship between the two (Vincent and Börjars 2010, 279). Theoretical assumptions and presuppositions penetrate every aspect of diachronic investigations, analytically and methodologically (Patten 2014, 110); it leads to certain types of phenomena being investigated and certain types of explanations being sought (Vincent and Börjars 2010, 296).

Extralinguistic factors, such as cultural change or changes in social practice, might be relevant in explaining certain linguistic changes (Baker 2009, 330; Leech et al. 2009, 49). The distinction between linguistic and extralinguistic factors might be useful in some ways, but we should keep the interaction between them in mind throughout (Biber, Conrad, and Reppen 1998, 7; Mair 2006, 38). When investigating linguistic change, there are different kinds of factors to consider in explaining specific changes, linguistic as well as extralinguistic. The method of inquiry should, then, accommodate all the relevant different factors.

THEORETICAL CONTEXTUALIZATION

In the framework of cognitive linguistics and usage-based approaches, using quantitative or, at least, quantifiable data is particularly useful (Bybee 2010b, 12) in contrast to the generative framework (Gisborne and Hollmann 2014, 3). In the past three decades, experimental and observational methods

55

produced great amounts of data that bring into question a number of assumptions: that language is a highly modular system, that people have and need significant innate structures for language acquisition, and that acceptability judgment data are reliable (Gries 2014, 16). Scholars have demonstrated that syntax interacts with phonology, semantics, and nonlinguistic cognition to a significant extent; that the seemingly impoverished linguistic input during acquisition is actually rich in probabilistic structure; and that judgments used to support certain theoretical frameworks are not as undisputed and uncontroversial as some would like to think, requiring more reliable data (Gries 2014, 16). The concept of *probabilism* is fairly important here and naturally involves usage frequency—the cumulative effect of frequency is a driving force behind change, where high and low frequencies have different effects on the behavior and change of linguistic elements (Croft 2000, 3; Deutscher 2005, 261; Leech et al. 2009, 90). This leads to linguistic change being gradual and continuous (Patten 2014, 91).

These gradual changes manifest as synchronic variation (Schøsler 2010, 216), and while the two (change and variation) cannot be equated, we should not disregard the fact of synchronic variation when focusing on diachronic developments (Leech et al. 2009, 210). The ongoing change, the variety and variation that is inherently part of human language, leads to the many ways there often are to express the same meaning. We can assume that there will always be redundancy in any language at any point in time—linguistic systems just aren't elegantly symmetrical to begin with (Croft 2000, 70; McMahon 1994, 50; Reppen, Fitzmaurice, and Biber 2002, vii). Methods to investigate linguistic change should account for all of these factors.

CORPUS LINGUISTICS

What Is Corpus Linguistics?

Corpus linguistics can briefly be described as the study or analysis of language use with electronic corpora (Leech et al. 2009, 24). By itself, corpora cannot tell us anything about language, but with the appropriate software, we can use it for electronic analyses (Evison 2010, 122). A corpus does not necessarily contain any new information about (a) language, but corpus analysis software provide a new perspective on the familiar (Biber, Conrad, and Reppen 1998, 234; Evison 2010, 122). In corpus analysis, the first two steps are typically (1) compiling a word list, which can be sorted according to frequency or alphabetically, and (2) generating concordance lists, containing examples of specific items in context (Evison 2010, 122). The relevant software allow us to search and find any item in the corpus and display the

results on a computer screen. Concordances are particularly useful in sorting the examples according to different variables, enabling the identification of patterns (Evison 2010, 129).

Corpus linguistics is inherently based on frequency—either whether something is attested in a corpus at all or how frequently it occurs, making it a distributional method (Gries 2014, 17). Linguists are interested not in the frequencies as such but, rather, in what the distributional characteristics of an element can tell us in terms of structural, semantic, and pragmatic properties of that element (Gries 2014, 17). Changes in frequency, or quantitative changes, can lead to or indicate qualitative changes in, for example, usage context, semantic nuances, and so on. Words and constructions are not randomly distributed in a corpus, making statistical analyses useful and relevant (Gries 2014, 25; Mair 2011, 6). It is sometimes claimed that corpus linguistics is more empirically oriented than theoretically oriented, where findings are deduced bottom up rather than top down. This is true to the extent that conclusions are based on real usage data, but questions, interpretations, and explanations are still theoretically situated (Leech et al. 2009, 31).

Corpus methods employ the ability of computers to do quick, accurate, and complex analyses, and these quantitative data are then qualitatively and functionally interpreted by the linguist (Biber, Conrad, and Reppen 1998, 233; Conrad 2010, 229; Hunston 2010, 158). Corpus linguistics does not present a goal in itself, but does offer different areas of linguistic inquiry a better method of investigation and is rather a means to an end (McCarthy and O'Keeffe 2010, 7). The basic traits of corpus methods and analyses are as follows:

- It is empirical, analyzing real usage patterns in natural texts
- It uses a large and structured collection of natural texts, called a corpus, as the basis for analysis
- It makes extensive use of computers for analysis, both automatic and interactive techniques
- It relies on both quantitative and qualitative analytical techniques (Biber, Conrad, and Reppen 1998, 4)

An important question arises about whether a corpus can be handled in the same way as a text (Tognini Bonelli 2010, 9). There is general consensus that, in spite of the shared point of origin between a corpus and a text, the way they should be handled differs in many respects (Tognini Bonelli 2010, 19). A corpus is not like a text but "more of the same"—it brings together many different texts and cannot be tied to one unique, coherent communicative event (Tognini Bonelli 2010, 19). The important contribution of a corpus

is the patterns of repetition and frequency, which is more important than the integrity of the text (Tognini Bonelli 2010, 19).

Advantages of Corpus Linguistics

One of the most important advantages of corpus linguistics is that you can analyze a great amount of data in terms of usage patterns while keeping track of different contextual factors (Biber, Conrad, and Reppen 1998, 3–4; Conrad 2010, 234; Tognini Bonelli 2010, 20). Corpus investigations typically have two aims: determining to what extent a certain pattern is used and determining which contextual factors influence its use (Biber, Conrad, and Reppen 1998, 3). Computers give consistent, reliable analyses—they don't change their minds and don't get tired during analysis (Biber, Conrad, and Reppen 1998, 4). There is also an interactive component allowing a human analyst to still make the more difficult judgments while the computer keeps record (Biber, Conrad, and Reppen 1998, 4).

Computers (can) do a number of things that human analysts can't: they can do more than one task at a time; store, read, and change numbers, strings, and records efficiently; sort data quickly and reliably, alphabetically or numerically; remember when things were done; and save links between pieces of information (Scott 2010, 136–38). There are also things that computers can't do, complementing human shortcomings and expectations: they do not know what they are doing, do not prefer one answer to another, and do not complain (Scott 2010, 139). These traits of computers and programs minimize human fallibility and preconceived ideas, while human judgment and discretion can be retained.

One aspect where corpus linguistics gives excellent correctives is the intuitions of linguists (and speakers) that are often inaccurate or even incorrect, typically because people tend to notice and remember unusual phenomena, giving them a distorted view of language use (Biber, Conrad, and Reppen 1998, 3; Bybee and Hopper 2001, 4; Reppen 2010, 31). Corpora are also usually compiled from texts by different speakers/authors, enabling generalizations that are not based on the idiosyncrasies of one speaker (Reppen, Fitzmaurice, and Biber 2002, vii). Identifying patterns in language use is furthermore difficult in casual observation—repetition in natural conversations is short lived, fleeting, and often does not make any lasting impression—so people's intuitions of usage frequency are often inaccurate, especially regarding something that is not visible on the surface of language use (Hunston 2010, 155).

Another pertinent matter that I already touched on is that the *existence* of an item cannot be equated to the *behavior* of that item (Poplack, Zents, and

Dion 2010b, 248), which is why the distribution of items is so important. Different from methods focusing on what is grammatically and semantically possible, corpus linguistics adds a distributional dimension that may also include probabilities (Kennedy 2002, 73). In emergent grammar, these probabilities can lead to insights about grammatical structures in a given language. In any language, we find different structural variants with overlap in meaning, leading to the question of what determines the choice between variants (Biber, Conrad, and Reppen 1998, 76–77). Frequency and distribution are also linked—generally, more frequent phenomena are spread more evenly (over texts/text types/contexts), but the relationship is not absolute, which is why distribution should always be taken into account together with frequency (Gries 2014, 40–41). If an item is used primarily in one domain or text type, we cannot make generalizations that are supposed to apply to language use in general.

Another advantage of corpus linguistics is that it takes authentic, natural language use as data (Mair 2011, 6)—language is then investigated in terms of what is really used, not what is theoretically possible (Gisborne and Hollmann 2014, 1).

Diachronic Corpus Linguistics

Even in other theoretical and methodological frameworks, the diachronic tradition has always relied on "real" linguistic data out of necessity (Gisborne and Hollmann 2014, 2; Trousdale 2014, 137). More recently, though, diachronic corpora have become more widely used in historical linguistics, enabling a number of different types of investigations. Even so, Leech et al. (2009, 50) claim "that frequency evidence is far more important in tracing diachronic change than has generally been acknowledged in the past." This encourages quantitative corpus methods for diachronic linguistics—frequency changes are often the quantitative precursor to qualitative changes in certain domains.

Another reason that corpus methods are appropriate for diachronic studies is that the gradual nature of grammatical change (instead of, say, lexical or phonological change) (Rosenbach 2010, 154) makes it more difficult to observe and can lead to its not being noticed by speakers at all (Leech et al. 2009, 7–8). Quantitative data and analyses help to identify gradual changes and determine its extent more accurately than any speaker or linguist can do on their own (Rosenbach 2010, 171). In cognitive linguistics, frequency is not necessarily just a manifestation of change but can also be an influence on change, being taken into account in explaining certain trends (Mair 2011, 3). The use of corpus data and methods is particularly important in the case

of grammaticalization (Leech et al. 2009, 17–18; Trousdale 2014, 116). It is important to note that there is a distinction between the original innovation and the gradual spread of that innovation—there is often a lag between the first use of an innovation and its appearance in corpus data, and only in this spread are corpus methods of any real value (Mair 2006, 13).

Diachronic studies are not just those investigating developments over the course of several centuries—a speaker's grammar can change in the course of one's lifetime (Patten 2014, 91). This recent and ongoing change is often in prescriptivists' line of fire, and corpus methods are in an excellent position to refine or correct perceptions of ongoing change (Mair 2006, 34). Systematic corpus-based studies show time and time again that current assumptions about ongoing changes are incomplete or partially incorrect and add information about the tempo of change (often overestimated) and the underlying processes that manifest on the surface of language use (Leech et al. 2009, 19).

There are certain parallels between the prescriptive traditions of Afrikaans and of English—even today, prescriptions that are linguistically unfounded are still enforced (Leech et al. 2009, 4), and sometimes, anecdotal observations and judgments are repeated so often that they develop a life of their own, accepted as folk knowledge and rarely if ever questioned, even in scholarly writing (Mair 2006, 18).

It is ironic that the method of investigation most often applied in these contexts—impressionistic comments based on anecdotal evidence—is the least reliable of the possible methods, even when applied by linguistically trained observers (Mair 2006, 15). Personal, unsystematic observations lead to incomplete views, even if authors refrain from emotional or aggressive rhetoric (Mair 2006, 17). From these types of observations, people tend to make three types of errors (Mair 2006, 155):

1. Proposed directional changes that are really long-term stable variation
2. Unnecessary emphasis on a small number of (frivolous) shibboleths
3. Focus on isolated, striking instances of change, leading to a lack in observing and accounting for underlying grammatical changes that lead to long-term developments in the core grammar of a language

Instead of approaching grammatical rules as natural laws, we can approach them as conventions that can be disregarded at times for rhetorical effect and for expressive force, sometimes leading to more extensive changes in convention. Corpus methods enable viewing these "transgressions," or innovations, in context, without necessarily going into matters of "correct" or "incorrect" or "good" or "bad"—its spread can then be the dissemination of an innovation rather than the spread of an "error."

Short-Term Diachronic Comparable Corpus Linguistics

A method that is particularly appropriate for investigating ongoing change is short-term diachronic comparable corpus linguistics (Leech et al. 2009, 24). The term *comparable* applies to two or more corpora that have the same composition but are from different time periods (Leech et al. 2009, 28). This method can, unfortunately, not be applied to a period much longer than a century, as genres or text types are not stable enough over such long periods—some extent of *short-termism* is an inevitable restriction on this method (Leech et al. 2009, 31). In the course of one century, it is unlikely that a grammatical change will start and finish, but we can observe certain trends that point toward considerable change (Mair 2006, 83).

In Mair's (2006) study on changes in English in the twentieth century, he uses corpora compiled from written standard language. While this choice is potentially problematic, Mair (2006, 1) claims that it is justifiable because of the prominence of the standard and its high level of codification. Weerman, Olson, and Cloutier (2013, 353–54) criticize this use of exclusively formal written texts in diachronic studies and encourage the compilation of more stratified corpora. Written language does represent *language* (Lass 1997, 47) but tends to be more conservative than spoken language (Lass 1997, 58). It is important, then, to keep the limitations of a given corpus in mind and not generalize beyond its abilities.

For the sake of compensating for possible corpus limitations, the principle of *bigger is better* is often used. There are instances, though, where it is not that simple, where smaller but carefully compiled specialist corpora can be of more use (Koester 2010, 66; McCarthy and O'Keeffe 2010, 6). There are no commonly accepted guidelines for large and small corpora, but a corpus of written language of fewer than 5 million words is generally regarded as small (Koester 2010, 67). Good investigations can be done with corpora as small as 250,000 words (Koester 2010, 67) if the corpus is compiled appropriately and the phenomenon is frequent enough. Smaller, specialized corpora can be particularly appropriate for investigating frequent grammatical items (McCarthy and O'Keeffe 2010, 6; Reppen 2010, 32) because all the instances of an item can be analyzed in depth, not just a sample (Koester 2010, 66). The most common linguistic items, such as personal pronouns, tense-related phenomena, and prepositions, can be investigated in a corpus with extracts of about 1,000 words, and text types or registers can be represented by ten (or even five) different texts (Koester 2010, 70). When ongoing grammatical change is investigated, smaller, well-structured corpora are preferable above larger, less structured data sets (Leech et al. 2009, 15).

EMPIRICAL DESIGN

Corpora of This Study

When I started with the research for this study, there were no diachronic corpora of written Standard Afrikaans (from the 1910s), which is why I compiled the relevant corpora myself. I investigated and considered a number of issues before data collection commenced. The first relates to the research questions and what type of data I would need to investigate grammatical and pragmatic changes in written Afrikaans from the early nineteenth to the early twentieth century. I worked with four intervals, represented by a decade each: 1911–1920 (corpus #1), 1941–1950 (corpus #2), 1971–1980 (corpus #3), and 2001–2010 (corpus #4). While the corpora used by Mair (2006) and Leech et al. (2009) use data from only one year, I decided on practical grounds to rather work with a decade. Some text types are hard to come by, which made the restriction of one year quite impractical. This leads to the next issue.

I next had to decide which registers or text types to include in the corpora, complicated by possible changes in genre conventions. An important difference compared to the corpora used by Mair (2006) and Leech et al. (2009) is that I included manuscript data (letters and diaries)—it is a small portion of the overall data, unfortunately, but at least unedited, informal language is represented at all (based on the recommendation of Weerman, Olson, and Cloutier 2013, 356). I could not include any spoken data in the corpora, mainly because there are no spoken data for the first two periods and very limited data for the subsequent two. The size of each section (text type) in the corpora were determined largely by the general availability of texts, taken as an indication of the extent to which it represents the written language in general. Two text types that are particularly heterogeneous—fiction and informational texts—are represented by expanded categories with particularly large data sets to account for possible stylistic variation without skewing the data. Broadly, the informational texts are popular nonfiction that does not fall into one of the more specific categories, such as biographical or religious texts, typically extracts from popular magazines, or books about sport, hobbies, or political issues. The distinction I made between news reports and informational texts were based on the traits of the texts themselves—news reports focus on short, dense reporting of events, typically found in newspapers but also occasionally in magazines, while informational texts are typically longer and more expanded and less business-like, focusing more on conveying different kinds of information and not just reporting on events, decisions, and so on. The categories according to text type, and the word count of each category in each of the corpora are as follows:

- Fiction (60,000)
- Popular nonfiction:
  ○ Biographical texts (20,000)
  ○ News reports (20,000)
  ○ Informational texts (60,000)
  ○ Religious texts (20,000)
- Academic texts:
  ○ Humanities (30,000)
  ○ Natural science (30,000)
- Manuscripts (letters and diaries) (21,000)

I took a maximum of about 2,000 words from one text, including at least ten different authors and texts in the smaller categories. The extracts were ended at the first suitable place after 2,000 words, causing some variation in the exact amount of words for each corpus. When working with longer texts, I deliberately took extracts from different parts of the different texts regarding the beginning and end of texts or chapters.

There is one gap in corpus #1 that I could not fill—in the 1910s, natural scientists in South Africa simply did not write in Afrikaans but, rather, in Dutch or English. After an extensive search at several libraries and archives, I found two natural science texts written in Afrikaans, and as an emergency measure, I took just under 5,000 words from each. There is, then, a 20,000-word gap in corpus #1. For this reason, among others, I work with normalized frequencies of occurrences per 100,000 words throughout the study.

I also took the subject of texts into account during corpus compilation—when two or three texts from the initial random selection were about the same or a similar subject, I chose different texts to avoid skewing the data. In corpus #1, however, it is inevitable that more texts converge in terms of topic, as many of the formal texts (primarily humanities) written *in* Afrikaans was *about* Afrikaans as language.

A final consideration was to "be realistic" (Biber, Conrad, and Reppen 1998, 250). Because I compiled the corpora mostly by myself, with the help of temporary student assistants at times, I had to stick to parameters that would be sufficient but realistic. Because of practical considerations, many of the published sources I included come from the library of the North-West University, where I work, but if the selection for a specific category could not be sufficiently diverse, I turned to other libraries (including the National Libraries of South Africa). All of these extracts were manually scanned, put through OCR software (text recognition), and manually edited. I also attained electronic texts for corpus #4 from the Centre of Text Technology

of the North-West University, made available to me with a license agreement and used here with permission. For corpus #1, I also attained already electronically converted copies of letters from Deumert's (2004) *Language Standardization and Language Change*. Manuscript data were obtained from different archives (including the National Archives of South Africa and the North-West University archive). Letters from the last period are not publicly available in archives yet, but the EPOG archive provided me with a large number of business letters in electronic format on the condition that they be anonymized during conversion to the appropriate format. I also received a number of personal letters and diary entries from private collections that I anonymized during electronic conversion to protect the identities and personal information of all parties involved.

Analysis Methods and Tools

Biber, Conrad, and Reppen (1998, 10) indicate that research questions of corpus studies often flow from other types of research. Leech et al. (2009, 18) find their questions in a variety of places, such as current ideas and assumptions about ongoing changes, often from prescriptive sources. These ideas can be used as a starting point for further investigation with corpus methods because they are often not completely unfounded, just not based on systematic data analysis (and so sometimes misplaced).

Because of the strong prescriptive tradition in Afrikaans, it is safe to assume that at least some of the principles will hold for investigating changes in Afrikaans since standardization. However, Leech et al. (2009, 79) also point out that by focusing only on obvious frequency changes, we could miss some more subtle changes that do not manifest in raw frequency. In order to avoid this, I decided to identify certain grammatical *categories* with potential changes from the prescriptive literature and basically analyze everything in those categories to see what emerges. A number of subcategories showed stability and not change, which I do not report here, but I also discovered some changes that are not apparent from raw frequencies and that I would not have found otherwise.

Something that also became clear from these broad analyses, as Leech et al. (2009, 268) also find, is that commentators are sometimes correct in their *identification* of a change but wrong in terms of the extent and cause of the change.

In preparation of the analyses, I investigated two types of sources to determine possible categories to analyze for change. First, I obviously looked at previous research in depth, identifying phenomena that have shown signs of change or variation during the past century but had not been investigated

recently or not systematically investigated with corpus methods. There are many such studies about a specific phenomenon and/or a particular text type—for example, Conradie (1992, 1998) and Van Wyk (2009) investigated changes in the tense system in specific text types; Van der Merwe (1997) and Ponelis (1979, 585) made contradicting claims about the relationship between *hy* 'he' and *dit* 'it,' pointing toward possible changes in the pronominal system; and Vink (1981) discussed the relationship between *van* 'of' and *se* (possessive) ''s,' indicating possible changes in the genitive system.

Second, I also took claims from prescriptive sources into account. There were many that pointed to possible changes embedded in the grammatical system—the use of *was* 'was/were' as an auxiliary in past-tense passives instead of *is* 'is/are' (e.g., Carstens 2018, 401; Müller and Pistor 2011, 744; Scholtz 1990, 13) and the use of *was gewees* 'was/were been' (Combrink and Spies 1986, 180; De Villiers 1983, 218; Müller 2003, 700; Van der Merwe and Ponelis 1991, 262; Van Schoor 1983, 144)—that might point to changes in the temporal system. The use of *hierdie* 'this' and *daardie* 'that' as independent demonstrative pronouns (Basson et al. 1968, 99; Carstens 2018, 78; Müller 2003, 101–2) and the "unnecessary" use of *-self* with reflexive pronouns (Carstens 2018, 75; Van der Merwe and Ponelis 1991, 197; Van Schoor 1983, 299) might be indications of variation or change in the pronominal system.

After I identified the possible starting points from the literature, I did some preliminary investigation into the different grammatical categories, and the corpus data showed sufficient signs of change that they were taken as the points of focus of this study.

Regarding the practical side of the analyses, I used the corpus analysis software WordSmith to compile word lists of every corpus. For all the analyses, I identified a word or group of words according to which I extracted concordances—for most instances, I could use one or a few lexical items, for example, *sal* 'shall' and *gaan* 'go' for future reference, *hulle* 'they' for third-person plural, *van* 'of' and *se* ''s' for the genitive, and so on. I also used the word lists to identify possible variants of items for all of the corpora but especially for corpus #1 from the time of initial standardization, when spelling rules were not yet established or implemented consistently. For example, in corpus #1, variants of *sal* 'shall' are *zal*, *zullen*, and *zult*; for *hy* 'he' there is also *hij*; for *het* 'have' there is also *'t*; and for *gaan* 'go' there is also *gaat*. These concordances I then analyzed manually. There are no reliable annotation software available for Afrikaans data, especially not historical data, which is why all the analyses were necessarily done by hand.

Whenever I encountered changes in frequency, I performed statistical tests to determine the significance of these changes. I particularly relied on

log-likelihood tests, using Rayson's (2018) log-likelihood calculator, which has been developed specifically for corpus data. This test indicates whether a change in frequency can be attributed to chance or is statistically significant. A result of less than 3.84 is insignificant ($p > 0.05$), between 3.84 and 6.63 indicates a low level of significance (p is between 0.01 and 0.05), and more than 6.63 is significant ($p < 0.01$) (Rayson 2018). This calculator automatically takes effect size into account.

In one section, I also modeled the data of alternating variants, the genitive *van* and *se*, with decision trees and logistic regression using the computer program R. These models were then used to determine the influence of several variables in the choice between the alternating variants. I discuss the necessary details in the section where I use these models.

CONCLUDING REMARKS

Corpus linguistics consists of research methods that enable linguists to describe language as it is actually used, which is particularly appropriate in a cognitive, usage-based framework. In this study, I use it to investigate language use in specific decades during the past century, tracing changes and developments of a number of phenomena. I further use statistical tests and models to either test for the significance of changes or determine how certain variables influence a choice between alternating variants. Finally, I include contextual factors in the explanations I propose for the course of certain changes or developments.

5

Paradigmatic Changes

PARADIGMS AND PARADIGMATIC CHANGE IN LANGUAGE

Large parts of languages and linguistic structures are paradigmatic in nature, especially where there are productive processes for the formation of morphemes, words, sentences, and so on. The Afrikaans verbal paradigm, for example, is largely regular, where the past tense is typically formed with the auxiliary *het* and the prefix *ge-* on the main verb. There are exceptions, of course: some verbs do not take *ge-* in the past tense, and passive constructions take a different auxiliary for the past tense. There is also a small subset of verbs that has preterit forms instead of using the regular past-tense paradigm. In certain varieties of Afrikaans, the subsets of irregular verbs are even smaller than in written Standard Afrikaans or possibly completely absent (Coetzee 2001, 23).

Different languages and varieties differ regarding paradigms and paradigmatic structure at the syntactic as well as the morphological level (Heltoft 2010, 11). Language change, especially grammatical change, might involve structuring and restructuring certain paradigms in a particular language or variety (Heltoft 2010, 11). If one would focus on only a singular process or instance that shows change, it might be misleading when it is part of a wider paradigmatic change, and this context is left out of the description and explanation of the change (Heltoft 2010, 12).

One process of linguistic change that is particularly relevant when thinking of paradigms is analogy. Analogy relates primarily to the link between sound and meaning and is typically viewed as a type of domestic device that, through regularization, cleans up at least *some* of the mess that sound changes cause (McMahon 1994, 70). The concept of regularization brought about by

analogy is central to Deutscher's (2005, 173) view of how language change works, as it is one of the consequences of speakers' spontaneous attempts at seeking order in the chaotic world around them. The ability to recognize patterns is not only critical in learning an additional language but also indispensable during young children's acquisition of their mother tongue(s) (Deutscher 2005, 174).

Traditionally speaking, there are two types of analogy: analogical leveling and analogical extension. Leveling entails the loss of alternation in a paradigm, and extension is when a particular alternation is transferred to a paradigm where it was previously absent (Bybee 2010b, 66). Both extension and leveling occur when a construction is applied to phenomena that previously participated in a different construction (Bybee 2010b, 69).

In analogical leveling, a paradigm loses alternation and becomes regularized (Bybee 2010a, 958). The regular form does not necessarily completely replace the old one, however (Bybee 2010a, 959), which is partly why we have interchangeable variants of certain words, for example, the plural of *doktor* 'doctor' (holding a PhD), which is *doktors* or *doktore*, where *dokter* 'doctor' (medical) only has *dokters* 'doctors.' One variant can be older than the other, where the newer one originated through analogy, using a common form of pluralization. According to Bybee (2010a, 959), leveling spreads on a lexical level, which is why irregularities are typically present in constructions with a high usage frequency that resist structural changes like these.

Analogical extension is when a paradigm that previously had no alternation develops alternation or where a paradigm shifts from one alternation to another (Bybee 2010a, 962). An Afrikaans example is the verb *bring* 'bring,' which took the vowel *breng* in early twentieth-century written language, developing alternation between *breng* and *bring*, and finally shifting to only *bring* (Kirsten 2013, 68–69). It is commonly accepted that extension is much less common than leveling (Bybee 2010a, 963).

There are several instances of paradigmatic change in Afrikaans in the course of the past century, the details of which I turn to now.

PRETERIT LOSS

The preterit can be described as a synthetic past tense where the past tense is indicated through inflection on the verb, such as *would* as the preterit of *will* or *did* as the preterit of *do*. In many languages, there is a distinction between the present tense, the preterit, and the perfect, such as in English with present tense *do*, preterit *did*, and perfect *have done*. In more analytical languages, such as Afrikaans, which has only one present tense and one past tense, finer

temporal distinctions often depend (at least partially) on other means than inflection, like adverbials, chronological order, and context (Conradie 1998, 41).

In the development of Afrikaans from Dutch, deflection was widespread, especially regarding the verbal system. The deflection caused the preterit to disappear almost completely, making way for the Dutch perfect as the new catchall past tense. Already in 1902, Du Toit (1902, 24) indicated that the "onfolmaak ferlede tyde" (literally 'incomplete past tense') had mostly disappeared—with the exception of a number of modals, the copula *wees* 'be,' and to a small extent the verbs *het* 'have' and *weet* 'know' (Conradie 1998, 37; 2006, 87; Ponelis 1979, 269). Conradie (1999, 20) gives an account of the remaining Afrikaans preterits with an indication of the extent to which it is still used:

Contemporary usage:

- Copula and auxiliary verb: *wees* 'be' with present *is* 'is/are' and preterit *was* 'was/were'
- Modal auxiliaries: *sal* 'shall' with preterit *sou* 'should,' *moet* 'must' with *moes* as preterit, *kan* 'can' with *kon* 'could,' and *wil* 'will' with *wou* 'would'
- Dubitative verb: *dink* 'think' with *dog/dag* 'thought'

Obsolescent forms:

- Main and auxiliary verb: *het* 'has/have' with *had* 'had' as preterit
- Modal auxiliary: *mag* 'may' with *mog* 'might' as preterit
- Main verb: *weet* 'know' with *wis* 'knew' as preterit

In the verbal paradigm of Afrikaans, the remainder of the preterit has been replaced by the Dutch perfect (used as a full-fledged past tense) and the historical present, which has been functionally extended to perform some of the earlier functions of the preterit (Conradie 1999, 21–22). The preterit is still fully functional in contemporary Dutch (Abraham 1999, 12), for example, but many other languages have been experiencing preterit loss, like Southern German and Yiddish in the West Germanic language family as well as Northern Italian, Hungarian, Polish, Czech, Russian, Ukrainian, and Slavonic (Abraham 1999, 13). The catalyst for preterit loss is often that the perfect develops into a more general past tense (Abraham 1999, 14). This is exactly what happened in Afrikaans, and there are additional developments that aided preterit loss in Afrikaans to an even greater extent. These additional developments include regularization of the Dutch verbs *hebben* 'have' and *zijn* 'be'

to only *het* and *is*, the regularization of the past participle to *ge-* (sometimes optional) + stem, and the functional extension of *het* to replace *is* as the past-tense auxiliary used with mutative verbs in active constructions (Conradie 1999, 22).

The process of preterit loss in the development of Afrikaans and all the possible contributing factors are explored in depth in Conradie (1999). He focuses primarily on the eighteenth and especially the nineteenth century, exploring the factors that contributed to early preterit loss in the formation of Afrikaans. He gives only a few brief remarks on preterit use in the 20th century, and other authors such as De Villiers (1971) and Ponelis (1979) do not focus on historical developments when discussing Afrikaans preterits.

OBSOLESCENT PRETERITS

In the previous section, I already indicated three obsolescent preterit forms: *had* 'had,' the preterit of auxiliary and main verb *het* 'has/have'; *mog* 'might,' the preterit of the modal auxiliary *mag* 'may'; and *wis* 'knew,' the preterit of the verb *weet* 'know.'

The first of these to be addressed is *had*. Early in the twentieth century, the use of *had* was already quite rare and declining even further (Kirsten 2013, 69–70). However, even in the fairly small corpora used in this study, there are examples of use throughout the century. The frequencies of use per 100,000 words are as follows:

- Corpus #1: 74 per 100,000 (47 main verb, 27 auxiliary)
- Corpus #2: 6 per 100,000 (mostly main verb, fewer than 1 per 100,000 auxiliary uses)
- Corpus #3: 2 per 100,000 (all main verb)
- Corpus #4: fewer than 1 per 100,000 (all main verb)

Compared to the form *het*, which is consistently used between 1,600 and 1,800 times per 100,000 words in each of the corpora, *had* and the related form *hadden* is already quite rare in corpus #1, contributing a mere 4 percent of the total uses of *het*-forms. Subsequently, its usage drops even further. This corresponds to the pattern of phenomena with (relatively) lower frequency being more susceptible to paradigmatic change than those with high frequency, in this case to a further decrease in usage frequency. In contrast with *het*, the preterit *had* is used more frequently as a main verb than as an auxiliary in corpus #1 and corpus #2, and in corpus #3 and corpus #4, it is used only as a main verb, but its occurrence is very infrequent.

The next obsolescent preterit on the list is the modal auxiliary *mog* 'might,' which is the preterit of *mag* 'may.' The frequencies of *mog* in the corpus data are as follows:

- Corpus #1: 9 per 100,000 words
- Corpus #2: 2 per 100,000 words
- Corpus #3: 0
- Corpus #4: fewer than 1 per 100,000 words

In contrast, the lowest rate of usage for the present tense form *mag* is 38 per 100,000 words in corpus #3 and corpus #4. An important difference between *had* and *mog* is that *had* has regular past-tense variants: *het* for auxiliary and *gehad* for main verb usage; *mog* does not have such a regular past-tense variant. This means that the present-tense form *mag* is appropriated for past-tense usage, but rather consistently only once or twice per 100,000 words of corpus data. Compared to the present-tense usage, then, *mag* is not used very frequently in the past tense; however, only in corpus #1 is *mog* used more frequently in the past tense than *mag*. This indicates that *mag* took over the functional load of *mog* in its absence, once again showing how a lower frequency can contribute to paradigmatic change.

The last obsolescent preterit to be discussed is *wis* 'knew,' the preterit of *weet* 'know.' The regular past-tense *geweet* has replaced the preterit to a large extent, the details of which are given in table 5.1.

Table 5.1.   Frequencies of *weet*, *wis*, and *geweet* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
weet	110	91	106	113
wis	6	3	0	1
geweet	10	16	16	21

From table 5.1, it is clear that *geweet* has taken over the functional load of *wis* almost completely, and it is the more frequent of the two in corpus #1 already.

All three of the obsolescent preterits show the same pattern: they are already somewhat low in frequency compared to other Afrikaans preterits (see the next section), and in the course of the century, their frequencies drop even further. The functional loads of the preterits are shifted onto regular past-tense variants or the present-tense variant. Next up are the much more frequently used modal auxiliaries with fully productive preterits.

Modal Auxiliaries

Before I investigate the use of the preterit of Afrikaans modal auxiliaries, I would like to give a brief overview of how I conceptualize modality in this section. As the focus of this section is not on modality but, rather, on the Afrikaans preterit, which happens to involve modal auxiliaries, I will not go into modality in any depth.

Van der Auwera and Plungian (1998, 80) define modality as those semantic domains that concern possibility and necessity. Within this framework, there are four domains: (1) participant-internal modality, which refers to ability and necessity internal to the participant; (2) participant-external modality, which refers to circumstances outside the participant, concerning possibility and necessity; (3) deontic modality, a subdomain of participant-external modality, which refers to permission or obligation outside of the participant; and (4) epistemic modality, which refers to the speaker's judgment on the likelihood of something (Van der Auwera and Plungian 1998, 80–81). Two additional concepts that are often regarded as within the purview of modality are volitive modality and evidential modality (Van der Auwera and Plungian 1998, 84). While Van der Auwera and Plungian (1998) do not accept these concepts within their deliberately restricted definition of modality, they admit that they are at least closely linked to the more central domains of modality above. For the purpose of this section, I will use modality as an umbrella term including also more peripheral types of (possible) modality.

There is a link between modality, tense, and aspect, which can be seen in, for example, Heine's (2003, 594) grammaticalization paths, and Patard (2014, 69) notes that the past tense can convey modal meaning in many languages. While imperfectivity would more often be associated with modality, the past tenses in Germanic languages that are used to convey modality are usually aspectually neutral (Patard 2014, 70). Patard (2014, 87) links the modal uses of past tenses with pragmatic inferences that become conventionalized to some extent. In many cases, the original temporal and aspectual reference is still possible during the transitional phase from temporal to modal meanings (Patard 2014, 88). In Afrikaans, for example, *sou* 'should' is in a transitional phase, as it is still occasionally used to indicate past future tense, although it is more often used with purely modal meaning in both the past and the present tense in the data. It is probably moving in a similar direction as English *should* (see Rossouw and Van Rooy 2012, 8), from a temporal to a modal meaning.

The modal auxiliary *sal* 'shall' indicates intention or prediction, usually with a future-tense implicature, and its preterit is *sou* 'should.' *Sou* is used in two main categories: one is to indicate the past future tense (5.1), and the

other is to indicate modality (5.2). The modal meanings of *sou* include hypotheticals and epistemic possibility, and it can also indicate intention in the past tense:

5.1. Hij <u>sou</u> nie die laaste wees deur die Hollands-Afrikaanse volk gehou, in sijn bestaan bedreigd deur die magtige Albion. (corpus #1, Religious text)
he <u>would</u> not the last be by the Hollands-Afrikaans nation held in his existence threatened by the mighty Albion
'He would not be the last one held by the Hollands-Afrikaans nation, in his existence threatened by the mighty Albion.'

5.2. Dit sou beter wees as hy hier sou gebly het om vir ons te help. (corpus #2, Fiction)
it <u>would</u> better be if he here <u>would</u> stayed have PTCL.INF for us to help
'It <u>would</u> have been better if he <u>would</u> have stayed here to help us.'

The inflected Dutch forms *zou* and *zouden* still occur a handful of times in corpus #1, but I regard them as equivalents for the purpose of the analyses as they are after all preterits.

A further distinction of usage contexts regards the tense of the sentence—while the modality itself is not bound by tense, the assertion or event can be presented in terms of the past or the present, as exemplified in examples 5.3 and 5.4:

5.3. 'n Paar jaar gelede <u>sou</u> niemand dit gewaag het nie. (corpus #3, Biographical text)
a couple years ago <u>would</u> nobody it dared have PPPTCL.NEG
'A few years ago, nobody <u>would</u> have dared.'

5.4. As jy regtig verder met hom wil werk (berading) <u>sou</u> ek voorstel dat jy hier begin. (corpus #4, Manuscript)
if you.SG really further with him want work (counselling) <u>would</u> I suggest that you.SG here start
'If you really want to work with him further (counselling) I <u>would</u> suggest that you start here.'

Table 5.2 gives the usage frequencies of the use of *sou* as (1) past tense, (2) modality in the past tense, and (3) modality in the present tense.

One clear change emerging from the table is the noticeable decline in overall frequency from corpus #2 on. A log-likelihood test shows that, while

Table 5.2.   Frequencies of *sou* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
past tense	19	31	19	20
modality in past tense	61	76	67	52
modality in present tense	76	59	36	29
TOTAL	156	166	122	101

the increase between corpus #1 and corpus #2 is not significant (0.86), the decrease from corpus #2 to corpus #4 is indeed (42.97).

Another trend presents itself: in corpus #1, modal *sou* is used in the present tense more frequently than in the past tense, but in corpus #2, the two tenses swap places. In corpus #3, the usage in the past tense increases even more and remains stable in corpus #4. While the most obvious change is the overall decline in usage of *sou*, a further change is the increasing tendency to rather use modal *sou* in the past tense than in the present tense.

The modal auxiliary *wil* 'will,' which indicates intention or desire in Afrikaans, takes the preterit form *wou* 'would.' In contrast with *sou*, as well as English *will* and *would*, both *wil* and *wou* are ever used only with modal meaning and not temporal reference. Both *wil* and *wou* are used for participant-oriented modality, and *wou* is not often used with other modal auxiliaries apart from *sou*. It is further used only in the past tense, illustrated in example 5.5:

> 5.5.   Ek <u>wou</u> hom die laaste ent uittrek, maar hy het byna histeries
>   geskree: "Moenie aan my vat nie!" (corpus #4, Fiction)
>   *I <u>wanted</u> him the last stretch out.pull but he have almost hysteri-*
>   *cally shouted must.not on me touch* PTCL.NEG
>   'I <u>wanted to</u> pull him out the last stretch, but he shouted almost
>   hysterically: "Don't touch me!"'

The frequencies of *sou* per 100,000 words in the corpora are as follows:

- Corpus #1: 42 per 100,000 words
- Corpus #2: 43 per 100,000 words
- Corpus #3: 41 per 100,000 words
- Corpus #4: 33 per 100,000 words

There is a slight decrease in usage of *wou* from corpus #3 to corpus #4, but a log-likelihood test shows it to be statistically insignificant (2.58). Thus, there seems to be no changes occurring with regard to *wou*.

The modal auxiliary *kan* 'can' conveys both participant-internal ability and participant-external possibility. The preterit *kon* 'could' places the ability or possibility in the past (5.6), except when a wish or (unrealistic) desire is being expressed (5.7) or if it is used in combination with *sou* (5.8):

5.6. Hy <u>kon</u> ook ure aaneen die ou Joodse geskrifte lees en met Josef daaroor gesels. (corpus #4, Religious text)
he <u>could</u> also hours on.end the old Jewish texts read and with Josef there.over discuss
'He <u>could</u> read old Jewish texts for hours on end and discuss it with Josef.'

5.7. Ek wens jy <u>kon</u> bietjie in my kop inklim en kyk wat gaan aan. (corpus #4, Manuscript)
I wish you.SG <u>could</u> little in my head in.climb and look what go on
'I wish you <u>could</u> climb into my head for a bit and see what's going on.'

5.8. En veral met John *sou* ek <u>kon</u> uitgaan—hy is immers verlief op jou! (corpus #2, Fiction)
and especially with John would I <u>could</u> out.go he be.PRS indeed in.love on you.SG
'And especially with John I <u>could</u> go out—he is after all in love with you!'

If example 5.7 would be reformulated with *kan*, the difference in meaning becomes clear where *kon* indicates a wish for an unrealistic or impossible matter, while *kan* indicates a wish for a real possibility. The original association of temporal distance between *kan* and *kon* is extended here to epistemic distance. The usage of *kon* in the data is as follows:

- Corpus #1: 123 per 100,000 words
- Corpus #2: 130 per 100,000 words
- Corpus #3: 110 per 100,000 words
- Corpus #4: 125 per 100,000 words

It seems that *kon* also shows no definite change.

The modal auxiliary *moet* 'must' indicates both deontic and epistemic modality, with the preterit *moes*, which is used solely in the past tense (5.9). While *moet* merges with *nie* 'not' in the negative to form *moenie*, *moes* remains separate as *moes nie* 'must.PST not' in the negative:

5.9. Die Transvaal wil knoei: hulle sê ons <u>moes</u> die Engelse nie gehelp *het* nie (corpus #1, Biographical text)

> *the Transvaal want tamper they say we* <u>must</u>.PST *the English not*
> *helped* have PTCL.NEG
> 'The Transvaal want to tamper: they say we <u>should</u> not have
> helped the English.'

The usage of *moes* is as follows:

- Corpus #1: 115 per 100,000 words
- Corpus #2: 108 per 100,000 words
- Corpus #3: 71 per 100,000 words
- Corpus #4: 87 per 100,000 words

The usage of *moes* is variable throughout the century but does not suggest continued change.

Regarding the preterits of modal auxiliaries—all of them either show no statistically significant change or are slightly on the decline in the course of the century.

The Verb *wees*

The copula *wees* 'be' can be realized in several different ways in Afrikaans language use—the infinitive *wees* 'be,' the most frequent present tense form *is* 'is/are,' the preterit *was* 'was/were,' and the regular past-tense *gewees* 'been.' Both *is* and *was* can combine with *gewees*, and *gewees* can also be used with modal auxiliaries and *het*. In corpus #1, *was* still occurs as a past-tense auxiliary with mutative verbs, as *het* had not yet replaced *is* and *was* completely in those contexts. The frequencies of *was* as auxiliary verb and copula are as follows:

- Corpus #1: 839 per 100,000 words
- Corpus #2: 581 per 100,000 words
- Corpus #3: 547 per 100,000 words
- Corpus #4: 547 per 100,000 words

It shows that while the use of *was* decreases quite drastically from corpus #1 to corpus #2, it remains fairly stable from corpus #2 on. A small part of the decline from corpus #1 to corpus #2 can be attributed to the loss of function as past-tense auxiliary, but that does not account for the rest of the decline. However, the usage remains stable afterward, indicating that even if there was a significant change early in the century, it did not continue to the second half of the century.

Trends in the Use of the Preterit in Afrikaans

It would seem that the verbal paradigm of Afrikaans shifted ever so slightly in terms of the use of the preterit, against the background of the former Dutch perfect as the new general past tense.

The verbs with a still productive use of the preterit in the early twentieth century (corpus #1) went one of two ways in the course of the century: those with particularly high frequencies retained the use of the preterit, such as *wou*, *sou*, *kon*, and *moes*, while those with lower frequencies lost the productive use of the preterit, such as *wis > geweet* and *mog > mag*. The number of verbs with a preterit form declined, causing the paradigm of verbal past-tense formation to become more regular than before.

There is one anomaly regarding frequency and regularization: the verb *het* 'have.' As a main verb, it has the infinitive *hê* and the partially regularized past-tense *gehad*, while the preterit *had* as a main and auxiliary verb has become obsolescent. As the second most frequent verb in written Afrikaans, would it not exactly have retained the preterit, like the case is with *wees* 'be,' especially since it is one of only two verbs that retained a separate infinitive form? A possible explanation for its preterit loss despite the odds can be found in the course according to which *het* developed from Dutch to Afrikaans: the Afrikaans form *het* originated from dialectal Dutch, particularly Hollands, which employed *het* as singular and *hewwe* as plural, rather than the more extended Standard Dutch paradigm with *heb*, *hebben*, *hebt*, and *heeft* (Conradie 2006, 89; Ponelis 1993, 386). The singular *het* then persisted into Afrikaans, while *hewwe* developed into the infinitive *hê* (Conradie 2006, 89; Ponelis 1993, 386). This means that *het* would have already started to regularize during the earlier development of Afrikaans, gaining particular momentum during the nineteenth century (Conradie 2006, 89–90). In this process, the preterit *had* lost its foothold during the widespread regularization of the verbal system despite the high frequency of *het*.

PRONOMINAL CHANGES

Third-Person Plural

The third-person plural pronoun in Afrikaans is *hulle* 'they.' Different from the second-person plural *julle* 'you.PL,' it did not develop from a similar, existing construction in Dutch. It probably formed in analogy to the second-person *julle*, using the Dutch possessive *hun* 'their' and adding *-lieden* or *-lui* 'people': *hunlui* developed into *hullie* and then *hulle* (Scholtz 1980, 66). Apart from the full form *hulle*, there is also a reduced variant *hul*, and,

as with all plural pronouns in Afrikaans, it has the same form whether used
in the subject (5.10) or object (5.11) position or as an attributive possessive
pronoun (5.12):

5.10.   <u>Hulle</u> het in tente gewoon en was geduring op soek na
        weivelde en drinkgate. (corpus #4, Humanities)
        <u>they</u> have in tents resided and be.PST constantly on search for
        pastures and drink.holes
        '<u>They</u> resided in tents and were constantly searching for pas-
        tures and water holes.'

5.11.   Hy het <u>hulle</u> op die plaas agtergelaat terwyl hy na Grahamstad
        gereis het om sir Henry Pottinger te gaan spreek. (corpus #3,
        Biographical text)
        he have <u>them</u> on the farm behind.left while he to Grahamstown
        traveled have PTCL.INF sir Henry Pottinger to go speak
        'He left <u>them</u> behind on the farm while he traveled to Graham-
        stown to go and speak to sir Henry Pottinger.'

5.12.   En <u>hulle</u> *werk* sal wis en seker vrugte dra vir ons volk. (corpus
        #1, Informational text)
        and <u>their</u> work shall true and sure fruit carry for our nation
        'And <u>their</u> work will truly and surely bear fruit for our nation.'

Other, less frequent uses of *hulle* include reference to inanimate objects, as a
generic pronoun, as a reflexive pronoun, and in the associative construction
(Den Besten 2012a), illustrated in example 5.13, where it is added to a proper
name or form of address to indicate a group with reference to a member of
the group:

5.13.   Dit lyk my <u>oom Koot-hulle</u> is ewe handig. (corpus #4, Bio-
        graphical text)
        it look me <u>uncle Koot-them</u> be.PRS equally handy
        'It looks to me like <u>uncle Koot and [his associates]</u> are equally
        handy.'

The issue at hand in this section regards the increased specialization of the re-
duced form *hul*, albeit only in written language. The two largest categories of
use for the full form *hulle* is the general third-person plural reference (referring
to humans and other animates and inanimate objects) and its use as attributive
possessive pronoun. The usage frequencies are indicated in table 5.3.

Table 5.3.   Frequencies of *hulle* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
third person	378	472	452	436
possessive	157	111	88	87

Regarding the reduced form *hul*, the usage frequencies for these two meanings are indicated in table 5.4.

Table 5.4.   Frequencies of *hul* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
third person	114	6	2	2
possessive	86	136	107	136

If the data of the two tables are put together, a trend becomes clear: the full form *hulle* is proportionally being used increasingly for anaphoric third-person reference, while the reduced form *hul* is being used increasingly as attributive possessive pronoun. In corpus #1, the functional load is still spread more evenly: both the full and the reduced forms are used more frequently for anaphoric reference, with the numbers of the full form being higher than that of the reduced form. From corpus #2 on, however, the reduced form is used quite infrequently for anaphoric third-person reference, but it is used more often as a possessive pronoun than the full form. Where 65 percent of the possessives are expressed with *hulle* in corpus #1, only 45 percent of them are expressed with *hulle* in corpus #2 and corpus #3, and in corpus #4, it drops even further to 39 percent.

In written Afrikaans, then, the subsection of the pronominal paradigm that is encompassed by the third-person plural forms shows increasing specialization of the two variants. By the early twenty-first century, the reduced form *hul* is used mostly as attributive possessive pronouns, while the full form *hulle* is used with many different meanings but mainly for third-person plural anaphoric reference.

Reflexive Pronouns

Ponelis (1979, 86) and De Stadler (1989, 414) regard reflexive pronouns not as a separate class of pronouns but, rather, as a specific use of the object form of personal pronouns. Whether classified separately or not, pronouns used

reflexively in Afrikaans have two forms: the bare object form (5.14) and the extended form with *-self* (5.15):

5.14. Ons mag <u>ons</u> voorstel dat die selle na twee kante toe funksioneer. (corpus #2, Natural sciences)
we may <u>us</u> imagine that the cells to two sides to function
'We may imagine for <u>ourselves</u> that the cells function to two sides.'

5.15. Daar word soms gesê dat 'n mens <u>jouself</u> tuis kan uitlewe. (corpus #3, Religious text)
there become.AUX.PASS.PRS sometimes said that a human <u>yourself</u> home can out.live
'It is sometimes said that you can live out your life at home.'

The extended form with *-self* is slightly controversial—while it is always deemed acceptable when particular emphasis is necessary or to disambiguate a statement (as in 5.16), its use in cases where it is technically unnecessary is dismissed by some (e.g., Carstens 2018, 75) and attributed to English influence (Carstens 2018, 75; Ponelis 1979, 88):

5.16. Arboreta leer ook sommer gou vir Silvie hoe om <u>haarself</u> soos 'n boom te laat lyk. (corpus #4, Fiction)
Arboreta teach also just quickly for Silvie how PTCL.INF <u>herself</u> like a tree to let look
'Arboreta also quickly teaches Silvie how to let <u>herself</u> look like a tree.'

Negative judgments of the so-called unnecessary *-self* form continue throughout the twentieth century and into the early twenty-first (Basson et al. 1968, 71; Botha and Burger 1926, 68; Carstens 1989, 82; Malherbe 1917, 81; Müller 2003, 189; Prinsloo and Odendal 1995, 170; Van der Merwe and Ponelis 1982, 169). In fact, the weight of negativity of the judgments continue to increase.

In order to determine the extent of use of reflexive pronouns, bare and with *-self,* I put all the personal pronouns used reflexively together. There are three categories of usage: first, the bare forms; second, the instances with *-self,* which would be regarded as 'correct' because it indicates emphasis or disambiguation; and, third, the instances with *-self,* where its inclusion is not absolutely necessary. I give the frequencies of the different categories in table 5.5.

The overall usage of reflexive pronouns (or reflexive uses of pronouns) decline for the first three corpora, after which it increases a bit between corpus

Table 5.5. Frequencies of reflexive pronouns per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
bare object form	145	94	48	52
-self obligatory	11	19	11	19
-self unnecessary	6	11	7	13
TOTAL	162	124	67	84

#3 and corpus #4. While the frequencies of the *-self* form uses are not necessarily meaningful in themselves, the proportions are more revealing. Both categories of *-self* forms increase proportionately throughout the century, and the "incorrect" usage goes from 4 percent of the overall frequency to 15 percent. It would seem, then, that the increasing concern of prescriptivists is based on proportional increase, if not numerical increase.

The next question is whether this proportional increase can be attributed to English influence. Ponelis (1993, 287) reports the use of the extended form instead of the bare form in letters from the late seventeenth century as well as sporadic use in the eighteenth—before British settlement and contact with English. However, he emphasizes that it was still a recessive variant, rising to prominence only in the twentieth century (Ponelis 1993, 288). The obligatory use of *-self* in English developed since Middle English, before which the paradigm was more similar to Afrikaans (Smith 2004, 589).

Since the extended forms have been used in Afrikaans since its origin, it cannot be attributed to overt transfer from English. However, current usage patterns are increasingly conforming to the English paradigm, which could be attributed to covert transfer from English. The fact that English followed a similar developmental trajectory than the one that Afrikaans is undergoing now leads to the question of whether Afrikaans would really need English influence for this development. As I have mentioned earlier, this type of question is incredibly difficult to answer for Afrikaans. Since there are no varieties of Afrikaans spoken today that have not been in long-term, continuous contact with English, we have no way of determining what this usage would look like without English influence. Still, the increasing similarity with the English paradigm is notable, and taking the high levels of bilingualism in the speech community into account, some covert transfer from English is likely.

Third-Person Inanimate

The third-person inanimate pronoun in Afrikaans is *dit* 'it,' and it can be used in many different ways. It can refer to inanimate referents that are singular, plural, or unspecified for number (5.17); it can refer anaphorically to the (or

a) previous sentence (5.18); it can refer cataphorically to the (or a) sentence or clause that is to follow (5.19); and it is often used as a dummy subject (5.20). Another possible use of *dit* is in demonstrative constructions with a preposition (5.21), but it is proscribed in the normative literature and fairly infrequent in written language, where the preferred variant is a compound with *daar-*, such as *daaraan* rather than *aan dit* (in 5.21):

5.17. Pryse wissel gedurig en alhoewel <u>dit</u> soms onvoorspelbaar styg moet vir stygings toegelaat word waar moontlik. (corpus #3, Humanities)
prices change constantly and although <u>it</u> sometimes unpredictably rise must for risings allowed become.AUX.PASS.PRS where possible
'Prices change constantly, and although <u>it</u> rises unpredictably sometimes, there should be made provision for increases where possible.'

5.18. Die kinders moet van kleins af voel dat hulle ouers in hulle belangstel. <u>Dit</u> wek vertroue. (corpus #2, Religious text)
the children must from small on feel that their parents in them interested <u>it</u> wake trust
'The children should feel that their parents are interested in them from a young age. <u>It</u> creates trust.'

5.19. Die doel maak nog glad nie 'n slegte middel goed nie, en <u>dit</u> is net 'n dwaas wat die slegste gereedskap uitsoek as hy een of ander werk wil verrig. (corpus #1, Humanities)
the goal make still all not a bad means good PTCL.NEG and <u>it</u> be.PRS just a fool that the worst tools out.search if he one or other work want done
'The goal does not make a bad means good, and <u>it</u> is only a fool that will select the worst tools if he wants to do some or other job.'

5.20. <u>Dit</u> begin reën en die water loop in slierte teen my gesig en in my nek af. (corpus #3, Fiction)
<u>it</u> start rain and the water run in trails against my face and in my neck down
'<u>It</u> starts to rain, and the water runs down my face and neck in streaks.'

5.21. . . . nadat jy gevra het om iets persoonliks agter te laat by jou sodat jy <u>aan dit</u> kan vashou. (corpus #4, Manuscript)
after you.SG asked have PTCL.INF something personal behind to leave by you.SG so.that you.SG <u>on it</u> can fast.hold
'. . . after you asked that I leave behind something personal with you, so you can hold <u>on</u> to <u>it</u>.'

The frequencies of the four most frequent uses of *dit* are indicated in table 5.6.

Two trends emerge from the table. The first is a decline, in broad strokes, of the use of *dit* as dummy subject. Ponelis (1979, 74) points out that the dummy subject *dit* can be omitted through inversion in certain contexts; for example, *Dit is vandag my verjaarsdag* (literally, 'It is today my birthday') can also be expressed by *Vandag is my verjaarsdag* 'Today is my birthday'; or *Dit suis in my ore* (literally, 'It whistles in my ears') can also be *My ore suis* 'My ears whistle.' The sentences without *dit* are shorter and could be increasingly the preferred choice because of its economy.

Table 5.6.   Frequencies of *dit* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
third-person inanimate	355	340	391	311
dummy subject	102	52	61	44
sentence anaphoric	293	265	323	358
sentence cataphoric	313	317	274	253

The second trend is the swap in relationship between sentence-anaphoric and sentence-cataphoric uses. In corpus #1 and corpus #2, cataphoric uses are more frequent, and in corpus #3 and corpus #4, anaphoric uses become more frequent. The use of *dit* for inanimate third-person reference and the use of most third-person pronouns are mostly anaphoric rather than cataphoric. The sentence-anaphoric use then fits into the broad paradigm more easily than the sentence-cataphoric use. This could be a case of analogical normalization, with increasing uniformity within a specific paradigm.

THE LOSS OF THE DUTCH MORPHOLOGICAL GENITIVE

In early Modern Afrikaans, there was still some limited use of the Dutch morphological genitive: the *s*-genitive and genitive articles. In one way, the *s*-genitive was used more freely than in contemporary Dutch—both before (5.22) and after (5.23) the noun:

5.22.   Ja dit was alleen <u>Gods reddende hand</u> wat my weer laat leef, want ik was afgewerk en klaar voor die siekte. (corpus #1, Manuscript)
yes it be.PST alone <u>God.s saving hand</u> that me again let live because I be.PST down.worked and done before the sickness

> 'Yes, it was only <u>God's saving hand</u> that let me live again be-
> cause I was downtrodden and done with the sickness.'

> 5.23.   Die oordeel begin by die <u>huis Gods</u>. (corpus #1, Religious text)
>   *the judgment start at the <u>house God.s</u>*
>   'The judgment starts at <u>God's house</u>.'

After corpus #1, I could only find a handful of instances of the *s*-genitive
in the data, most of which are in religious texts or contexts. Unfortunately,
there is a practical issue making the identification of *s*-genitives in the data
quite difficult. One of the most productive plural suffixes in Afrikaans is -*s*,
which means that sorting the word lists of the corpora in reverse order gives
one a very long list of words ending with -*s*, very few of which are actually
instances of the morphological genitive. Finding the instances of *s*-genitives
by hand was then necessary because the corpus data are not annotated, but
with the large amounts of data I had to sift through, it is unfortunately quite
likely that I missed some instances. I used the ones that I did manage to pin-
point to direct my search further, and it did help me find other instances that
I would not have seen otherwise. However, I cannot claim to have found all
the instances of the *s*-genitive, and the rest of the section is based on what I
did manage to identify from the data.

First, the *s*-genitive is not the only remnant of the Dutch genitive in the
data. In corpus #1, there are some instances of genitive articles to indicate
possession (5.24–5.25):

> 5.24.   Terwijl nou ook die wetenskap ongetwijfeld in die waarheid
>   <u>der</u> dinge haar oorsprong moet soek. . . (corpus #1, Humani-
>   ties)
>   *while now also the science undoubtedly in the truth <u>of things</u>*
>   *her origin must seek*
>   'While science should also seek her origins in the truth <u>of</u>
>   things . . .'

> 5.25.   Aan die kruis het Hy as Soon <u>des</u> Mensen die doodsbeker
>   uitgedrink ter oorwinning. (corpus #1, Religious text)
>   *on the cross have He as Son <u>of</u> Humans the death.cup empty.*
>   *drank for victory*
>   'On the cross, he emptied the cup of death for victory as Son
>   <u>of</u> Man.'

In the later corpora, these remain in a few fixed expressions or in construc-
tions with some open slots (5.26–5.28):

> 5.26.   "Smyt weg jou roer of jy is 'n <u>kind des doods</u>," beveel 'n stem
>   kil. (corpus #3, Fiction)

throw away your.SG gun or you.SG be.PRS a child PTCL.GEN death
order a voice chilly
'"Throw away your gun or you will be <u>dead</u>," a cold voice or-
ders.'

5.27. Dit was <u>in der waarheid</u> geen drama meer nie. (corpus #2, In-
formational text)
it be.PST <u>in PTCL.GEN truth</u> no drama anymore PTCL.NEG
'<u>In truth</u>, it was no drama anymore.'

5.28. Ons sal <u>des te gelukker</u> wees, my hartjie. (corpus #2, Fiction)
we shall <u>PTCL.GEN to happier</u> be my heart.DIM
'We will be <u>all the happier</u>, my sweetheart.'

In corpus #1, there are also a few examples of the genitive *-s* used together
with *des* (5.29):

5.29. Ons sien hier hoe dit sal staan met die Kerk, as die Bruide-
gom kom; as die geroepene ingaan tot die avondmaal van die
bruilof <u>des Lams</u>. (corpus #1, Religious text)
we see here how it shall stand with the Church when the Bride-
groom come when the called.ones in.go to the evening.meal of
the wedding PTCL.GEN <u>Lamb.s</u>
'We see here how it will be with the Church, when the Bride-
groom comes; when the called ones go in to have the commu-
nion of the wedding <u>of the Lamb</u>.'

The remnants of the Dutch genitive found in the data are summarized in table
5.7.
   From the table, it becomes clear that the Dutch genitive was still used
somewhat productively in corpus #1, after which the uses are only in certain
fixed expressions or constructions, becoming very few by corpus #4. An
important observation is that, even in corpus #1, most of the instances are
used in a religious context. The *s*-genitive is used sixty-two times with *Gods*
'God's' and several other instances in religious contexts, such as *Kristus'*
kerk 'Christ's church,' *bruilof des Lams* 'wedding of the Lamb's,' and *saak*
des gebeds 'matter of prayer's.' The articles *der* and *des* are used more freely
in other contexts, such as *Senaat der Unie van Zuid-Afrika* 'Senate of Union
of South Africa' and *de erns des tijds* 'the seriousness of the times.' In cor-
pus #2 and corpus #3, all the identified instances are in religious contexts,
with one exception: *kole vuurs op my hoof* 'coal fire's on my head,' used in
fiction in corpus #2. However, the expression has a biblical origin, linking it
with a religious context once again. In corpus #2 and corpus #3, *der* and *des*
are typically used in fixed expressions and formulaic language, examples of

Table 5.7.   Frequencies of Dutch genitives per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
-s after possessum	11	0	1	0
-s before possessum	65	12	8	0
der (article)	18	6	9	2
des (article)	17	7	10	1
des with *-s*	6	0	0	0
TOTAL	117	25	28	3

which include *in der ewigheid* 'in all eternity' and *des te meer/vuriger/gelukkiger* 'all the more/fiercer/happier' and so on, and the three uses in corpus #4 are two instances of *jare der jare* 'years and years' and one of *des te meer* 'all the more.'

At the time of corpus #4, the *s*-genitive is too infrequently used to be in the corpus data at all, even in religious contexts where it was still used to a limited extent in corpus #2 and corpus #3. The articles *des* and *der* are used exclusively in fixed expressions and formulaic language in corpus #4, apart from surnames such as *Van der Merwe*, which indicates that language users do not need to understand the original grammatical functions of the articles or how to use them productively anymore.

In the course of a century, the Dutch genitive went from productive, if limited, use to almost absent apart from fixed expressions. This means that a whole paradigm of genitive marking all but disappeared in Afrikaans language use, and only the new genitive particle *se* and the continuation of Dutch *van* remains. The grammaticalization of *se* and its relationship to *van* are described in depth in the next chapter.

CONCLUDING REMARKS: PARADIGMATIC CHANGES

In this chapter, I reported on a number of (mostly subtle) paradigmatic changes in the Afrikaans linguistic system. First, the Afrikaans verbal system continued to regularize. While the more frequent modal auxiliaries and the copula *wees* retained their inflected preterit forms, the preterit forms of the infrequent modal auxiliaries and the verb *het* declined visibly in the course of the century, sometimes to the point of being absent from the corpus data.

Regarding pronouns, the reduced form *hul* of the full third-person plural pronoun *hulle* 'they' is increasingly used as an attributive possessive pronoun, while its other uses declined, and the inanimate pronoun *dit* 'it' is used increasingly for sentence-anaphoric reference rather than sentence-cataphoric reference, conforming to its other patterns to a greater extent. When personal

pronouns are used reflexively, they tend to be increasingly used in the extended -*self* form even when it would not be necessary, indicating that the Afrikaans reflexive paradigm is developing in the same direction as English.

The final paradigmatic change regards the remnants of the Dutch morphological genitive. While the Dutch morphological genitive was still used productively, although infrequently, in the early twentieth century, its only remaining usage contexts are certain fixed expressions. This means that the genitive paradigm of Afrikaans shifted from morphological marking and periphrastic constructions to only periphrastic constructions.

While there are no large paradigm shifts or extensive changes, these instances illustrate the effect of frequency on grammatical change and how a low frequency can aid paradigmatic changes while a high frequency resists it.

<div align="right">

6

</div>

Grammaticalization

GRAMMATICALIZATION AS A PROCESS OF LINGUISTIC CHANGE

In general terms, grammaticalization can be defined as the process through which a less grammatical linguistic item becomes a more grammatical linguistic item. Some linguists distinguish between primary and secondary grammaticalization, where primary grammaticalization refers to *lexical* items developing into *grammatical* items and secondary grammaticalization refers to the development of less grammatical items into more grammatical items.

Using corpora and corpus linguistic methods is a very good fit for investigating grammaticalization. A number of phenomena that can occur in corpus data may point toward grammaticalization:

- The spread of an item changes as the functions of the item change (Bybee 2010a, 964).
- Different items increasingly fuse, and/or the internal structure of an item or construction is reduced (Bybee 2010a, 972).
- The meaning of an item or construction becomes bleached and loses aspects of the original lexical meaning (McMahon 1994, 160).
- The grammaticalizing item increasingly loses independence (Trousdale and Traugott, 2010, 3).
- Tying all of the previous points together, the item or construction becomes increasingly grammatical in nature.

All of these processes are not necessarily involved in every phase of every instance of grammaticalization, but the one component that remains at the center is that the construction becomes increasingly more grammatical.

There are some important issues to keep in mind when using corpora to investigate grammatical change and specifically grammaticalization, which is often tied to changes in frequency. Szmrecsanyi (2016, 153) points out that fluctuation in frequency can be due to other factors than grammatical change, such as environmental change (as he calls it); that is, the context within which something is used changes in frequency, not the construction as such. For example, an increase in the use of pronouns marked as female is due not necessarily to a change in the pronominal system but, rather, to a change in the prominence of women in society. As Szmrecsanyi (2016, 154) points out, changes in text frequencies may be related to culture as much as they are related to language itself.

One of the ways to distinguish between *contextual* or *environmental* changes and *grammatical changes* is to focus on the probabilistic conditioning of variants (Szmrecsanyi 2016, 153–54). If the contexts within which a variant is used expand, it is a more reliable indication of grammaticalization than, for example, if the frequency of the contexts themselves increased. A probabilistic change should be *habitat independent*, or not tied to a specific text or genre, in order to be regarded as real grammatical change (Szmrecsanyi 2016, 156). This factor will be important in determining whether a change is grammaticalization, together with the other indicators listed above.

FUTURE REFERENCE

There are several different ways to reference the future in West Germanic languages, many of which are periphrastic constructions.[1] In English, for example, there are *will* and *shall*, often shortened to *'ll*, and *going to*, often shortened to *gonna*. The German future construction is formed with the verb *werden* 'become' (Hilpert 2008, 131), and Dutch alternates between the use of *zullen* 'shall/will' and *gaan* 'go' (Hilpert 2008, 106). Apart from these expressions of future tense, the above-mentioned Germanic languages (and others) have a futurate present—where the present tense is used to make reference to the future (Hilpert 2008, 157). Afrikaans fits in well with its close family members regarding future reference.

There are three ways to express future reference in Afrikaans. The one is the modal auxiliary *sal* 'will/shall,' with a strong future implicature, inherited from Dutch *zullen*. Another way to indicate future reference is the verb *gaan* 'go,' inherited from Dutch *gaan*, which is used as a lexical verb and an auxiliary in both Afrikaans and Dutch. That these auxiliaries (*sal* and *gaan*) are appropriated for future reference is not surprising, as *sal* followed the

volition path (volition/desire > intention > future) and *gaan* the movement path (movement toward a goal > intention > future) of grammaticalization, both of which are rather common grammaticalization routes (Bybee 2010a, 967; Heine 2003, 594). Finally, Afrikaans also enables the use of the futurate present, where the present tense is used to reference the future, which is not surprising for a Germanic language (Hilpert 2008, 157). While there were no drastic changes in the recent history of Afrikaans future reference, corpus data suggest that it might not have been all that stable in the past century as would seem on the surface.

The Auxiliary *sal*

The auxiliary *sal*, an invariable form developed from Dutch *zullen*, is regarded as a modal auxiliary in the first place (De Stadler 1992, 94; Wybenga 1993, 21), although it is used mainly with a future reference implicature. *Sal* corresponds to English *shall* and *will* to some extent, which are also modal auxiliaries that were appropriated for future reference. As with *shall* and *will*, the modal function of *sal* is used, on the one hand, to indicate participant-internal intention or volition, illustrated in example 6.1, and on the other to indicate participant-external modality regarding prediction, illustrated in example 6.2:

> 6.1. Ek <u>sal</u> maar my kans afwag. (corpus #4, Biographical text)
> *I <u>shall</u> but my chance await*
> 'I <u>will</u> await my chance.'
>
> 6.2. Hiertoe <u>sal</u> die volgende beide eksperimente groteliks opklaring verskaf. (corpus #2, Natural sciences)
> *here.to <u>shall</u> the next both experiments greatly clarity provide*
> 'Both of these experiments <u>will</u> provide clarity in this regard.'

Sal is also occasionally used to indicate participant-external modality without a future implicature (Ponelis 1979, 249), illustrated in example 6.3. These instances that lack implicit future reference, which were only a handful, were discounted in this section, as the focus is on future reference:

> 6.3. 'n Jong man? O, ja, dit <u>sal</u> Willem du Preez wees, om na die gewel te kom kyk. (corpus #3, Fiction)
> *a young man oh yes it <u>shall</u> Willem du Preez be PTCL.INF at the gable to come look*
> 'A young man? Oh, yes, it would be Willem du Preez, to come and look at the gable.'

The distinction between the two general categories of modality—participant–internal and participant-external—is relevant when looking at *sal* in terms of future reference. While intention or volition can obviously imply the future, participant-external modality is more objective and, as De Stadler (1992, 95) points out, "purer" future reference.

This distinction is not simple or easy to make, however, because intention or volition is by implication internal to the speaker, who uses the same tools to indicate intention/volition and prediction. In order to make a meaningful and consistent distinction between the two types of modality (for further information in this regard, see Van der Auwera and Plungian 1998), I compiled the criteria in table 6.1.

Table 6.1.   Criteria for discerning participant-internal and participant-external uses of *sal*

Category	Participant-internal	Participant-external
Semantic context	Explicit indication of personal assurance, intention, etc.	Explicit indication of objective prediction
	Requests	Conditions
Accompanying modal auxiliaries	*moet* (must), *wil* (will), subjective *kan* (can)	objective *kan* (can)
Subject of the clause	Human subject (passive constructions excluded)	Any type of subject
Nature of the main verb	Only action verbs	Action, reception and stative verbs
	Emotional verbs	Any type of verbs
Meaning of negative	Negation indicates nonintention	Negation indicates falseness
Pragmatic context	Interpersonal texts	Factual texts
	Interpersonal situations within texts	Transmission of factual information
Adverbs and adverbials	Indicates intention or obligation	Indicates possibility or likelihood

The first category of criteria I used is the semantic context. When there is explicit indication of intention or volition, such as example 6.4, or when a request is made, such as example 6.5, I took it as participant-internal, while I marked explicit indications of objective predictions or conditions, such as example 6.6, as external:

> 6.4.   Ons <u>sal</u> graag na u vergadering van 22 Februarie aandag aan die verpligtinge skenk. (corpus #4, Manuscript)

> we *shall* gladly after your.HON meeting of 22 February attention
> to the responsibilities pay
> 'We <u>will</u> gladly attend to the responsibilities after your meeting
> of 22 February.'

6.5.   Dan <u>sal</u> jij vir mij ook speel? (corpus #1, Fiction)
       then <u>shall</u> you.SG for me also play
       '<u>Will</u> you then play for me as well?'

6.5.   Soos later baie duidelik <u>sal</u> blyk, het hy hiervoor alles feil ge-
       had. (corpus #2, Biographical text)
       as later very clear <u>shall</u> seem have he here.for all fault had
       'As <u>will</u> be very clear later, all of this was his fault.'

When additional modal auxiliaries were used with *sal*, I took the use of *moet* 'have to' and *wil* 'want to' as an indication of participant-internal modality, while *kan* 'can' are used with both objective ability (taken as external) and subjective intention (taken as internal). Nonhuman syntactic subjects usually excluded participant-internal modality, while human subjects could be used with both.

   If the main verb was a copular verb or if it expressed an action where the subject received something or something was done to the subject, such as example 6.7, I also took it to exclude intention or volition, while other action verbs could be used with both types of modality. Verbs indicating emotion would also typically be associated with participant-internal modality (see example 6.8):

6.7.   Ag, kom bog man, dit <u>sal</u> jou nie dronk maak nie . . . (corpus
       #1, Fiction)
       oh come nonsense man it <u>shall</u> you.SG not drunk make PTCL.NEG
       'Oh, nonsense, it <u>will</u> not make you drunk.'

6.8.   Dat God ons liefhet net soos ons is en dat Hy nooit <u>sal</u> ophou
       om ons lief te hê nie. (corpus #3, Religious text)
       that God us love.have just like we be.PRS and that He never <u>shall</u>
       stop PTCL.INF us love to have PTCL.NEG
       'That God loves us just as we are and that He <u>will</u> never stop
       loving us.'

When a sentence was not in the negative, negating it also helped to determine the type of modality: if negation indicated nonintention rather than falsehood, it was taken as participant-internal. I also took the pragmatic context into consideration—interpersonal texts and contexts typically lend themselves more to participant-internal modality, while factual texts and the transfer

of factual information would rather be associated with participant-external modality. Adjuncts or adverbs indicating intention, like *graag* 'gladly' (as in example 6.4), or objective possibility or likelihood, like *duidelik* 'clearly' (as in example 6.6), were also taken into account.

In every instance, all the criteria were taken into account for the sake of analyses that are as accurate and consistent as possible. There were a couple of rhetorical questions that I disregarded for the analysis. A summary of the analysis is given in table 6.2.

Table 6.2.   Frequencies of *sal* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
participant-internal future reference	96	89	84	103
participant-external future reference	314	265	260	226
TOTAL	410	355	344	329

The total frequency of *sal* declines significantly (log likelihood of 23.07), and while the proportions remain fairly stable in the first three corpora, corpus #4 shows a proportional increase in participant-internal uses and consequently a decline in external uses. The frequency of the external uses, however, decreases throughout and quite significantly so (log likelihood of 34.34)—the participant-external uses in corpus #4 are a mere 72 percent of that of corpus #1. In spite of this significant decline in frequency, the participant-external uses remain more than two-thirds of the total uses even in corpus #4. This shows that, on the one hand, *sal* is still used for purer future reference much more often than not but that, on the other, exactly this use shows a consistent decline throughout the century.

The Verb *gaan*

The Afrikaans verb *gaan* is another direct inheritance from Dutch. At the time of initial Dutch colonization in southern Africa, future time reference with *gaan* was occasionally used, although it "was not yet the primary function of *gaan* with a non-finite verbal complement" (Hilpert 2008, 114) in continental Dutch. In a recent study of contemporary Dutch, Van Olmen and Mortelmans (2009, 363) indicate that all auxiliary uses of *gaan* make up 38 percent of total uses in spoken Dutch, which means that future reference is still a secondary meaning of *gaan* today. As will become clear soon, that is not the case in Afrikaans, where *gaan* has attained a higher level of grammaticalization.

Different from the verb *sal*, which is used only as an auxiliary in Afrikaans, the verb *gaan* functions both as a lexical verb with more concrete and more abstract meanings and as two types of auxiliaries—what Heine (2003, 589) calls layering. *Gaan* can be used as a movement verb (example 6.9), as a more abstract verb meaning "go" or "be about" (see examples 6.10 and 6.11), as a linking auxiliary, as shown in example 6.12, and as a future auxiliary, such as in example 6.13:

6.9. Na 'n rukkie staan hy op en <u>gaan</u> na hul toe; maar eers dro hy sy oë af. (corpus #1, Humanities)
after a while stand he up and <u>go</u> to them to but first dry he his eyes off
'After a while he gets up and <u>goes</u> to them; but first, he dries his eyes.'

6.10. Met my werk <u>gaan</u> dit redelik goed. (corpus #2, Manuscript)
with my work <u>go</u> it rather well
'With my work it <u>is going</u> rather well.'

6.11. Dit <u>gaan</u> oor sy privaatlewe. (corpus #2, Informational text)
it <u>go</u> over his private.life
'It <u>is</u> about his private life.'

6.12. Ek dink ook dit is reg dat jy hulle <u>gaan</u> haal. (corpus #3, Fiction)
I think also it be.PRS right that you.SG them <u>go</u> fetch
'I also think it is good that you <u>go</u> fetch them.'

6.13. Dit <u>gaan</u> egter nie maklik wees nie. (corpus #4, Manuscript)
it <u>go</u> however not easy be PTCL.NEG
'It's not <u>going to</u> be easy, though.'

Ponelis (1979, 244) and Van Rensburg (1998, 166) do not distinguish the two auxiliary uses from each other—they treat the auxiliary function of *gaan* as a "direkte skakelwerkwoord" (direct linking verb), which may or may not include future implication. Carstens (1998) discusses Afrikaans auxiliaries in terms of family resemblance and grammaticalization and identifies several morphosyntactic characteristics according to which Afrikaans auxiliaries can be classified in terms of levels of grammaticalization (Carstens 1998, 77). While Carstens also has only one entry for *gaan*, indicating that she does not distinguish two separate auxiliaries either, some of her distinctions can be useful to separate the two uses: Afrikaans auxiliaries that are grammaticalized to only a limited extent would make way for a more grammaticalized auxiliary to take the V2 position, while it would occur at the end of the clause with

the main verb, as in example 6.14, or directly follow the more grammatical-ized auxiliary but precede the main verb as in example 6.15:

6.14.   Pa *het* al die werksmense <u>gaan</u> roep. (corpus #3, Fiction)
        dad have *already the work.people go call*
        'Dad already went to <u>go</u> and call the workers.'
6.15.   Nee wat, Grahamstand *kan* <u>gaan</u> slaap. (corpus #4, Fiction)
        no that Grahamstown can *go sleep*
        'No, really, Grahamstown *can* <u>go</u> to sleep.'

Example 6.14 also illustrates an important difference between *gaan* as a linking auxiliary and as a future auxiliary, as it is used in the past tense with the (more grammaticalized) past-tense auxiliary *het* 'have,' which takes the V2 position. While the boundary between the two auxiliary uses is not clear cut, as is often the case (Carstens 1998, 78), it was maintained as far as pos-sible. When it was unclear whether a specific instance functions as a linking auxiliary or a future auxiliary, the sentence was reformulated to determine whether it could be converted to the past tense (in which case, it would be a linking auxiliary, like example 6.14) and/or other auxiliaries were added to the sentence—if *gaan* remained in the position directly following the subject, it was classified as a future auxiliary (such as example 6.16, with the linking auxiliary *kom* 'come'), but if the modal auxiliary directly followed the sub-ject, it was classified as a linking auxiliary (like example 6.14):

6.16.   Hulle het gesê ons <u>gaan</u> in die Kaap *kom* sukkel. (corpus #3,
        Informational text)
        they have said we go in the Cape come *struggle*
        'They said we <u>will</u> *come* and struggle in the Cape.'

Table 6.3 gives the frequency details of *gaan* in the corpus data.
    Table 6.3 shows that, while the overall discourse frequency of *gaan* fluctuates without showing clear signs of change, the different uses of *gaan*

Table 6.3.   Frequencies of *gaan* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
movement verb	87	67	75	39
abstract main verb	66	52	42	42
linking auxiliary	70	55	91	76
future auxiliary	23	31	53	87
TOTAL	246	205	261	244

show more definite change. If the two main categories, lexical verb and grammatical verb, are compared, it shows that the lexical verb category was the larger of the two in corpus #1, after which the grammatical uses became more prevalent. The prototypical use of *gaan* changed from being lexical to being grammatical. The greatest contributions to the changes in the different categories come from the movement verbs and future auxiliaries. While the movement verbs contribute more than one-third of the total uses of *gaan* in corpus #1, it subsides thereafter, shrinking to the smallest category in corpus #4—a very significant change (log likelihood of 50.40). Use for future reference is the smallest category in corpus #1 and grows consistently to be the largest in corpus #4—another significant change (log likelihood of 127.72). These trends show that, even though the overall discourse frequency of *gaan* does not increase, the internal proportional frequencies show a clear trajectory of ongoing grammaticalization—the lexical uses decrease, while the grammatical uses increase. In corpus #4, the most grammatical use as future auxiliary contributes the largest proportion of all the categories.

The Futurate Present

Apart from grammatical future reference, with either inflection or auxiliaries, all Germanic languages can use the present tense to make reference to the future (Hilpert 2008, 157). In some Germanic languages, the futurate present is even more common than grammatical future reference devices (Hilpert 2008, 7–8). This means that a study of Afrikaans future reference cannot ignore the futurate present.

However, extracting examples of the futurate present from the corpora is not a simple procedure. Because the corpus data are not annotated in any way, I cannot extract a sample of present-tense sentences and determine the proportion that is used to reference the future. Luckily, Hilpert (2008) provides an alternative method for investigating the futurate present in corpus data that also avoids certain other caveats.

Following Hilpert (2008, 158), the futurate present is (narrowly) defined as a present-tense verbal construction in the matrix clause of a sentence referring to the future. This excludes some contexts that are sometimes included in the definition of futurate presents, for example, syntactically subordinate verbal constructions, examples with modal verbs, and imperatives (Hilpert 2008, 158). For analysis, I further follow Hilpert (2008, 160) in focusing on examples "that unambiguously encode future time reference"—those with future-time adverbials, such as *more* 'tomorrow' or *volgende jaar* 'next year.' The samples are divided into three categories: (1) the futurate present

Table 6.4.   Frequencies of clauses with future adverbials per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
sentences with *sal*	28	28	31	27
sentences with *gaan*	3	3	5	10
futurate present	8	16	13	15
TOTAL	40	47	48	52

constructions, (2) instances with *sal* or *gaan*, and (3) those that in fact do not refer to the future, which are excluded from further analyses.

In some Germanic languages, such as German and Dutch, the futurate present is used more often than other future constructions (Hilpert 2008, 157). However, it does not seem to be the case for Afrikaans. In all the instances extracted with future adverbials, *sal* and *gaan* collectively account for far more instances of future reference than the futurate present, and even *sal* on its own is used much more often than the futurate present throughout. While these instances of *gaan* show a similar trajectory of increase to the overall use of future *gaan*, it is still not used as often as the futurate present even in corpus #4.

Unfortunately, the data sets are quite small, with sets of between 22 and 42 identified futurate presents per corpus. Constructions with *sal* from the same data sets range between 70 and 81 and *gaan* from eight to 26. The small size of the sets does not lend itself to much further analysis, although I can make a few remarks about the available data.

The data suggest two possible changes in preference regarding the type of verbal constructions used in the futurate present. First, it shifts from a strong preference for transitive constructions in the first half of the century to more or less equal amounts of transitive and intransitive constructions in the second half. Second, a strong preference for animate syntactic subjects in corpus #1 to corpus #3 changes to a slight preference for inanimate subjects in corpus #4. However, as I already mentioned, the data sets are too small to make too much of this apart from mentioning that it seems like the preferences for constructions with the futurate present differ from those of future constructions with *gaan* (see the following section for details of future constructions with *gaan*).

Comparison: Future Reference with *sal* and *gaan*

The data indicate that the future auxiliary frequencies of *sal* and *gaan* show opposite patterns, and I investigate the relationship between these two auxiliaries in further detail here.

I have already mentioned that *sal* is a modal auxiliary in the first place, which is used with an implicature of future reference. In contrast, while *gaan* can sometimes be used as a marker of certainty, it is even in those cases still a future auxiliary in the first place (De Stadler 1992, 92). In his authoritative work on Afrikaans syntax, Ponelis (1979, 244) claims that *gaan* has a stronger factual connotation than *sal* and can be paraphrased with "will definitely." It is on these grounds that De Stadler (1992, 95) claimed that *gaan* is a "purer" future reference than *sal*. In a close investigation of the use of *gaan* in future reference, I have found that it is used not with participant-internal modality but, rather, for objective, epistemic predictions about the future. The best way to illustrate this difference is with questions. Example 6.17 is a question from corpus #4, paraphrased in example 6.18 with *gaan* instead of *sal*:

6.17. <u>Sal</u> jy my leer toor, Arboreta? (corpus #4, Fiction)
       shall you.SG me teach magic Arboreta
       '<u>Will</u> you teach me to perform magic spells, Arboreta?'
6.18. <u>Gaan</u> jy my leer toor, Arboreta?
       go you.SG me teach magic Arboreta
       'Are you <u>going to</u> teach me to perform magic spells, Arboreta?'

Just as in English, example 6.17 would be interpreted as a request, while example 16.18 would be an inquiry that refers not to intention but to likelihood. This means that the use of *gaan* with future reference is competition not for the use of *sal* with participant-internal modality but only for epistemic future reference. Table 6.5 shows the frequencies of *gaan* and *sal* used in participant-external future reference.

Table 6.5. Frequencies of *sal* and *gaan* with participant-external future reference per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
sal	314	265	260	226
gaan	23	31	53	87

While *sal* is still used much more frequently for participant-external future reference than *gaan* in written Standard Afrikaans, there is a continuous, significant growth in the use of *gaan* for the same function. It is clear, then, that *sal* is making way for *gaan* and that the total frequency of future reference remains fairly stable. It cannot be that simple, however, and many studies have illustrated that a grammaticalizing item does not spread equally

through different genres and contexts and levels of formality. Unfortunately, the corpora I used in this study are fairly small, and dividing the data into the different genres fractures them beyond use. There is a prescriptive source that indicates that *gaan* is less formal than *sal* (Müller 2003, 85), and in order to see how that might affect the spread of *gaan*, I divided the corpora into three levels of formality: the formal letters and academic texts fall into the formal category, personal letters and diary entries fall into the informal category, and the rest falls into a neutral category. I then calculated the use of *gaan* per 100,000 words in each category of formality, the details of which are indicated in figure 6.1.
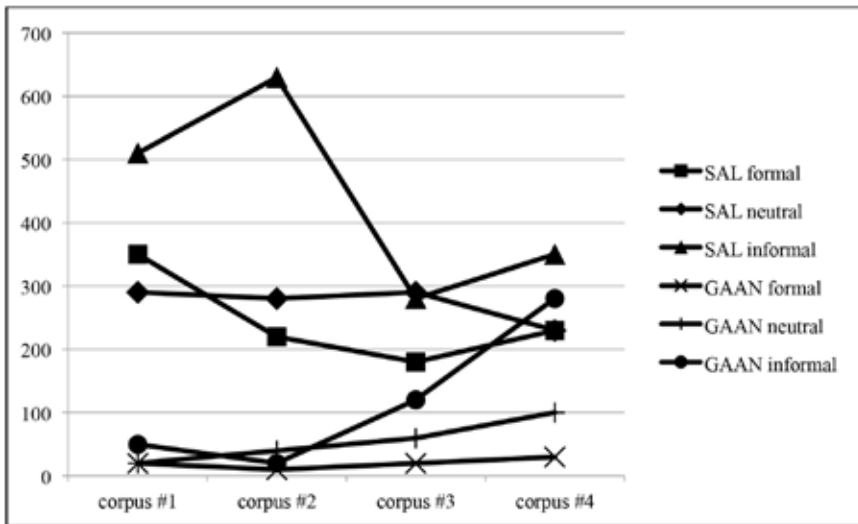
Figure 6.1. Patterns of formaliy in the use of *sal* and *gaan*

Figure 6.1 shows that, even though the frequency of future *gaan* increases in all three categories from corpus #2 on, the increase is the greatest in informal texts. On the other hand, *sal* is more variable, but in broad strokes, it shows the greatest decrease in informal texts. As is often the case, the newer grammaticalizing variant is spreading faster in informal language use, in this instance even overtaking *sal* in formal and neutral contexts.

The analyses above show clearly that *gaan* is continuing on its grammaticalization path, with the resulting increase in frequency of its use for future reference and its spread from informal to formal contexts.

PRONOUNS

A handful of Afrikaans pronouns and pronominal constructions show signs of grammaticalization, which can be seen mainly in the reduction of forms or the increasing frequency of reduced variants.

Third-Person Inanimate Pronoun

In one section of the previous chapter, I focused on the third-person inanimate pronoun *dit* 'it' and a small change in its functional spread. This section, however, focuses on the use of *dis*, which is a contracted form of *dit is* 'it is,' similar to English *it's*. The variant *dis* is used in most of the same functions as its purely pronominal counterpart *dit*: it can refer to inanimate referents that are singular, plural, or unspecified for number (example 6.19), it can refer anaphorically to the (or a) previous sentence (example 6.20), it can refer cataphorically to the (or a) sentence or clause that is to follow (example 6.21), and it can be used as a dummy subject (example 6.22):

> 6.19. Begin jou dag met hawermout—<u>dis</u> eintlik superkos . . . (corpus #4, Informational text)
> *start your.SG day with oats it.s actually super.food*
> 'Start your day with oats—<u>it's</u> actually super food.'
> 6.20. 'n Mens wag jou hier dood. Maar <u>dis</u> darem nie oral so nie! (corpus #2, Fiction)
> *a human wait you.SG.REFL here dead but it.s at.least not everywhere so PTCL.NEG*
> 'You wait yourself to death here. But <u>it's</u> not like that everywhere!'
> 6.21. <u>Dis</u> in sij belang, dat die mens glo, dat hij nie meer leef nie. (corpus #1, Religious text)
> *it.s in his interest that the human believe that he not anymore live PTCL.NEG*
> '<u>It's</u> in his interest, that humanity believes, that he does not live anymore.'
> 6.22. <u>Dis</u> sommer koud & effense buie val. (corpus #3, Manuscript)
> *it.s just cold & slight showers fall*
> '<u>It's</u> cold with light showers falling.'

The frequencies of *dis* per 100,000 words, according to the different functions, are given in table 6.6.

Table 6.6.   Frequencies of *dis* per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
third-person inanimate	19	15	34	28
dummy subject	10	9	10	10
sentence anaphoric	31	26	37	43
sentence cataphoric	43	48	39	43
TOTAL	103	97	121	124

While the frequency of the different functions fluctuate or remain more or less stable, the overall increase in frequency is statistically significant (log likelihood of 13.19). This increase suggests grammaticalization, as the contracted variant *dis* (without an apostrophe, as is used with other contractions) shows a reduction in form and a loss in independence of the two original items *dit* and *is* in this specific context.

The Generic Pronoun *('n) mens*

Many Germanic languages have (and others had) a generic pronoun (also called *human impersonal pronoun* in, e.g., Van Olmen and Breed 2018) etymologically derived from *man* (human), such as German, Swedish, and Norwegian *man* and Dutch *men*, referring to "people in general" (Coussé and Van der Auwera 2012, 121; Hoekstra 2010, 32). In certain languages, this variant can refer to a definite referent, such as Swedish *man*, but in others, it cannot, such as Dutch *men* (Coussé and Van der Auwera 2012, 122). These forms may sometimes also include the speaker, such as Fries *men*, in which case it refers to "people like us" (Hoekstra 2010, 37). The Dutch *men* is used mainly in formal language and is in competition with *je* 'you,' *we* 'we,' and *ze* 'you' (Coussé and Van der Auwera 2012, 123; De Hoop and Tarenskeen 2015, 164).

After the establishment of Standard Afrikaans, the already infrequent Dutch *men* (17 uses per 100,000 words) ceased to be used at all in the subsequent corpora. A new generic pronoun, or human impersonal pronoun, developed in its place, however: *('n) mens*, literally '(a) human.' This is not uncommon in Germanic languages, as mentioned above, and the new variant is already widely used in corpus #1.

In the corpus data, the lexical item *mens* 'human' is used in several different ways. It is used as a noun in reference to a concrete, specific person, as in example 6.23. It is also used as a more general noun, referring to "humanity" (example 6.24), thus already generic to some extent but not yet pronominal. This meaning forms a bridge to the grammaticalized, pronominal use, often

with the indefinite article that would accompany its use as a pronoun (example 6.25) but also without the article, as a pure pronominal (example 6.26):

6.23. Sy begin haar self as aparte <u>mens</u> te ontdek. (corpus #2, Humanities)
she begin her self as separate <u>human</u> to discover
'She starts to discover herself as separate <u>person</u>.'

6.24. Die <u>mens</u> met sy beperkte en verduisterde vermoëns is ook soeker na skoonheid. (corpus #3, Religious text)
the <u>human</u> with his limited and darkened abilities be.PRS also seeker of beauty
'<u>Humanity</u> with its limited and darkened abilities also seek beauty.'

6.25. Ek weet watter plek 'n mens in die dorp nie drank kan kry nie! (corpus #1, Fiction)
I know which place <u>a human</u> in the town not liquor can get PTCL.NEG
'I know where in town <u>you</u> cannot find liquor.'

6.26. Min dinge skok <u>mens</u> nog. (corpus #4, Informational text)
few things shock <u>human</u> still
'Few things still shock <u>people/you</u>.'

Especially when used without the article, as bare *mens*, the pronominal character is clear. The more typical structure of a noun phrase is replaced with a bare form, and only in this context, when used as a pronoun. The frequencies of the different uses of the lexical item *mens* are given in table 6.7.

Table 6.7. **Frequencies of *mens* per 100,000 words**

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
concrete noun	21	23	20	24
"humanity" noun	20	25	51	27
generic pronoun	95	60	56	73

The pronominal use of *mens* is the largest category throughout, being used more often as a pronoun than a lexical noun in all the corpora except corpus #3. In corpus #3, the more abstract use referring to "humanity" is used more than twice as much as the concrete use, which still points to more abstract meanings being prototypical rather than the more concrete meaning. The data do not show any visible proportional or frequency changes, only some fluctuations. However, as I indicated above, there are two variants used as

pronouns: the extended phrase *'n mens* 'a human' and the reduced, more grammaticalized bare form *mens* 'human.' The usage frequency of the two competing pronominal variants are given in table 6.8.

Table 6.8.   Frequencies of the two pronominal forms of *('n) mens* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
'n mens	81	48	43	46
mens	14	13	13	27

The overall frequency of pronominal uses fluctuate without definite change, and the specific frequencies also do not change consistently: the extended variant drops quite drastically from corpus #1 to corpus #2 but then fluctuates for the rest of the century. The bare variant remains fairly stable from corpus #1 to corpus #3 but then rises sharply in corpus #4. The general pattern, however, is a drop in the extended variant and a rise in the bare variant, and both of these overall changes are statistically significant (decline of *'n mens* from corpus #1 to corpus #2: log likelihood of 21.90; increase of *mens* from corpus #3 to corpus #4: log likelihood of 13.20). This reduction in form, or increasing use of the reduced variant, is typical of grammaticalization. Still, the extended variant makes up almost two-thirds of the overall frequency, indicating that the extent of grammaticalization might still be limited. It is important to note, though, that the corpora consists entirely of written material. In the absence of spoken data, the distinction between data from published material and data from manuscript sources can be used to gauge the measure of influence regarding the formality of the context. The frequencies of bare *mens* per 100,000 words are given in figure 6.2.

It is clear that the increase in use of bare pronominal *mens* is much more drastic in manuscript texts than in published texts. This is often the case in the spread of new (grammaticalized) variants, where it is accepted in more informal contexts first and then spreads to more formal contexts later. This is also apparent from figure 6.2, where the frequency in published texts drops slightly from corpus #1 to corpus #3 but then also starts to rise, quite sharply, in corpus #4.

Indefinite Pronouns with *enig-*

The class of indefinite pronouns share, obviously, the characteristic of being indefinite. However, as De Stadler (1989, 340) points out, definiteness and indefiniteness are relative, and definiteness should be viewed as a nondiscrete semantic value. Afrikaans has a core group of indefinite pronouns: *iemand*
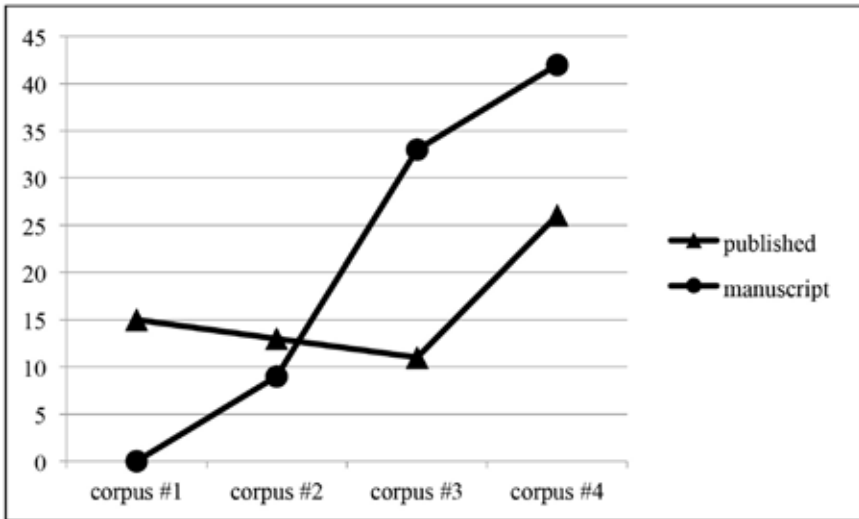
Figure 6.2. **The use of pronominal bare *mens* in published vs. manuscript texts**

'somebody,' *iets* 'something,' *niemand* 'nobody,' and *niks* 'nothing' (Ponelis 1979, 101). The class itself remains open, however, and new forms are added continually, typically from sources such as noun phrases, indefinite numerals, and Q- and W-words (Ponelis 1979, 100). Other words frequently used as indefinite pronouns in the data include *alles* 'everything,' *almal* 'everybody,' *elkeen* 'everyone,' and a group—the focus of this section—of forms merged with *enig-* 'any-,' such as *enigiemand* 'anybody,' *enigiets* 'anything,' and *enigeen* 'anyone.'

Ponelis (1979, 118) indicates that the quantifier *enige* 'any,' sometimes used as premodifier with several indefinite pronouns, not only are used as a separately written premodifier with *iemand* 'somebody,' *iets* 'something,' and *een* 'one' but also, as in English, can merge with the pronouns, forming the variants *enigiemand* 'anybody' (6.27), *enigiets* 'anything' (6.28), and *enigeen* 'anyone' (6.29). Different from English, *een* 'one' is not used pronominally on its own, but it acquires pronominal meaning when merged with *enig-*. The word *enige* on its own can be used as a free-choice item with universal reading, similar to English *any*, while this is not the case in Dutch, where *enig(e)* has a more limited use (Hoeksema 2010, 842):

6.27. Enigiemand, van babas tot bejaardes, kan voordeel trek uit die gebruik van die 38 blomme-geneesmiddels. (corpus #4, Informational text)

> *anybody from babies to elderly can benefit pull from the use of
> the 38 flower-medicines*
> 'Anybody, from babies to the elderly, can benefit from the use
> of the 38 flower medicines.'

6.28.   Of julle dan eet of drink of <u>enigiets</u> doen, doen alles tot ver-
heerliking van God. (corpus #2, Religious text)
*whether you.PL then eat or drink or <u>anything</u> do do everything
to glory of God*
'Whether you eat or drink or do <u>anything</u>, do everything for
God's glory.'

6.29.   As jy my toe 'n rewolwer gegee het, sou ek <u>enigeen</u> kon gaan
doodskiet, <u>enigeen</u> wat vaagweg agter sy dood gestaan het.
(corpus #3, Fiction)
*if you.SG me then a revolver gave have would I <u>anyone</u> could
go dead.shoot <u>anyone</u> that vaguely behind his death stood have*
'If you gave me a revolver then, I would have shot <u>anyone</u>
dead, <u>anyone</u> that was remotely involved in his death.'

The frequencies of indefinite pronouns merged with *enig-*, per 100,000
words, are as follows:

- Corpus #1: less than 1 per 100,000 words
- Corpus #2: 3 per 100,000 words
- Corpus #3: 7 per 100,000 words
- Corpus #4: 11 per 100,000 words

While the frequencies remain fairly low, the increase is statistically signifi-
cant (log likelihood of 35.36). The merged form that occurs first is *enigiets*,
occurring once in corpus #1, and its frequency remains the highest of the
three related forms. The form *enigeen* occurs for the first time in corpus #2,
and finally *enigiemand* emerges in corpus #3. With all three, the frequency
rises in every corpus from the first occurrence. This rise in the frequency of
merged forms indicates ongoing grammaticalization of this type of indefinite
pronouns.

GENITIVE: THE GRAMMATICALIZATION OF *SE*

In contemporary Afrikaans, there are two nonpronominal ways to express
genitive relations. The most common way, in written Afrikaans at least, is

with *van* (as in example 6.30), similar to Dutch *van*, English *of*, and German *von*:

> 6.30. Ek het deur die lees <u>van</u> 'n traktaatjie tot bekering gekom. (corpus #4, Religious text)
> *I have through the read <u>of</u> a tract.DIM to conversion came*
> 'I converted by reading a tract.'

There is some disagreement among linguists about whether to call the genitive use of *van* a specific category of preposition usage or a genitive construction. Scholars such as Ponelis (1979, 126) include only *se* ''s' (see below) and possessive pronouns in the term *genitive*, with all *van* uses prepositional, and Hantson (2001, 10) calls this phenomenon in Germanic languages prepositional as well. However, Van Rooy (1984) makes a case for distinguishing between the prepositional use of *van* in Afrikaans and the genitive use. While I do not intend to involve myself in this debate, I do want to establish that I include the relevant *van*-constructions when referring to Afrikaans genitive constructions.

Another way to express genitive relations in Afrikaans is with the genitive particle *se*, as in example 6.31:

> 6.31. Pappie sit die rystoel op die eetkamer <u>se</u> tapyt met die rooi-roosrand neer. (corpus #3, Fiction)
> *Daddy put the wheelchair on the dining.room PTCL.GEN carpet with the red.rose.edge down*
> 'Daddy puts the wheelchair on the dining room<u>'s</u> carpet with the red rose edge.'

These two constructions, with *van* and with *se*, are interchangeable in certain contexts but not in others. This means that, in many cases, speakers make a choice between either a *se*-genitive or a *van*-genitive. In other languages with similarly alternating constructions, such as Dutch and English, several factors influence the choice in different ways. This alternation in Afrikaans has received only limited attention in the literature.

Originally, Proto-Germanic would have had an extensive case system that included the vocative, nominative, accusative, dative, instrumentive, and genitive (Scott 2011, 106). In contemporary Germanic languages, the case systems have simplified to a large extent and in some cases disappeared as a system, even if retaining some remnants of (among others) the genitive. In North Germanic, there is still an *s*-genitive with some ties to the original case (Börjars 2003; Perridon 2013). The West Germanic languages have similar

constructions—English, German, and Dutch have *s*-genitives, even if there are semantic differences and different diachronic development paths that led to the contemporary constructions in the different languages (Hendriks 2012; Scott 2011; Weerman and De Wit 1999).

By the end of the Middle Dutch period, before the origin of Afrikaans, Dutch had lost most of its formal case system (Scott 2011, 106; Weerman and De Wit 1999, 1157), and today the *s*-genitive retains only limited use. It has a more limited distribution than the English equivalent, as it is usually used with proper names or forms of address (Hantson 2001, 7; Weerman and De Wit 1999, 1165–67), and it is not usually repeated in the same construction (Hantson 2001, 10). This *s*-genitive has disappeared from contemporary Afrikaans language use.

Another Dutch construction to express genitive relations is the pronominal doubling construction, typically exemplified by *Jan z'n boek* 'Jan his book' (Weerman and De Wit 1999, 1170), with female and plural forms as well. This type of construction has been in use since the fourteenth or fifteenth century (Weerman and De Wit 1999, 1173), which means it was transplanted to South Africa with seventeenth-century Dutch. This is relevant because of the origin of the Afrikaans *se*-genitive construction.

The Afrikaans *se*-genitive originated partially from the Dutch pronominal doubling construction, specifically the singular male form *zijn* 'his' and its reduced form *z'n*. While Afrikaans did not retain any of the reduced pronouns used in Dutch, the sound change from *zijn* [zəin] to [zən] to [sə] is not unexpected since children acquiring Dutch use a similar reduction until the extended system is learned (Weerman and De Wit 1999, 1173). There are similar forms in some Western Flemish varieties (Haegeman 2013, 221).

The pronominal doubling construction in Dutch is fairly restricted in its use, as is the Dutch *s*-genitive. However, this is not the case with Afrikaans *se*, which is not restricted in terms of the animacy, definiteness or number of the possessor, the length of the possessor, the mode (i.e., written or spoken), or the formality of the situation. Den Besten (2012d, 9) notes that Afrikaans *se*-genitives underwent a "tremendous expansion" in expressive power in its development from Dutch, often ascribed to influence from indigenous Khoikhoi and the Creole Portuguese and Pasar Malay of the slaves (Roberge 1996, 130). In many ways, it is closer in usage contexts and parameters to the English *s*-genitive than to the Dutch *s*-genitive or pronominal doubling construction, a point I will return to later.

In my corpus data, the frequency of *se* has increased quite significantly during the century represented by the data (see details below), something that is often associated with continuing grammaticalization. New (gram-

maticalizing) linguistic variants often do not spread at the same speed in all contexts (linguistic and otherwise), however, so it is pertinent to investigate in which contexts *se* has been gaining ground on *van* and how the use of *se* has expanded during the past century. In order to achieve this, I focus on the alternation between *se* and *van* in interchangeable contexts and how it relates to certain linguistic and other textual variables. More specifically, I investigate how the contribution of these variables to the choice of either *se* or *van* changed in the course of the past century, showing context expansion in the case of *se*. To this end, I model the data with decision trees on the one hand and logistic regression on the other. The results from these models enable a nuanced description of the different factors contributing to genitive choice and the ongoing grammaticalization of *se*.

Regarding extracting the relevant data from the corpora, for corpus #1, I extracted the variants of *se* apart from the forms *se* and *van* used in the rest of the corpora. The variants include *z'n* and *s'n* and a handful of constructions that were closer to pronominal doubling constructions, that is, with gender and number agreement.

While there were only a handful of instances of *se* in each corpus that were discarded (typically because it was a typing or spelling error and not really a *se*-genitive), it was a more complex process to identify the relevant *van* instances. I used three categories to classify *van* in the first round of analysis: (1) those that are not genitives (mostly other prepositional uses as well as some others, like being part of a surname), (2) those that are genitives but are not interchangeable with *se*, and (3) those that are interchangeable with *se*. Examples 6.32 to 6.35 illustrate nongenitive uses of *van*: examples 6.32 to 6.34 have prepositional uses, and example 6.35 shows *van* as part of a surname:

6.32. Sy is 'n gawe, die grootste gawe <u>van</u> God aan hom. (corpus #3, Religious text)
she be.PRS a gift the biggest gift <u>from</u> god to him
'She is a gift, the greatest gift <u>from</u> God to him.'

6.33. Met beste groeten <u>van</u> huis tot huis. (corpus #1, Manuscript)
with best greetings <u>from</u> house to house
'With greetings <u>from</u> house to house.'

6.34. Ek hou vreeslik baie <u>van</u> jou, N . . . (corpus #2, Manuscript)
I like very much <u>of</u> you.SG N
'I like you so much, N . . .'

6.35. (<u>Van</u> Niekerk, 2001:54–57) (corpus #4, Humanities—a text reference)

There are many more contexts where genitive *van* can be used but not *se* than the other way around. Because of the similarities with English, I used as a starting point the criteria of exclusion and inclusion of the many studies of the English equivalents (Heller, Szmrecsanyi, and Grafmiller 2017; Hinrichs and Szmrecsanyi 2007; Rosenbach 2002; 2014). As I mentioned earlier, the extent and parameters of uses of Afrikaans *se*-genitives resemble English *s*-genitives more closely than Dutch *s*-genitives or pronominal doubling constructions. In fact, in analyzing thousands of Afrikaans genitives (both *se* and *van*), I have found that the guidelines for exclusion supplied in Bresnan et al. (2017) for English can be applied to Afrikaans almost entirely, the only exceptions being parameters that do not apply to Afrikaans genitive constructions at all: Afrikaans *se*-genitives are not used as classifying genitives, such as "widow's peak" in English, and are also not used in double genitive constructions, such as English "a friend of John's" (see Bresnan et al. 2017).

The *van*-genitives that were excluded are appositives, partitives, descriptive genitives (including those attributing material/content or origin/source), genitives with possessums starting with a determiner other than the definite article *die* 'the' (or no determiner at all), and idiomatic or fixed expressions. Following Rosenbach's (2014, 225) recommendation, all other genitives were included whether they are *likely* to be expressed by both constructions or not: "rather than testing each individual expression for genuine interchangeability we should rather proceed from possible classes of genitive constructions, excluding all those contexts which we know categorically ban one of the variants." As Hinrichs and Szmrecsanyi (2007, 446) comment on their analysis of English genitive alternation: "Our notion of 'interchangeability' does not imply that the *s*-genitive and the *of*-genitive would have been equally 'felicitous' in the context at hand. . . . The aim of our analysis is indeed to explain what makes one genitive construction more felicitous than the competing one." The only *se*-genitives that were excluded from the final analyses are those used in idiomatic or fixed expressions, which were only a handful in all of the corpora. The following examples show uses of genitive *van* that are not interchangeable with *se*: example 6.36 shows an appositive genitive, example 6.37 a partitive genitive, example 6.38 a descriptive genitive, and example 6.39 a *van*-genitive with a possessum without the definite article *die*, and examples 6.40 and 6.41 provide fixed expressions that cannot be replaced with *se*-genitives. Examples 6.42 and 6.43 show the use of *se* with fixed expressions that cannot be replaced with a *van*-construction:

> 6.36.   . . . die daad <u>van</u> misdaadpleging. (corpus #3, Humanities)
> *the deed <u>of</u> crime.committing*
> 'the deed <u>of</u> committing crime.'

6.37.   . . . die feit dat baie <u>van</u> die Afrikaanse skole enorm gegroei
         het . . . (corpus #1, News report)
         the fact that many <u>of</u> the Afrikaans schools enormously grown
         have
         . . . 'the fact that many <u>of</u> the Afrikaans schools have grown
         enormously . . .'

6.38.   . . . die leer <u>van</u> die regverdigmaking deur die geloof . . . (cor-
         pus #2, Religious text)
         the teach <u>of</u> the sanctification through the faith
         . . . 'the teaching <u>of</u> sanctification through the faith . . .'

6.39.   Daar was voorstanders en teenstanders <u>van</u> die inhoud van die
         mosie. (corpus #4, Humanities)
         there be.PST supporters and opponents <u>of</u> the content of the mo-
         tion
         'There were supporters and opponents <u>of</u> the content of the
         motion.'

6.40.   Ek meen ook om <u>een van die dae</u> my skietgeweer te hê. (cor-
         pus #2, Manuscript)
         I mean also PTCL.INF <u>one of the days</u> my fire.arm to have
         'I also mean to get my fire-arm <u>one of these days</u>.'

6.41.   Die Nasionale Departement <u>van</u> Minerale en Energie (corpus
         #4, Natural sciences)
         The National Department <u>of</u> Minerals and Energy

6.42.   "Dienswillig <u>se stert</u>!" snork Grootraad. (corpus #4, Fiction)
         dutiful PTCL.GEN tail snort Great.council
         '"Dutiful <u>my ass</u>!" snorts Great-One.'

6.43.   Wêreldbeker-babbelaas? <u>Se voet in 'n visblik</u>! (corpus #4,
         News report)
         world.cup-hangover PTCL.GEN foot in a fish.can
         'World cup hangover? <u>Not a chance</u>!'

After the exclusion of all the noninterchangeable classes of both *se* and *van*, I coded the data for a number of variables, again using as a starting point many of the previous studies of English genitive alternation (Ehret, Wolk, and Szmrecsanyi 2014; Heller, Szmrecsanyi, and Grafmiller 2017; Jankowski and Tagliamonte 2014; Rosenbach 2002; 2008; 2014). Because of the already established similarities between the Afrikaans and English constructions, I used the variables from the studies on English as an exploratory starting point. These variables are animacy of the possessor, definiteness of the possessor, semantic relation expressed by the genitive, length of both the possessor and the possessum, and the text type within which the particular instance occurs.

First, I will provide some descriptive statistics for the distribution of the two genitive variants according to the individual variables. Then, in order to compare how the different factors affect the choice of using either *se* or *van*, as well as how they interact with each other, I model the data from each corpus with a decision tree as well as logistic regression. I compare the results from the two different models and specifically focus on where they diverge and converge. All of this is instrumental in revealing the grammaticalization of *se* and how its use is expanding in many different contexts.

Descriptive Statistics of Afrikaans Genitive Alternation

In the past century, the relationship between *se* and *van* has changed quite drastically, primarily because of the increased use of *se*. Table 6.9 shows the usage of interchangeable *se* and *van* per 100,000 words in each of the corpora.

Table 6.9. Frequencies of *se* and *van* in interchangeable contexts per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
se	133	223	426	580
van	1,274	1,473	1,305	1,072
TOTAL	1,406	1,696	1,730	1,651

In corpus #1, *se* contributes only just over 9 percent of the total interchangeable genitives, and by corpus #4, it contributes 35 percent. This shows a very significant increase in usage (log likelihood of 953.51), which is worth exploring further in more detail to determine how the spread of *se* proceeds in different contexts and text types.

While *se* is the preferred form for the majority of prenominal genitives in corpus #1, the form *s'n* was still used fairly regularly. The corpus also contains a number of other pronominal forms, still marked for gender and number, in the same construction, as in examples 6.44 to 6.47, such as *haar* 'her' and the reduced form *h'r*, *hulle* 'their' and the reduced *hul*, and the fuller masculine forms *zijn/sijn/sij/sy* 'his':

6.44. Maar net om Johanna <u>haar</u> ontwil. (corpus #1, Manuscript)
        but just for Johanna her sake
        'But only for Johanna*'s* sake.'
6.45. . . . die openlik verkondigde doel van die Engelse <u>hul</u> intervensie. (corpus #1, Informational text)

> *the openly preached goal of the English <u>their</u> intervention*
> . . . 'the openly preached goal of the English<u>'s</u> intervention.'

6.46. Die professore <u>s'n</u> kritiek is baie skerp. (corpus #1, Informational text)

> *the professors <u>his.CONTR</u> criticism be.PRS very sharp*
> 'The professors<u>'</u> criticism is very sharp.'

6.47. Oom <u>sijn</u> kennis is nie al te groot nie . . . (corpus #1, Fiction)

> *uncle <u>his</u> knowledge be.PRS not all too big PTCL.NEG*
> 'Uncle<u>'s</u> knowledge is not very broad . . .'

I included the forms that were used in the same way as *se*—they were still regarded as refined forms but equivalents of *se* in the early twentieth century (Roberge 1996, 138). Because of the small number of pronominal forms, its functional similarities to *se*, and the fact that it was regarded as equivalents of *se*, I include them in my overall analyses, along with *s'n* and *se*. When I refer to *se* in the context of corpus #1, then, these equivalents are included.

Possessor Animacy

The animacy of the possessor has been shown to be an important factor in genitive choice in both Dutch and English. In Dutch, there is a strong constraint on prenominal possessors, mostly allowing only human possessors, with some room for personified animals (there is some variation in different varieties of Belgian Flemish, however, according to Haegeman 2013, 229). Rosenbach (2005, 627) has shown that animate possessors strongly favor the use of the *s*-genitive in English.

In some studies, animacy is not regarded as a simple binary between animate and inanimate, while other researchers find it useful to collapse the fine-grained categories, especially when the data sets are too small. Heller, Szmrecsanyi, and Grafmiller (2017, 10) do not find a more fine-grained distinction relevant for their study, but Rosenbach (2008), on the other hand, does. She maintains a continuum of animacy (Rosenbach 2008, 153):

> *animate*   inanimate
> human N > animal N > collective N > temporal N > locative N > common inanimate N

Collective nouns, such as *company*, are more to the animate end of the continuum because they can be conceptualized as both inanimate (an institution) and animate (a group of people) (Rosenbach 2008, 153). Temporal and locative nouns have also been shown to act differently from common inanimates as possessors in English and are therefore analyzed in their own

categories (Rosenbach 2008, 153–54). In my analysis, I collapsed *human* and *animal*, including only higher animals, typically associated with humans and regarded as individuals, corresponding to the *animate* category used in Haegeman (2013, 229) in a study on prenominal possessors in West Flemish. Others also sometimes include all *animates* into one category when studying English genitive alternation (e.g., Ehret, Wolk, and Szmrecsanyi 2014, 269; Grafmiller 2014, 477; O'Connor, Maling, and Skarabela 2013, 97). Afrikaans genitives with possessors in the different animacy categories are exemplified below: animate possessor, example 6.48 with *van* and example 6.49 with *se*; collective possessor, example 6.50 with *van* and example 6.51 with *se*; temporal possessor, example 6.52 with *van* and example 6.53 with *se*; locative possessor, example 6.54 with *van* and example 6.55 with *se*; "other inanimate" possessor, example 6.56 with *van* and example 6.57 with *se*:

6.48.   Pestalozzi het gebouw op die sielkunde <u>van Aristoteles</u>. (corpus #1, Informational text)
         Pestalozzi have built on the psychology <u>of Aristotle</u>
         'Pestalozzi built on the psychology <u>of Aristotle</u>.'

6.49.   <u>Notiti se</u> lyf het die watervlak gebreek. (corpus #3, Fiction)
         <u>Notiti</u> ᴘᴛᴄʟ.ɢᴇɴ body have the water.surface broke
         '<u>Notiti's</u> body broke the water surface.'

6.50.   Werk vir die vooruitgang <u>van ons Korps</u>! (corpus #1, News report)
         work for the advancement <u>of our Corps</u>
         'Work for the advancement <u>of our Corps</u>!'

6.51.   Ek het die kamp van <u>Universiteit L se</u> C.S.V. verlede naweek geniet. (corpus #2, Manuscript)
         I have the camp of <u>University L ᴘᴛᴄʟ.ɢᴇɴ</u> C.S.O. last weekend enjoyed
         'I enjoyed the camp of <u>L University's</u> C.S.O. last weekend.'

6.52.   Dit is alreeds die Maandagmôre <u>van die tweede week</u>. (corpus #2, Fiction)
         it be.ᴘʀs already the Monday.morning <u>of the second week</u>
         'It is the Monday morning <u>of the second week</u> already.'

6.53.   Miskien <u>vandag en vanaand se</u> baie huiswerk. (corpus #2, Humanities)
         maybe <u>today and tonight ᴘᴛᴄʟ.ɢᴇɴ</u> much homework
         'Maybe <u>today and tonight's</u> lots of homework.'

6.54.   Ja, varings, in die middel <u>van die Groot Karoo</u>. (corpus #4, Fiction)

> *yes ferns in the middle <u>of the Great Karoo</u>*
>
> 'Yes, ferns, in the middle <u>of the Great Karoo</u>.'

6.55. Omtrent al <u>Abidjan se</u> taxibestuurders is Moslems. (corpus #3, Biographical text)

about all <u>Abidjan PTCL.GEN</u> taxi.drivers be.PRS Muslims

'Almost all of <u>Abidjan's</u> taxi drivers are Muslims.'

6.56. Dis 'n kans om die positiewe dinge <u>van ons volkwees</u> te beklemtoon. (corpus #3, Informational text)

it.s a chance PTCL.INF the positive things of our nation.being to emphasize

'It's an opportunity to emphasize the positive things <u>of being our nation</u>.'

6.57. 'n Mens moet net versigtig wees of <u>die kar se</u> wiele is in die lug. (corpus #1, Manuscript)

a human must just careful be or <u>the car PTCL.GEN</u> wheels be.PRS in the air

'One should be careful otherwise <u>the car's</u> wheels end up in the air.'

The use of *se* with inanimate possessors in the Afrikaans data remains low, even in corpus #4, at 9 percent of all interchangeable genitives with inanimate possessors, much lower than the overall proportion of *se* (35 percent in corpus #4). Temporal possessors are next, rising from *se* comprising 5 percent in corpus #1 to 27 percent in corpus #4. The use of *se* with locative possessors rises from 7 to 36 percent and with collectives from 2 to 37 percent—both are close to the overall proportion of *se*-genitives in corpus #4. This shows that collective, temporal, and locative possessors do behave differently in Afrikaans language use, as it does in English. The most significant change, however, regards the use of *se* with animate possessors (log likelihood of 489.00), illustrated in table 6.10.

The use of *se* as a proportion of all the interchangeable genitives with animate possessors rise from one-quarter to almost two-thirds—much higher than the overall proportion of *se* (9 to 35 percent). This shows that even though animacy does not impose an almost categorical restriction on the use

Table 6.10. Relative frequencies of *se* and *van* with animate possessors

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
se	24%	34%	58%	66%
van	76%	66%	42%	34%

of the different genitive variants, as it does in Dutch, it would probably be a significant factor in the choice for either of the two, as in English.

Possessor Definiteness

In Dutch, prenominal genitives are used only with proper names and forms of address (Scott 2011, 125; Weerman and De Wit 1999, 1167), although there is variation in some varieties of Dutch (Haegeman 2013, 221). In English, definiteness of the possessor NP is also a significant factor in genitive alternation (Rosenbach 2008, 154) to the point where human proper names take 99 percent *s*-genitives in one variety (Jankowski and Tagliamonte 2014, 331). The three categories of definiteness used in the analysis are *indefinite*, as in example 6.58, *definite* as in example 6.59, and *proper noun* as in example 6.60:

> 6.58.   My onthou is soos 'n skip se ruim . . . (corpus #4, Fiction)
> *my remember be.PRS like a ship PTCL.GEN bulge*
> 'My memory is like a ship's bulge . . .'
>
> 6.59.   Die hoofmanne sidder oor the ruwe plan van die koning . . .
> (corpus #1, Fiction)
> *the head.men shudder about the rough plan of the king*
> 'The chiefs shudder at the rough plan of the king . . .'
>
> 6.60.   Dat sy regtig sulke woorde uit Jan se mond hoor! (corpus #2,
> Fiction)
> *that she really such words out John PTCL.GEN mouth hear*
> 'The she's really hearing such words out of Jan's mouth!'

While there is definitely not an almost categorical restriction on the definiteness of the possessor in Afrikaans genitive constructions as there is in Dutch, its relevance in genitive alternation in other closely related languages, such as English, does make a case for investigating it in Afrikaans.

The use of *se* with indefinite possessors remains fairly infrequent, starting at 5 percent of all genitives with indefinite possessors and rising to only 16 percent in corpus #4. Regarding definite possessor NPs, *se* constitutes only 8 percent of the total in corpus #1 and rises to 30 percent in corpus #4, which is still a slightly smaller proportion than the overall use of *se*. In contrast, proper noun possessors show a sharp, significant rise in the already frequent use with *se* (log likelihood of 515.40), as shown in table 6.11.

In corpus #1, 15 percent of the genitive constructions with proper noun possessors are expressed by *se*, which rises to 55 percent in corpus #4. This indicates that the definiteness of the possessors contributes to the choice between *se* and *van*.

Table 6.11. Relative frequencies of *se* and *van* with proper noun possessors

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
se	15%	26%	38%	55%
van	*85%*	*74%*	*62%*	*45%*

Semantic Relations

Rosenbach (2014, 229) points out the difficulty surrounding semantic relations in genitives: "Certainly, the 'toughest nut' to define of all the factors in genitive variation is the semantic relation that holds between possessor and possessum." Scholars differ significantly in the relations they include and exclude and in how fine grained the distinction between different relations is (Conradie 2001; Nikiforidou 1991, 150; Rosenbach 2002, 60; 2014, 229). A distinction that has been used quite fruitfully is that between *prototypical* and *nonprototypical* relations (see, e.g. Grafmiller 2014, 477; Jankowski 2013, 85; Jankowski and Tagliamonte 2014, 315; Rosenbach 2002, 120–23). Prototypical relations include kin terms such as in example 6.61, body parts as in example 6.62, legal ownership as in example 6.63, and part/whole relations as in example 6.64; all others are grouped into nonprototypical relations as in example 6.65 (Rosenbach 2014, 229). In English, *s*-genitives occur more frequently with prototypical relations than with nonprototypical relations. Based on the other similarities between English and Afrikaans, I would expect to see *se*-genitives more frequently with prototypical relations than with nonprototypical relations:

> 6.61. Verskeie slagoffers, onder wie <u>Smit se seun</u> Koos, is aangerand. (corpus #4, News report)
> *several victim under who <u>Smit PTCL.GEN son</u> Koos be.AUX.PASS. PST assaulted*
> 'Several victims, among which <u>Smit's son</u> Koos, were assaulted.'
>
> 6.62. Ons <u>altwee se oë</u> is baie seer . . . (corpus #3, Manuscript)
> *us <u>both PTCL.GEN eyes</u> be.PRS very sore*
> 'Both of <u>our eyes</u> hurt badly . . .'
>
> 6.63. Dan gaan hulle saam na <u>die bruidegom se huis</u>, alwaar die bruilof gevier word. (corpus #1, Religious text)
> *then go they along to <u>the bridegroom PTCL.GEN house</u> where the wedding celebrated become.AUX.PASS.PRS*
> 'Then they go to <u>the groom's house</u> together, where the wedding is being celebrated.'
>
> 6.64. Seuns moet sorg dat die duim op <u>die naat van die broek</u> kom. (corpus #2, Informational text)

> *boys must ensure that the thumb on <u>the seam of the pants</u> come*
> 'Boys should ensure that their thumbs reach <u>the seam of the pants</u>.'

6.65. Gebooie is <u>die morele en wetlike basis van die Westerse beskawing</u>. (corpus #4, Religious text)
commandments be.PRS <u>the moral and legal basis of the Western civilization</u>
'Commandments are <u>the moral and legal basis of Western civilization</u>.'

The expectation is largely borne out by the Afrikaans data. In corpus #1, *se* is used with only 8 percent of the nonprototypical relations, climbing to 11 percent in corpus #2, 20 percent in corpus #3, and 30 percent in corpus #4, remaining lower than the overall proportion of *se* throughout. In contrast, *se* is used with 24 percent of the prototypical relations in corpus #1 already and with 70 percent by corpus #4, a change that is, unsurprisingly, statistically significant (log likelihood of 272.05). See table 6.12 for the details.

Table 6.12. Relative frequencies of *se* and *van* with prototypical semantic relations

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
se	24%	37%	59%	70%
van	76%	63%	41%	30%

To gain better insight into the spread of *se* with prototypical relations, it is worth looking at a more fine-grained analysis distinguishing the different prototypical relations specified above. Part/whole relations seem to be the one favoring *se* the least, starting out at 9 percent and ending with 35 percent, equivalent to the overall proportion of *se*. Kin terms are the second-smallest category of prototypical relations used with *se* but still much bigger than part/whole relations, rising from one-third in corpus #1 to two-thirds in corpus #4. Body parts also rise consistently, from 33 percent in corpus #1 to 89 percent in corpus #4 (the largest category in the last corpus). Legal ownership is the category with the largest proportion of *se*-genitives in corpus #1, at 36 percent, and it consistently rises to reach 85 percent in corpus #4. This shows that even though the binary distinction between prototypical and nonprototypical relations is useful in Afrikaans genitive alternation, the finer-grained distinction between different types of prototypical relations are worth the trouble because the use of *se* with kin terms, legal ownership, and body parts differs quite significantly from its use with part/whole relations.

There is an important different between the semantic relations kin terms, legal ownership, and body parts on the one hand and part/whole relations and nonprototypical relations on the other: the first group typically occurs with animate possessors by virtue of the type of relations it expresses, and the second group often occurs with inanimate possessors. This implies that the variable *semantic relation* might not function independently from the variable *possessor animacy*—this is also apparent in the statistical models discussed later.

Text Type

Some studies have shown the significance of situation or medium, genre, or text type in English genitive alternation (Grafmiller 2014; Heller, Szmrecsanyi, and Grafmiller 2017; Hinrichs and Szmrecsanyi 2007; Jankowski 2013; Jankowski and Tagliamonte 2014; Szmrecsanyi 2013). Typically in language change, the spread of a variant does not occur at the same speed or in the same way in different media (spoken or written) and text types, especially in some specialist text types or when (in)formality is concerned. Indeed, Vink (1981, 438) remarks that Afrikaans *se* has a more informal connotation and *van* a more formal one.

Like many diachronic studies, this one uses only written data, which is in many ways unfortunate. Furthermore, the relevant corpora almost exclusively represent Standard Afrikaans. However, the corpora still consist of entries from several different text types, which might give us at least some indication of how text type and/or formality interact with the choice between using *van* and *se* in Afrikaans genitives. The relative frequencies of *se* (compared to *van* for a specific section of a particular corpus) are given in table 6.13.

The text type with the smallest proportion of *se*-genitives to *van*-genitives throughout is texts from the natural sciences, starting at 1 percent and rising

Table 6.13. Relative frequencies of *se* compared to *van* in different text types

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
biographical texts	12%	16%	41%	41%
fiction	22%	40%	60%	60%
humanities	5%	12%	17%	21%
informational texts	9%	10%	22%	45%
manuscripts	16%	13%	39%	29%
natural sciences	1%	4%	5%	9%
news reports	3%	8%	12%	47%
religious texts	5%	6%	18%	34%
OVERALL PROPORTION	9%	13%	25%	35%

to a mere 9 percent in corpus #4. The other category of academic texts, from the humanities, is the second-smallest category in corpus #4 at 21 percent, which is still much higher than the natural sciences but also lower than the overall proportion of *se*. The usage in manuscript texts fluctuate throughout, which is not totally unexpected in the only unedited category in the corpus. The text type that consistently shows the largest proportion of *se*-genitives compared to *van*-genitives is fiction, starting at 22 percent in corpus #1 and rising to a stable 60 percent in corpus #3 and corpus #4, the only text type where *se* overtakes *van*. Interestingly, the combined frequency of *se*- and *van*-genitives per 100,000 words is the lowest in fiction in all the corpora except corpus #1, where fictional texts seem a little less genitive averse. So, while fiction does not contribute a particularly large portion of the overall uses of the genitive, the contribution it does make favors *se* over *van* quite significantly in contemporary usage.

News reports depart from typical English usage in an interesting way. Where newspapers lead the way in its usage of *s*-genitives in English (Grafmiller 2014, 483), its proportion of *se*-genitives compared to *van* remains one of the smallest in corpus #1 to corpus #3, suddenly jumping to 47 percent in corpus #4. In the latest corpus, news reports have the second-largest proportion of *se*-genitives next to fiction, more closely conforming to findings for English usage than the earlier corpora.

Syntactic Weight

In the choice between *se* and *van*, the possessor and possessum do not occur in the same order. The possessum precedes *van* and the possessor succeeds it, where the possessor precedes *se* and the possessum succeeds it. Because of this, it is expected that long possessums would favor the use of *se* and that long possessors would favor the use of *van*, as Vink (1981, 439) suggests. This is the case in English alternation, where long possessums favor the *s*-genitive and long possessors the *of*-genitive, as reported by several studies (Ehret, Wolk, and Szmrecsanyi 2014, 298; Jankowski and Tagliamonte 2014, 322; O'Connor, Maling, and Skarabela 2013, 104; Rosenbach 2005, 631).

An important question in determining the role of length or weight in binominal constructions such as the genitive is how to measure such length or weight. The number of stresses or letters have been used, but the number of words is also an efficient measure (as used in Ehret, Wolk, and Szmrecsanyi 2014, 266), and it poses fewer practical challenges in manual coding of data than other measures. Furthermore, Rosenbach (2014, 227) reports that there has been a very high correlation between results using number of words and other measures, such as number of syntactic nodes or phrases, so it seems safe

to use the number of words as a measurement of syntactic weight (Wasow and Arnold 2003, cited in Rosenbach 2014).

However, it is not merely the actual length of the possessor and possessum that is at stake but also the relative length comparing the possessor and the possessum, as implemented by Rosenbach (2005, 627) and Heller, Szmrecsanyi, and Grafmiller (2017, 17). In this section, I report briefly on the proportion of *se* versus *van* in terms of relative length. Importantly, the definite article *die* 'the' of the possessum in *van*-genitives was not counted (as in Hinrichs and Szmrecsanyi 2007, 453), as the equivalent with *se* would not include the article, similar to English (Heller, Szmrecsanyi, and Grafmiller 2017, 4; Rosenbach 2014, 224). The relative frequencies of *se* in the three categories of relative length are given in table 6.14.

Table 6.14. Relative frequencies of *se* (compared to *van*) regarding relative length

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
equal in length	12%	20%	32%	44%
longer possessor	8%	9%	18%	27%
longer possessum	15%	26%	38%	54%

In genitive constructions with possessors and possessums equal in length, the proportion of *se*-genitives compared to *van*-genitives is somewhat higher than the total proportion of *se*-genitives at 12 percent in corpus #1, climbing to 44 percent in corpus #4. This suggests that possessor length and possessum length might not equally contribute to the choice between *se* and *van*.

Genitive constructions containing possessors that are longer than the possessums overwhelmingly favor *van*-genitives, as could be expected, with a 92 percent proportion of *van*-genitives in corpus #1 falling to 73 percent in corpus #4, remaining higher than the overall proportion of *van* relative to *se*. It is also by far the largest category of the three, indicating that possessors tend to be longer than possessums in Afrikaans genitive constructions.

As expected, the constructions with possessors that are shorter than the possessums are more favorable to *se*-genitives, starting at 15 percent proportion of *se* relative to *van* in corpus #1, rising in each corpus to reach 54 percent, more than half, in corpus #4.

Regarding syntactically more complex possessors, I noted all the coordinated possessors in the data, as well as the possessors with postmodification, and very complex possessors, including, for example, coordination with postmodification or nested genitives. The proportion of *se*-genitives with postmodified possessors compared to *van* remains low, starting at 1 percent in corpus #1 and ending at 10 percent in corpus #4. Coordinated possessors

are less *se*-adverse, where the *se*-constructions constitute 3 percent of the total in corpus #1, rising steadily to 21 percent in corpus #4. These proportions remain lower than the overall proportion of *se*-genitives in all interchangeable contexts, which illustrates the importance of the end-weight principle, confirmed by the fact that *se*-genitives are never used with very complex possessors in the data despite of occurrences with *van* ranging from 27 to 40 per corpus (raw numbers).

Different Factors in Combination

In order to investigate how the different factors influence the choice between *se* and *van* together, I modeled it with the decision trees in the *partykit* package in RStudio (RStudio Team 2016), running the models with different combinations of independent variables to determine how the different variables contribute to the choice between *se* and *van*. I also modeled the data with logistic regression, using the *glm* command in RStudio and related processes.

While many others have used logistic regression to calculate the probability of different choices in genitive alternation (e.g., Ehret, Wolk, and Szmrecsanyi 2014; Grafmiller 2014; Heller, Szmrecsanyi, and Grafmiller 2017; Szmrecsanyi 2013, 2016), a decision tree has an important advantage of automatically taking interactions between variables into account. This is important while investigating alternations that might be influenced by different variables that might also interact with one another. For example, Rosenbach (2005) found that possessor length influences the choice between the *s*-genitive and *of*-genitive in English differently for animate and inanimate possessors. Interactions such as these are automatically accounted for in decision trees, while possible interactions have to be specified in logistic regression models.

A decision tree partitions the data—it first finds the variable that accounts for the largest majority of the variation and splits the data accordingly; each of the resulting branches is then divided further based on the values of other independent variables until all of the data are accounted for (Eddington 2010, 267). In this model, a variable that appears higher up in the tree influences the dependent variable more than the one below it (Eddington 2010, 269). Decision trees are also very powerful and able to handle very large data sets or unbalanced data sets with many variables (something that can cause problems for regression models), and they find generalizations that might not be apparent otherwise (Eddington 2010, 270).

Logistic regression is widely accepted as a method in studies where the response variable has two possible outcomes (Speelman 2014, 487), such as *se* and *van* in this study. Logistic regression analyses calculate the effects

of individual factors on the choice between two variants (Ehret, Wolk, and Szmrecsanyi 2014, 279), such as estimating the probability of a genitive being realized as either *se* or *van*, given certain conditions. Logistic regression analyses can give more fine-grained results than decision trees (Eddington 2010, 272), and the two can be used in a complementary way to get different perspectives on the same data.

After experimenting with including and excluding some of the factors as variables in the models, the following were included:

(i)   Animacy of the possessor, with the categories (a) animate, (b) collective, (c) locative, (d) temporal, and (e) inanimate

(ii)   Definiteness of the possessor, with the categories (a) indefinite, (b) definite, and (c) proper noun

(iii) Semantic relation, with the categories (a) nonprototypical, and prototypical split into (b) body parts, (c) kin terms, (d) legal ownership, and (e) part/whole

(iv)   Text type, according to the stratification in the corpora

(v)   Possessor length in number of words

All of the variables above occur fairly high up in at least one of the corpora's decision trees, but possessum length does not; regression analysis also showed that possessum length does not make a significant contribution to the choice between *se* and *van*; based on this, it was left out of the final models.

Corpus #1

In the decision tree of corpus #1, there is only one single decision branch that leads to a preference for *se* over *van*:

(a).   possessor animacy: animate › possessor length: ≤ 2 words › text type: fiction & manuscripts › possessor length: ≤ 1 word

This decision branch accounts for 77 instances, with a 36.4 percent error rate (meaning 36.4 percent of the 77 instances, or 28, are *van* instead of *se* within the same parameters). All the other end nodes (11 in total) lead to a preference for *van* with different levels of accuracy. One other branch has only a 53.6 percent accuracy in preference for *van*, but it still accounts for only 69 genitives. These findings are not unexpected when accounting for the small number of *se*-genitives in the corpus. The specific branch is visualized in figure 6.3.
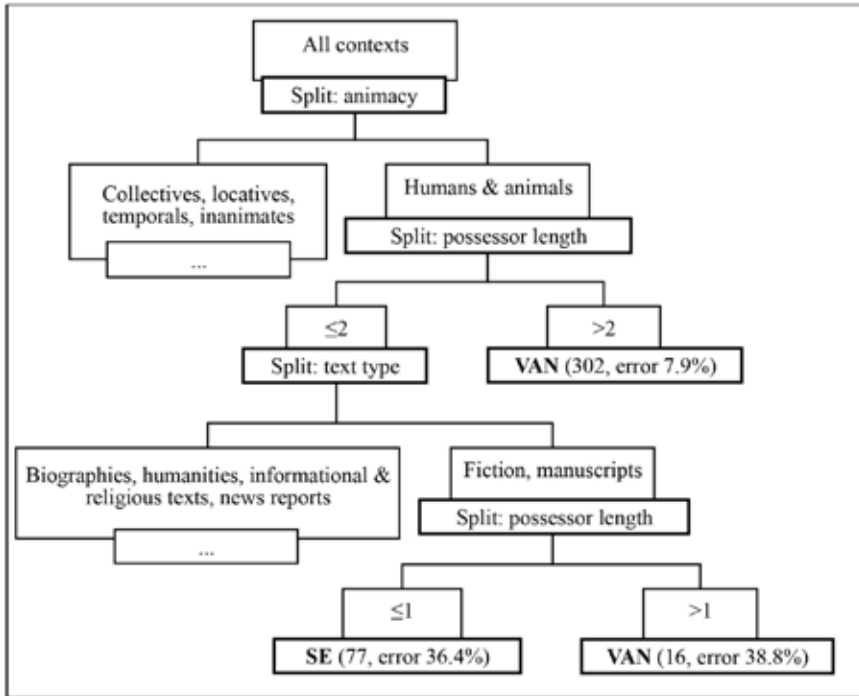
Figure 6.3.   Decision branch leading to *se* in corpus #1

This branch also illustrates how possessor length functions differently in certain contexts—possessors longer than two words typically take *van* even with animate possessors, and one-word possessors favor *se* only in certain text types.

The logistic regression model indicates that possessor animacy makes the most significant contribution to the choice of genitives (with $p < 0.001$), where animate possessors raise the probability of *se* being used quite significantly; this is followed by possessor length, where longer possessors tend to favor the use of *van*, and then text type, where the chances of using *se* are most probable in fiction. This corroborates the findings from the decision tree. Figure 6.4 shows an effects plot of animacy. The plot point for animate is much lower than the others, showing a preference for *se*. The plot point for inanimate is the highest, showing the highest preference for the use of *van*.

Figure 6.5 shows the effects plot of text type on genitive choice, with the plot point for fiction the lowest and that for manuscripts the second lowest, further confirming the findings of the decision tree above.
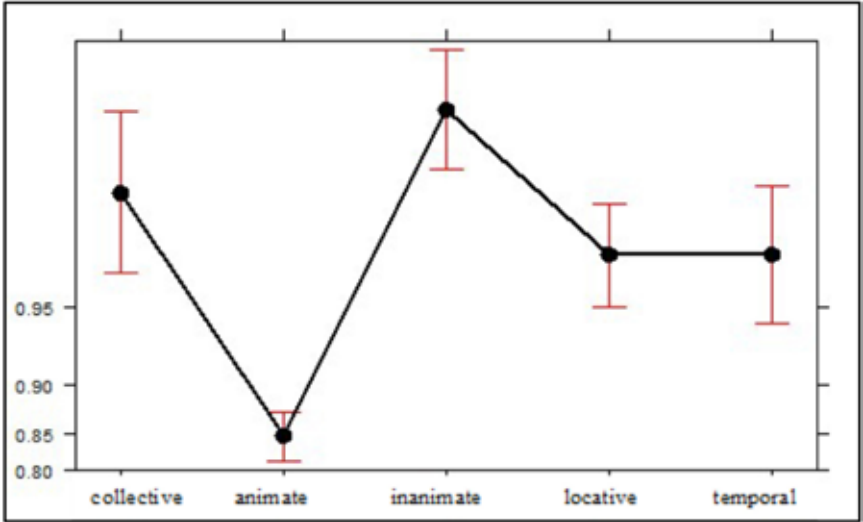
Figure 6.4. Corpus #1 effects plot of possessor animacy
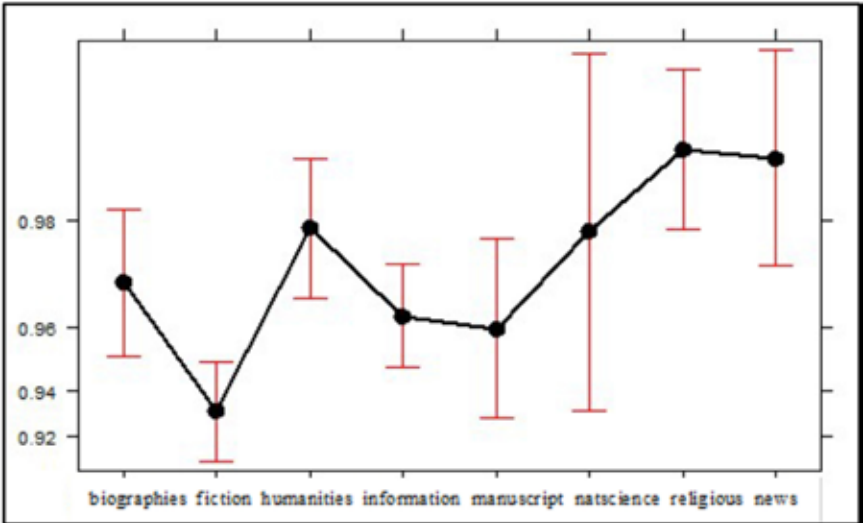
Figure 6.5. Corpus #1 effects plot of text type

Corpus #2

For corpus #2, the decision tree gets drastically more complex. Of the 22 end nodes, five lead to a preference for *se* rather than *van*. The first split, different from corpus #1, relates to the semantic relation: nonprototypical and part/whole relations to one side and body parts, kin terms, and legal ownership to the other. The latter branch leads to three end nodes, two of which favor *se*, both from the text type fiction. The other text types all favor *van*, even when used with the relevant semantic relations. This branch is visualized in figure 6.6.
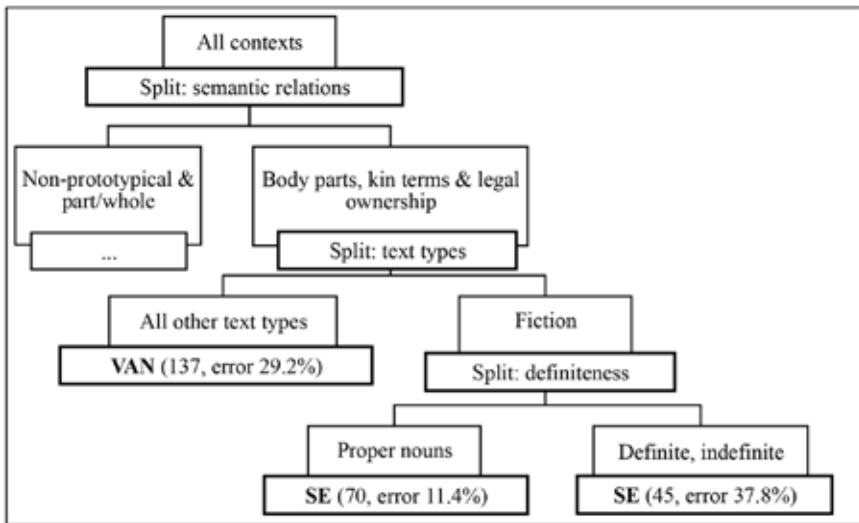
Figure 6.6.   Decision branch 1 leading to *se* in corpus #2

In the other branch from this split, there is one branch that leads to two nodes favoring *se*. The splits are as follows:

(b).   semantic relation: nonprototypical & part/whole › possessor length: ≤ 1 word › definiteness: definite & proper noun › possessor animacy: animate › text type: all excluding religious texts and news reports

This branch is represented in figure 6.7. In the final split, fiction favors *se* strongly (76 instances, 15.8 percent inaccuracy), while the other text types also favor *se*, accounting for 122 instances but with at 41 percent error rate. For genitives with longer possessors, the one context where *se* is still favored
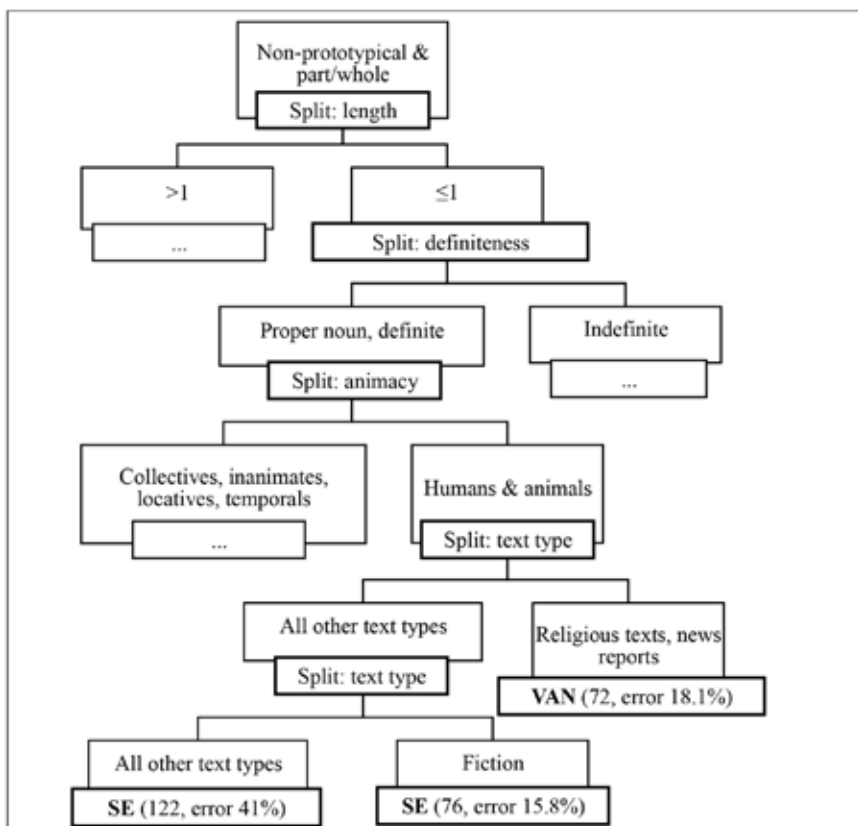
Figure 6.7. Decision branch 2 leading to *se* in corpus #2

slightly is with animate possessors, in fiction, if it is two-word possessors and not longer (71 instances, 42.3 percent error rate).

Here we can see how the tree shows the interactions of the different variables. In figure 6.6, certain semantic relations expressed in specifically the text type fiction lead to a preference for *se*. In the other branch, possessor length (one word) combined with definiteness (definites and proper nouns) and animacy (humans and animals) leads to a preference for *se* in all but two text types (although fiction still has the highest success rate).

Different from the decision tree, logistic regression analysis shows that the strongest predictor for genitive choice in corpus #2 is still possessor animacy, as in corpus #1—the decision tree indicates semantic relation as the strongest predictor. However, as I pointed out earlier, the groups of semantic relations split roughly according to animacy: in genitive relations that express kin terms, legal ownership, and body part relations, the possessors

are typically animate; in genitive relations that express part/whole relations and nonprototypical relations, the possessors are frequently inanimate. This shows that the effects of possessor animacy and semantic relation cannot be strictly separated and together form the strongest predictor of genitive choice. The effects plot of possessor animacy from the logistic regression analysis is shown in figure 6.8, and the animate point is, once again, the lowest plot point (strongest contributor to the choice for *se* at an even lower value than in corpus #1), and the highest point is that of inanimate possessors.
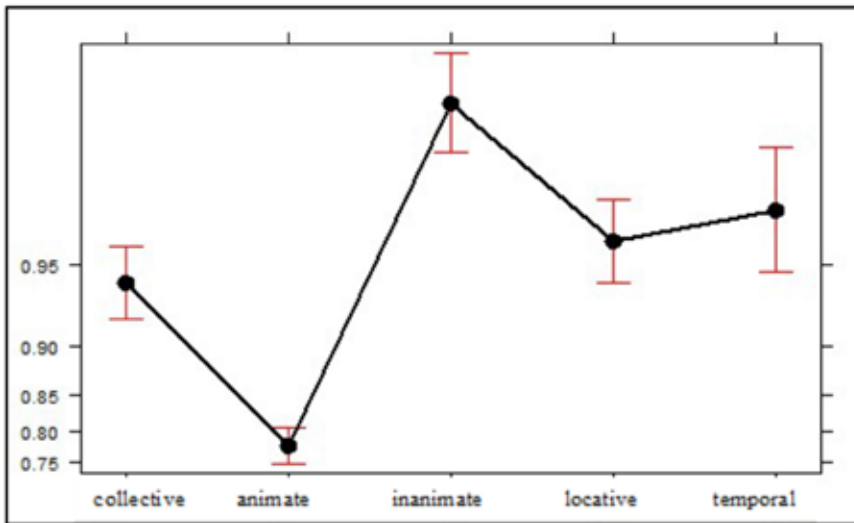
Figure 6.8.   Corpus #2 effects plot of possessor animacy

Otherwise, the logistic regression analysis confirms that all of the variables make a significant contribution in Afrikaans genitive choice, with *se* being favored most in fiction, with short, proper noun possessors and *van* being favored with longer possessors, especially in religious texts, similar to what the decision tree indicates.

Corpus #3

In corpus #3, the decision tree becomes even more complex, with 36 end nodes, 14 of which favor *se* over *van*. The first split is, like corpus #2, between the semantic relations body parts, kin terms, and legal ownership to the one side and nonprototypical and part/whole relations to the other. The former branch leads to four end nodes, three of which favor *se*, represented

in figure 6.9. Genitives in biographical texts, fiction, manuscripts, and natural sciences favor *se* regardless of possessor length but with a very low error rate for one-word possessors (3.1 percent, 128 instances) and a higher one for longer possessors (19.6 percent, 97 instances). Texts from the humanities, informational texts, religious texts, and news reports only slightly favor *se* with possessors of three words and shorter (96 instances, 40.6 percent error rate), while longer possessors favor *van* (13 instances, 23.1 percent error rate).
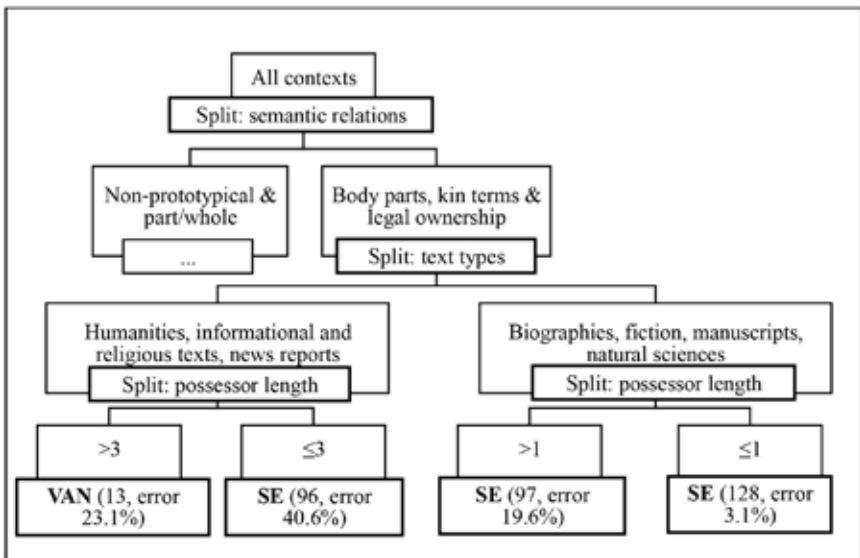
Figure 6.9.   Decision branch leading to *se* in corpus #3

Factors otherwise favoring *se* over *van* in different branches are possessor animacy (animate), text type (fiction mostly, manuscripts often, biographical texts occasionally, some others in very specific contexts), possessor length (shorter rather than longer), and possessor definiteness (proper nouns).

In the logistic regression analysis, possessor animacy is, once again, indicated as the strongest contributing factor to genitive choice. As is the case in corpus #2, the decision tree split with semantic relations also coincides with possessor animacy. Figure 6.10 shows the effects plot of possessor animacy, with the plot point for animate possessors even lower than before and that of inanimate possessors still the highest.

The rest of the results from the logistic regression analysis confirm that of the decision tree: text type makes a very strong contribution to genitive
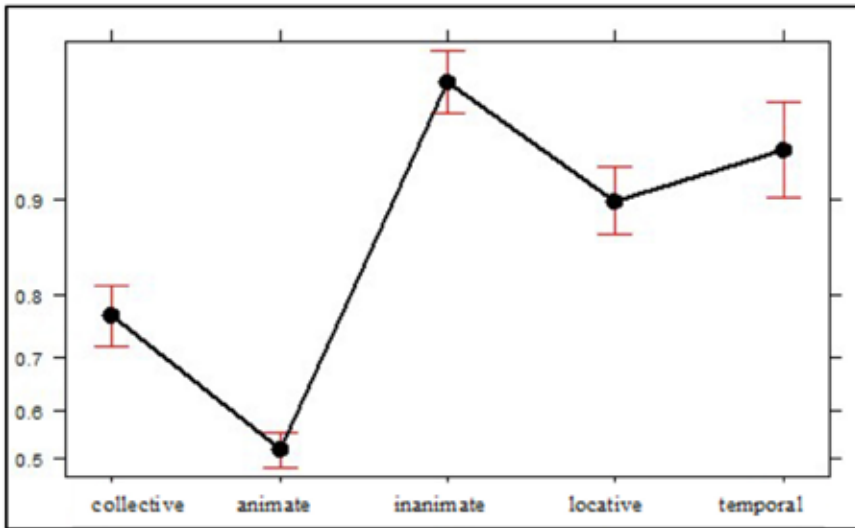
Figure 6.10.   Corpus #3 effects plot of possessor animacy

choice, followed by possessor length, and indefinite possessors favor *van* in particular.

Corpus #4

Finally, the decision tree modeling genitive alternation in corpus #4 becomes slightly less complex than that of corpus #3, with 25 end nodes, nine of which favor *se* above *van*. Once again, the first split is based on semantic relation, with body parts, kin terms, and legal ownership to one side and non-prototypical and part/whole relations to the other. The relations body parts, kin terms, and legal ownership favor *se* in all the end nodes (accounting for 372 instances in total with error rates between 0 and 22.2 percent) except for instances with possessors longer than three words. See figure 6.11 for in illustration of the branch.

The different ways certain variables influence choice in different circumstances become apparent in this decision branch as well. While all of the end nodes in this branch except one lead to *se*, the level of accuracy is influenced by possessor length and text type after a previous split for length and text type—one of the splits leads to the only *se*-branch with a perfect score.

Regarding the nonprototypical and part/whole relations, *se* is still favored with proper noun, animate possessors (257 instances with 9.3 percent error in fiction and biographical and informational texts and 303 instances with a
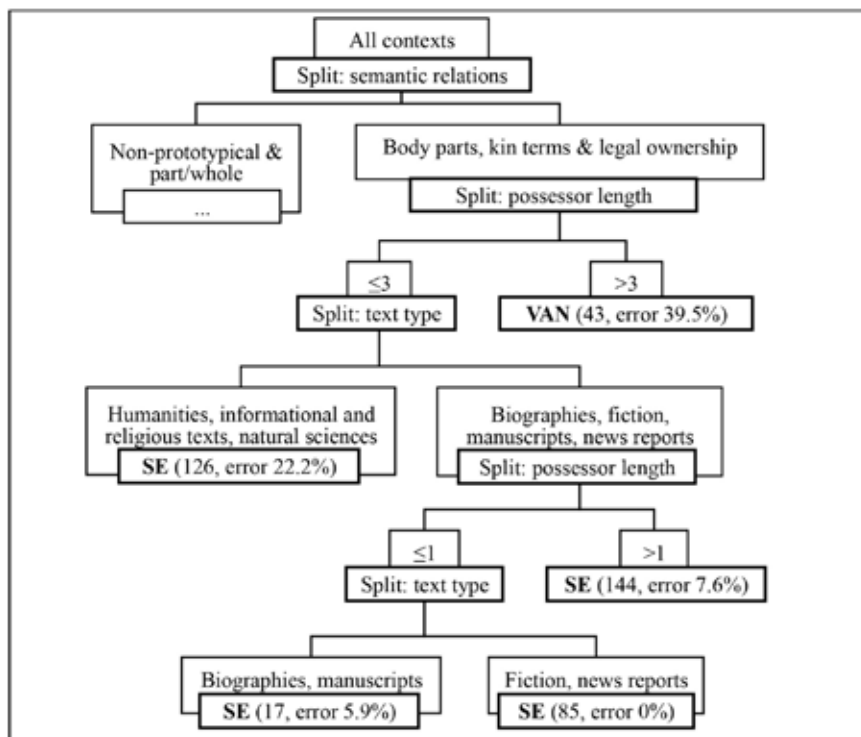
Figure 6.11.   Decision branch leading to *se* in corpus #4

high 47.2 percent error rate in the other text types) except for those possessors longer than three words. With inanimate proper noun possessors, the text type fiction leads to a 50 percent split between *se* and *van* in 10 instances. Otherwise, the factors favoring *se* are animacy (especially animate possessors but also collectives, locatives, and temporals in some contexts), text type (fiction, informational texts, manuscripts, religious texts, and news reports), and possessor length (two words if animate, one word if not).

As with all the other corpora, the logistic regression analysis indicates that possessor animacy is once again the strongest predictor of genitive choice. And, once again, the semantic relation split in the decision tree coincides with the animacy split in the regression analysis. The effects plot of animacy is given in figure 6.12, the lowest point yet for the animate plot point.

The new contribution of news reports in favor of *se* is also corroborated by the regression analysis as well as the tendency to prefer *van* in academic text types and religious texts.
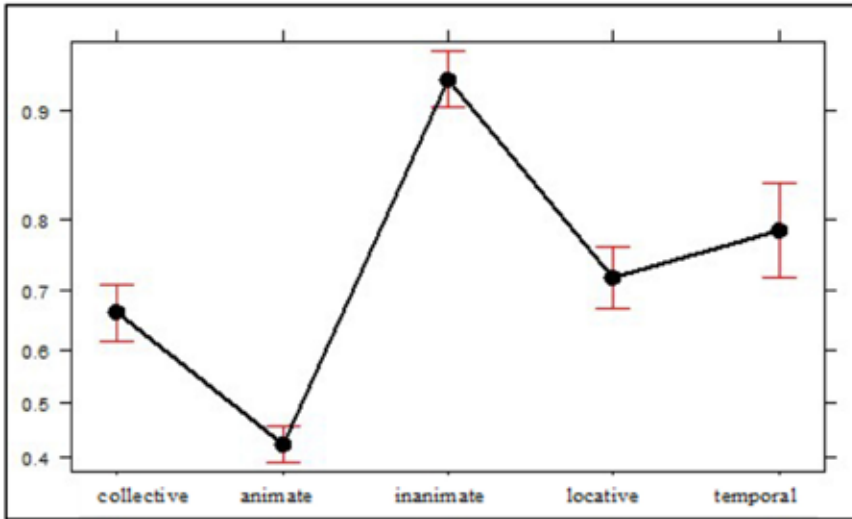
Figure 6.12.   Corpus #4 effects plot of possessor animacy

Brief Comparison with English

There are important differences between but also similarities with the findings of studies on English genitive alternation. Animate possessors are consistently one of the strongest factors encouraging the use of the *s*-genitive in English (e.g., Hinrichs and Szmrecsanyi 2007; Rosenbach 2005). It would seem to be similar in Afrikaans regarding the use of *se*, even if it so consistently coincides with semantic relations, blurring the lines somewhat. Possessor length also contributes to genitive choice in roughly the same way in English and Afrikaans.

An important difference between English and Afrikaans lies in the contributions of the different text types. Where news reports tend to significantly favor the *s*-genitive in English (Grafmiller 2014, 483), fiction makes a similar contribution to favoring the *se*-genitive in Afrikaans more than any other text type, with news reports favoring *se* more visibly only in the most recent data set.

Summary: Grammaticalization of *se*

The data in this study reveal two types of changes in Afrikaans genitive alternation since the early twentieth century, both pointing to the grammaticalization of *se*.

The one change can be described as the increasing use and spread of the *se*-genitive. Where the *se*-genitives (and related constructions) constitute only 9 percent of all the interchangeable genitives with the other 91 percent realized as *van*-genitives in the early twentieth-century data, it rises to 35 percent in the early twenty-first-century data. Fictional texts lead the way in this spread, consistently having the highest proportion of *se*-genitives throughout the century. The raw numbers of *se*-genitives and its proportion in the total number of interchangeable genitives rise fairly consistently throughout the century.

The other type of change the data point toward regards the factors that favor the use of *se*-genitives over *van*-genitives, or cases of context expansion. From the 1940s on, possessor animacy and semantic relation coincide strongly, and together they remain the most significant factors in genitive choice. However, there are some changes within certain variables in terms of which specific ones favor *se* or not.

First, while fiction consistently contributes to favoring *se* and manuscripts remain relevant, news reports start favoring *se* as well in corpus #4.

Regarding possessor length, only one-word possessors favor *se* under strict circumstances in corpus #1, while two-word possessors also favor *se* under specific circumstances in corpus #2. In corpus #3 and corpus #4, the distinction shifts to that between possessors longer than three words favoring *van* and shorter than or equal to three words favoring *se* in several branches of the decision trees. This means that while possessor length remains relevant for the choice between *se* and *van*, the cutoff length favoring *se* moved from one word to three in the course of the century—this could be an indication that weight is becoming a less important factor in genitive choice over time and/or that *se* is expanding further in usage context.

Possessor definiteness becomes significant only from corpus #2 on and varies in the strength of its contribution. It is consistently proper noun possessors that favor *se*, with indefinite possessors generally favoring *van*.

A last point about this section on the grammaticalization of *se*: I also showed how strong the interactions between certain variables can be. The combination of decision tree analysis and logistic regression analysis enabled us to see two different sides of the same coin and how possessor animacy and semantic relation correlate strongly in their contribution to genitive choice in Afrikaans.

CONCLUDING REMARKS: GRAMMATICALIZATION

There are several cases of ongoing grammaticalization in Afrikaans, some of which are subtle and small scale, while others are more significant develop-

ments. The more subtle cases include the increased fusion of *dit is* 'it is' to into *dis* 'it's,' the increased fusion of the quantifier *enig-* 'any-' with indefinite pronouns to form new ones such as *enigiemand* 'anybody' and *enigiets* 'anything,' and the erosion of the generic pronoun *'n mens* 'a human' to *mens*.

The more large scale developments of grammaticalization regard future reference and the genitive. The verb *gaan* 'go' shows clear signs of grammaticalization, including an increase in its use as a future auxiliary compared to its use as lexical verb and linking verb and an expansion in its context of use. There are similar developments for the genitive particle *se* ''s' in addition to its very significant increase in raw frequency and relative frequency compared to genitive *van* 'of.' Both *gaan* and *se* are in competition with alternatives: *gaan* with *sal* 'shall/will' and *se* with *van* 'of.' While the alternatives are still more frequently used in written Afrikaans, both *gaan* and *se* are steadily gaining a larger proportion of the functional load of future reference on the one hand and genitive expression on the other.

NOTE

1.   The section on future reference relies on research that I did for an article published in *Tydskrif vir Geesteswetenskappe* (see Kirsten 2018).

7

Discursive and Sociocultural Changes

FOCUS ON DISCURSIVE AND SOCIOCULTURAL CHANGES

Discourse can be conceptualized in many ways, but the relevant focus here is on language use above and beyond the syntactic level (Biber, Connor, and Upton 2007, 1). Items that signal certain relationships between different clauses or sentences (or paragraphs, sections, and so on) function on a discourse level to aid in creating a coherent text. Leech et al. (2009, 19) point out that it can also link the linguistic content with extralinguistic context, pointing to an even broader focus also attended to in this chapter.

Because language is typically used in social contexts, the immediate or broader social context can and does influence how language is used. In this chapter, I also explore how certain sociocultural developments and changes contribute to changing language use and how this is reflected in a century of written Afrikaans.

CONNECTIVES

When working with language use on a text or discourse level, items that connect and/or structure the discourse become particularly important. Two assumptions are relevant in this context: (1) discourse structure is complex, with hierarchical and linear relationships between different units and types of units, and (2) certain linguistic items are involved in marking this coherence structure of a discourse (Kroon 1997, 19). Coherence refers to how different parts of a text or discourse combine to form larger parts and eventually a whole structure (Redeker and Gruber 2014, 2).

135

The label *discourse marker* is often used to refer to a group of expressions that indicate in some way how a textual unit is integrated into the larger textual structure or context (Kroon 1997, 17), although other terms are also used, such as *discourse particle* or *pragmatic marker*. Discourse markers form a heterogeneous group, including any elements that perform other functions but are used as discourse markers when they do not form part of the propositional content of a sentence or utterance (Aijmer and Simon-Vandenbergen 2011, 227). When these items are used specifically to signal coherence relations, they are called *connectives* (Redeker and Gruber 2014, 6; Van der Vliet and Redeker 2014, 24).

There are three main categories of connectives: coordinators, subordinators, and conjunctive adverbials (Bosman and Otto 2012, 72; Kunz and Lapshinova-Koltunski 2014, 234; Messerschmidt and Messerschmidt 2006, 10). In Afrikaans, there is a structural difference between coordinators and subordinators, where the typical SVX order is maintained when using a coordinator, but it is inverted (SXV) when using a subordinator (Messerschmidt and Messerschmidt 2006, 10; Van Rooy and Kruger 2016, 104). This makes the distinction between the two types of conjunctions fairly straightforward: coordinators connect two main clauses, while subordinators connect a subordinate clause (with the subordinate SXV order) with a main clause. There are other important differences as well: the order of occurrence with coordinators and conjunctive adverbials is fixed, while it can be changed with subordinators (Messerschmidt and Messerschmidt 2006, 10); the syntactic connection between two clauses is also tighter when using a subordinator than a coordinator or conjunctive adverbial (Messerschmidt and Messerschmidt 2006, 11).

Another way to distinguish different types of connectives refers to the type of relations they express between the clauses they connect (Kunz and Lapshinova-Koltunski 2014, 233):

- Additive connectives, linking two events that are true (or not) at the same time, such as *en* 'and'
- Adversative connectives, contrasting two events or giving an alternative, such as *tog* 'yet' or *alhoewel* 'although'
- Causal connectives, relating cause and effect or dependence between events, such as *want* 'because' or *daarom* 'that's why'
- Temporal connectives, indicating the temporal relation between events, such as *terselfdertyd* 'at the same time' or *daarna* 'afterwards'

Fouché, Van den Berg, and Olivier (2017, 834) add a category of Afrikaans connectives: connectives signaling a projected complement clause, with two entries, *dat* 'that' and *of* 'if.'

It is important to distinguish the use of certain connective items as something other than *discourse* markers—Stede (2014, 126) refers to the terms *discourse reading* and *sentential reading*, where the former refers to an element that contributes to discourse interpretation and the latter to sentence meaning. An example of this is the Afrikaans connective *en* 'and,' which, like its English counterpart, can connect two independent clauses (7.1), but also, for example, items in a list within a clause (7.2).

> 7.1. Idalette het weer gelaai, <u>en</u> vir die res van die nag het die klein huiskring brandwag gesit. (corpus #2, Fiction)
> *Idalette have again loaded, <u>and</u> for the rest of the night have the small house.circle guard sat*
> 'Idalette loaded again, <u>and</u> for the rest of the night, the small house circle stood guard.'
> 7.2. Sy is klein <u>en</u> fyn gebou. (corpus #4, Informational text)
> *she be.PRS small <u>and</u> finely built*
> 'She has small <u>and</u> delicate build.'

It also happens that certain connectives can express more than one relation, sometimes with a level of ambiguity where either meaning is plausible or perhaps both in parallel (Stede 2014, 127).

Because connectives are typically not obligatory and there are often several that can express similar relations between clauses, it is an area of language use that could change quite significantly in the course of a century. Importantly, the use (frequency, choice of particular items, etc.) of connectives is tied quite closely to genre or text type, particularly where formality and informality are involved (Redeker and Gruber 2014, 7).

of to Signal Complement Clauses

There are two subordinators in Afrikaans used to signal a projected complement clause: *dat* 'that' and *of* 'if.' In the Afrikaans literature, there are different terms and descriptions for these items: subordinate conjunctions (Carstens 2018, 53), bare subordinators (Ponelis 1979, 439), and structure words (Combrink 1995, 40). A description of these terms is that they signal the start of a complement clause and link the complement clause to a main clause (Fouché, Van den Berg, and Olivier 2017, 830–31). Both *dat* and *of* are involved in projection, where the interpretation of the main clause is dependent on the complement clause that follows (Fouché, Van den Berg, and Olivier 2017, 831). The use of *dat* typically signals a statement (7.3), and *of* signals a question or some kind of inquiry (7.4). Other than that, the terms

are fairly neutral and semantically underspecified (Fouché, Van den Berg, and Olivier 2017, 830):

7.3.   Die koloniste het geweet <u>dat</u> hij hul belange op die hart gedra het . . . (corpus #1, Biographical text)
the colonists have knew <u>that</u> he their interests on the heart carried have
'The colonists knew <u>that</u> he took their interests to heart . . .'

7.4.   Hy wonder <u>of</u> Ann ooit sal sin hê in die eenvoudige ou plekkie. (corpus #3, Fiction)
he wonder <u>if</u> Ann ever will sense have in the simple old place.
<small>DIM</small>
'He wonders <u>if</u> Ann would ever like this simple old place.'

This section focuses on the use of *of* in the course of the twentieth and early twenty-first centuries.

The item *of* has two distinct functions—one can be translated as 'if' (7.5), and one can be translated as 'or' (7.6):

7.5.   Die een groot vraag is natuurlik <u>of</u> ek in die onmiddellike toekoms die SVB-werk mag los. (corpus #2, Manuscript)
the one big question be.<small>PRS</small> naturally <u>if</u> I in the immediate future the SVB-work may leave
'The one big question is of course <u>if</u> I can leave the SVB-work in the immediate future.'

7.6.   Moenie enige verpligting aanvaar <u>of</u> beloftes maak wat jy nie kan hou nie. (corpus #4, Religious text)
must.not any responsibilities accept <u>or</u> promises make that you.
<small>SG</small> *not can keep* <small>PTCL.NEG</small>
'Don't accept responsibilities <u>or</u> make promises that you cannot keep.'

The frequencies of use per 100,000 words in the consecutive corpora are given in table 7.1.

Table 7.1.   Frequencies of *of* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
of connective	93	72	68	60
of alternative	391	317	311	327
TOTAL	484	389	379	387

The use of *of* to indicate alternatives drops significantly from corpus #1 to corpus #2 (log likelihood of 18.99) but then fluctuates without showing change in a particular direction. On the other hand, the use of *of* as a connective continues to decline from corpus #1 right through to corpus #4, amounting to a significant change (log likelihood of 18.96). As I mentioned earlier, the use of connectives is closely linked to different text types. It might be useful, then, to explore the use of connective *of* in the different text types in the corpus in more detail. The frequencies per 100,000 words for each text type are given in table 7.2.

Table 7.2. Frequencies of *of* as connective per text type

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
biographical texts	39	84	84	59
fiction	147	158	108	91
humanities	40	33	67	20
informational texts	88	31	70	58
manuscripts	191	136	81	122
natural sciences	43	33	23	20
news reports	48	20	15	72
religious texts	44	29	35	20

The frequency of connective *of* is not particularly high, so the spread of its use is quite erratic in many of the text types. Natural sciences show a consistent decline in use, and religious texts have a bit of a zigzag but generally also show decline in the course of the century. However, there is no clear general pattern discernable from the data in table 7.2. An alternative fine-grained analysis of the use of connective *of* regards its use in different constructions and with different types of verbs in the matrix clause.

A slightly archaic use of *of* as connective is similar to what *asof* 'as if' is used for today, in a comparative construction (7.7). *Of* is also occasionally used as a *coordinator*, not a *subordinator*, in constructions that indicate a more temporal or consecutive meaning (7.8). The subordinator *of* is sometimes used to signal a clause that is not directly embedded into a matrix clause, used similarly to English 'whether' (7.9). The other, more typical uses of *of* as a subordinator can be categorized according to the type of verb in the matrix clause, including concrete verbs (7.10), communication verbs (7.11), mental verbs (7.12), abstract verbs (7.13), and copular verbs (7.14):

7.7.   Ons moet eenvoudig Afrikaans praat, net <u>of</u> dit so behoort te wees, en dit behoort ook so. (corpus #1, Informational text)

*we must simply Afrikaans speak just <u>if</u> it so supposed to be and
it supposed also so*
'We just have to speak Afrikaans, just <u>as if</u> it is supposed to be
like that, and it is.'

7.8.   Hy is skaars verby, <u>of</u> die slang staan met 'n groot geblaas
        orent. (corpus #3, Biographical text)
        *he be.PRS barely passed <u>or</u> the snake stand with a great hissing
        up*
        'He barely passed, <u>or</u> the snake rises up with a big hiss.'

7.9.   Die duiwel is daar <u>of</u> ons nou al aan hom glo of nie. (corpus #1,
        Religious text)
        the devil be.PRS there <u>if</u> we now already in him believe or not
        'The devil is there <u>whether</u> we believe in him or not.'

7.10.   Daar aangekom, **proe** hy eers <u>of</u> die brandewyn goed is. (cor-
        pus #2, Fiction)
        *there arrived **taste** he first <u>if</u> the brandy good be.PRS*
        'Arriving there, he **tastes** <u>if</u> the brandy is good.'

7.11.   . . . dus kan daar nie **gesê** word waar dit is nie en <u>of</u> dit in 'n
        werkende toestand is nie. (corpus #4, Natural sciences)
        *so can there not **said** become.AUX.PASS.PRS where it be.PRS PTCL.
        NEG and <u>if</u> it in a working condition be.PRS PTCL.NEG*
        .'.. '. . . so there cannot be **said** where it is and <u>if</u> it is in work-
        ing order.'

7.12.   Sy wil **weet** <u>of</u> ek al negeduisend druppels opium ingeneem
        het. (corpus #3, Fiction)
        *she want **know** <u>if</u> I already nine.thousand drops opium taken
        have*
        'She wants to **know** <u>if</u> I have ever taken nine thousand drops
        of opium.'

7.13.   In hierdie oordeel **maak** dit nie **saak** <u>of</u> 'n persoon uit 'n be-
        voorregte agtergrond kom met uitstekende opleiding nie. (cor-
        pus #4, Informational text)
        *in this judgment **make** it not **matter** <u>if</u> a person out a privileged
        background come with excellent education PTCL.NEG*
        'In this judgment it doesn't **matter** <u>if</u> a person comes from a
        background of privilege with exceptional education.'

7.14.   Dit **is** egter vir my 'n vraag <u>of</u> ek iets meer as blote vrindskap
        van jou mag verwag. (corpus #2, Manuscript)
        *it **be.PRS** rather for me a question <u>if</u> I something more than
        bare friendship from you.SG may expect*

> 'It **is**, however, a question for me <u>if</u> I may expect more than
> just friendship from you.'

The frequencies per 100,000 words of each of these types of uses can be seen
in table 7.3.

Table 7.3.   Frequencies of types of connective *of*

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
comparative	8	5	1	0
coordinator	8	4	2	2
quasi-independent	8	9	3	3
concrete verb	5	8	9	12
communication verb	16	9	8	10
mental verb	20	17	22	15
abstract verb	4	2	4	2
copular verb	24	18	19	15

Three of these types of connective *of* show significant decline: the com-
parative uses (log likelihood of 34.74), coordinator uses (log likelihood of
15.70), and uses in copula constructions (log likelihood of 16.06). As the
overall frequency of connective *of* is declining, such developments are hardly
unexpected—it does, however, show that certain uses are falling into disuse
quicker than others. Comparative and coordinator uses become marginal by
the twenty-first century, and the pattern for the independent uses is not so
clear but develops in roughly the same direction. The use of connective *of* in
copula constructions, however, remain one of the largest categories, even if it
declines from corpus #1 to corpus #2 and from corpus #3 to corpus #4. This
probably relates more directly to the general decline in use of connective *of*,
being one of the more typical categories.

A development in direct contrast to these is the increasing use of connec-
tive *of* with concrete verbs (log likelihood of 8.29), even when other uses
show signs of decline. Usage with the other dynamic verbs either fluctuate
throughout (mental and abstract verbs) or remain stable after standardiza-
tion (communication verbs). The pattern does seem to suggest that even as
connective *of* declines in frequency, its profile is undergoing subtle changes,
where the remaining uses are with dynamic verbs more than anything else
and with copular verbs. The less typical uses continue to decline.

Because of the broad nature of the connection expressed by connective *of*,
it is quite possible that the decline in overall usage is mirrored by an increase
in bare complement clauses. The frequency of *dat* 'that,' as described in the
following section, adds to this suspicion.

dat and -dat Compounds

In this section, the focus is on *dat* 'that' and compounds formed with *-dat* as the second part. As I already mentioned above, *dat* is the main complementizer in Afrikaans. Statement complement clauses can be introduced with *dat*, in which case they are subordinate clauses with the SXV word order, or they can be formed without a particular connective, in which case the typical SVX order of Afrikaans remains (Van Rooy and Kruger 2016, 104). The bare form *dat* is the most neutral subordinator in Afrikaans, and its function is to introduce a statement complement clause (Messerschmidt and Messerschmidt 2006, 9), as illustrated in example 7.15:

> 7.15.   Die merkwaardige is <u>dat</u> die versamelaar weinig daarvan
>   gekoop het. (corpus #3, News report)
>   *the remarkable be.PRS <u>that</u> the collector little thereof bought*
>   *have*
>   'The remarkable part is <u>that</u> the collector bought little of it.'

The frequencies of *dat* are as follows:

- Corpus #1: 860 per 100,000 words
- Corpus #2: 885 per 100,000 words
- Corpus #3: 791 per 100,000 words
- Corpus #4: 707 per 100,000 words

While there is a slight rise in frequency between corpus #1 and corpus #2, the decline from corpus #2 through corpus #4 is consistent and significant (log likelihood of 50.26). As with *of*, it seems that complement clauses are increasingly being used without a subordinator.

A study by Van Rooy and Kruger (2016) undertakes a statistical analysis of the different factors that contribute to a complement clause taking *dat* or being bare (omission of *dat*). The most important contributing factors are specific lexical verbs used in the main clause and the text type in which the sentence is found (Van Rooy and Kruger 2016, 114). Communication and mental verbs with a particularly high frequency, such as *sê* 'say,' *dink* 'think,' and *weet* 'know,' favor the omission of *dat*, while verbs with a lower frequency favor the use of *dat* (Van Rooy and Kruger 2016, 114). Regarding text types, the more informal texts types of fiction, news reports, and magazines favor the omission of *dat*, while the more formal text types of academic texts and published popular books favor the use of *dat* (Van Rooy and Kruger 2016, 114).

A different context of the element *dat*, however, shows in increase in frequency—its use in morphological compounds that form connectives, where -*dat* is the second morpheme (Messerschmidt and Messerschmidt 2006, 8). The first morphemes in these compounds are often prepositions, such as *na-dat* 'after.that,' *totdat* 'until.that,' *deurdat* 'through.that,' and so on, but they can also be adverbs, such as *noudat* 'now.that,' *sodat* 'so.that,' and *hoedat* 'how.that' (Messerschmidt and Messerschmidt 2006, 9). Similar to bare *dat*, -*dat* compounds are also subordinators, used to incorporate a subordinate clause into a matrix clause (Messerschmidt and Messerschmidt 2006, 10).

To show the extent of the compounds, I give an example of each of the different forms in my corpus data (7.16–7.28):

7.16. Dat sy werklik liggaamlik in die lewe teruggebring is, word bevestig <u>deurdat</u> sy iets kry om te eet. (corpus #3, Religious text) ['**through.that**,' method]
that she really bodily in the life back.brought be.AUX.PASS.PST become.AUX.PASS.PRS confirmed <u>through.that</u> she something get PTCL.INF to eat
'That she is really brought back to life bodily, is confirmed <u>through</u> her receiving something to eat.'

7.17. En <u>doordat</u> hij sig bewus is van wat hij doet het hij s'n onderskeidings vermogen geoefen. (corpus #1, Informational text) ['**through.that**,' method, obsolete spelling]
and <u>through.that</u> he him.REFL aware be.INF of what he do have he his discernment ability practiced
'And <u>through</u> him being aware of what he's doing, he practiced his discernment.'

7.18. As u in die koerante lees <u>hoedat</u> die wetteloosheid alreeds sy opmars in ons vaderland loods . . . (corpus #3, News report) ['**how.that**,' method]
if you.HON in the newspapers read <u>how.that</u> the lawlessness already his up.march in our fatherland launch
'If you read in the papers <u>how</u> the lawlessness already marches into our father land . . .'

7.19. 'n Gevoel van meerderwaardigheid ontstaan by die leser, <u>om-dat</u> sy vermoedens bevestig word. (corpus #4, Humanities) ['**to.that**,' reason]
a feeling of superiority arises by the reader <u>to.that</u> his suspicions confirmed become.AUX.PASS.PRS
'A feeling of superiority arises in the reader, <u>because</u> his suspicions are confirmed.'

7.20. Ik zal hem echter uw naschrift ter hand stellen, <u>opdat</u> hij in de toekomst voorzichtig zijn zal. (corpus #1, Manuscript) ['**on.that**,' consequence]

I shall him rather your.HON postscript in hand put <u>on.that</u> he in the future careful be.PRS shall

'I will, however, give him your postscript, <u>so that</u> he will be careful in the future.'

7.21. Sy Skrifkennis was deeglik, gesond en suiwer, <u>sodat</u> predikante wat voor hom opgetree het, baie deeglik in ag geneem het wat hulle sê. (corpus #2, Biographical text) ['**so.that**,' consequence]

his Scripture.knowledge be.PST thorough healthy and pure <u>so.that</u> ministers that before him acted have very thorough in consideration took have what they say

'His knowledge of Scripture was thorough, healthy and pure, <u>so that</u> ministers that acted in front of him, took great care in what they said.'

7.22. Laat ons betyds weet <u>zodat</u> ons jou kan afhaal. (corpus #1, Manuscript) ['**so.that**,' consequence, obsolete spelling]

let us in.time know <u>so.that</u> we you.SG can down.pick

'Let us know in time <u>so that</u> we can pick you up.'

7.23. Maar <u>eerdat</u> dit kon geskied, toe gebeur daar 'n jammerlike ding . . . (corpus #3, Manuscript) ['**before.that**,' temporal]

but <u>before.that</u> it could happen then happened there a sorry thing

'But <u>before</u> it could happen, there happened a sorry thing . . .'

7.24. Dit duur 'n oomblik <u>voordat</u> sy kan glo wat sy gehoor het. (corpus #4, Fiction) ['**before.that**,' temporal]

it take a moment <u>before.that</u> she can believe what she heard have

'It takes a moment <u>before</u> she can believe what she heard.'

7.25. Die res van die strome sluit aan <u>nadat</u> die Magaliesrivier en die Swartspruit op Hekpoort 122 verenig het. (corpus #2, Natural sciences) ['**after.that**,' temporal]

the rest of the streams lock on <u>after.that</u> the Magaliesrivier and the Swartspruit on Hekpoort 122 converged have.

'The rest of the streams connect <u>after</u> the Magaliesriver and the Swartspruit converged on Hekpoort 122.'

7.26. Kyk, Willem, <u>noudat</u> jy so gestyg het in my agting, kan jy gerus wees. (corpus #2, Fiction) ['**now.that**,' temporal]

look Willem <u>now.that</u> you.SG so rose have in my regard can

you.SG at.ease be

'Look, Willem, <u>now that</u> your esteem has risen so much in my eyes, you can be at ease.'

7.27. Sy het nog elke jaar <u>vandat</u> sy in Kraaifontein werk na haar ma toe gekom, sy sal die Karoo ook mis. (corpus #4, Fiction) ['**from.that**,' temporal]

she have still every year <u>from.that</u> she in Kraaifontein work to her mom to came she will the Karoo also miss

'Every year <u>since</u> she's been working in Kraaifontein, she came to her mother, she will miss the Karoo as well.'

7.28. Hy wil probeer om eerlik te bly, <u>totdat</u> hy die dag sterf. (corpus #2, Fiction) ['**until.that**,' temporal]

he want try PTCL.INF honest to remain <u>until.that</u> he the day die

'He wants to try to remain honest, <u>until</u> the day he dies.'

The process of forming compounds with *-dat* seems to be somewhat productive—an additional compound, *indat* 'in.that,' is not present in the data or in dictionaries, but it can be found occasionally in Afrikaans language use (Messerschmidt and Messerschmidt 2006, 16). The frequencies of the different *-dat* compounds are given in table 7.4.

In many cases, the individual frequencies cannot be interpreted in a meaningful way because they are too low and the corpora too small. There are two observations, however. First, a few items disappeared in the course of the century (or became too infrequent to occur in corpus #4 at least), such as *eerdat* 'before.that' and *opdat* 'on.that,' and the old spellings of *doordat*

Table 7.4. Frequencies of *-dat* compounds per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
deurdat ('through.that')	3	6	5	3
doordat ('through.that')	<1	0	0	0
hoedat ('how.that')	2	3	3	3
omdat ('to.that')	72	83	83	83
opdat ('on.that')	2	0	0	0
sodat ('so.that')	37	38	53	50
zodat ('so.that')	2	0	0	0
eerdat ('before.that')	<1	<1	<1	0
voordat ('before.that')	10	23	17	30
nadat ('after.that')	13	28	22	42
noudat ('now.that')	<1	2	2	<1
vandat ('from.that')	1	2	2	3
totdat ('until.that')	23	23	17	12
TOTAL	166	209	204	225

'throught.that' and *zodat* 'so.that' fell into disuse. Second, the total frequency of *-dat* compounds increased significantly in the course of the century, even if certain items and spellings disappeared (log likelihood of 5.28, indicating low-level significance). To gain more meaningful insight into the use of *-dat* compounds, I give the frequencies of the different types of meaning expressed by the compounds in table 7.5.

Table 7.5.   Frequencies of types of *-dat* compounds per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
method (*deurdat, doordat, hoedat*)	5	9	8	5
temporal (*eerdat, voordat, nadat, noudat, vandat, totdat*)	48	79	60	87
reason (*omdat*)	72	83	83	83
consequence (*sodat, zodat*)	41	38	53	50

    None of these categories show a clear trajectory of increase. The method compounds remain fairly low in frequency, the temporal and consequence compounds show variability but roughly point to an increase in usage, and *omdat* rises in frequency from corpus #1 to corpus #2 and then remains completely stable in the data. Importantly, while there is a small decline in the number of different forms used, the overall frequency increases. This shows that the popularity of particular compounds continues to increase in broad strokes, like *sodat*, *voordat*, and *nadat*, even if the compounding process does not serve to supply new forms at the same rate that old forms fall into disuse.

    Finally, even though the overall frequency of *-dat* compounds continues to rise, it is not sufficient to even out the loss of bare *dat* reported earlier in the section. Together with the decline of *of* 'if,' it would seem that the bare complement clause is on the increase in Afrikaans, while the use of complementizers is declining.

Selected Further Connectives

    In this final section on connectives, I aim to give a broad overview on changes in the general use of one-word connectives in written Afrikaans. As a starting point, I took the categories in Carstens (1997) and Fouché, Van den Berg, and Olivier (2017) of what they call *konjunksiemerkers*, which directly translates to 'conjunctive markers' or, as I call it, connectives. In this section, I do not take the connectives from the previous sections into account.

    The first category is that of additive connectives, performing a fairly straightforward linking function. The items I investigated are *boonop*

'moreover,' *bowendien* 'moreover,' *buitendien* 'besides,' *byvoorbeeld* 'for example,' *daarbenewens* 'in addition,' *daarby* 'in addition,' *daarnaas* 'next,' *dienooreenkomstig* 'accordingly,' *naamlik* 'namely,' *naas* 'next,' *ook* 'also,' *verder* 'further,' *vervolgens* 'furthermore,' and *voorts* 'further.' Several of these words are also used with other meanings or functions, in which case I identified the connective uses and excluded the rest. Examples of how each one is typically used are given in examples 7.29 to 7.42:

7.29. <u>Boonop</u> het hy nie minder nie as 15 kongresse bygewoon. (corpus #3, News report)
<u>moreover</u> have he not less PTCL.NEG than 15 conference attended
'<u>Moreover</u>, he attended no less than 15 conferences.'

7.30. <u>Bowendien</u> is hy die taak opgelê om die tussenkoms van die Europese regerings te probeer verkry om vrede te bewerkstellig, met behoud van hul onafhanklikheid. (corpus #4, Biographical text)
<u>moreover</u> be.AUX.PASS.PST he the task on.laid PTCL.INF the between.coming of the European governments to try get PTCL.INF peace to effect with maintenance of their independence
'<u>Moreover</u>, he was given the task to try and obtain the intervention of the European governments to effect peace, while maintaining their independence.'

7.31. Suid-Afrika is <u>buitendien</u> geen provinsie van Nederland nie, en sy belange en toestande is totaal verskillend . . . (corpus #2, Biographical text)
South-Africa be.PRS <u>besides</u> no province of Netherland PTCL.NEG and his interests and conditions be.PRS totally different
'South Africa is, <u>besides</u>, no province of the Netherlands, and its interests and conditions are completely different . . .'

7.32. Mantashe self het <u>byvoorbeeld</u> regters as "teen-revolusionêre" verdag probeer maak. (corpus #4, News report)
Mantashe self have <u>for.example</u> judges as "against-revolutionary" suspicious try make
'Mantashe himself has <u>for example</u> attempted to draw suspicion to judges as "anti-revolutionary."'

7.33. <u>Daarbenewens</u> veroorsaak dit 'n gevoel van euforie of 'n verhoogde gevoel van welsyn. (corpus #3, Informational text)
<u>in.addition</u> cause it a feeling of euphoria or a heightened feeling of well.being

'<u>In addition</u>, it causes a feeling of euphoria and a heightened feeling of well being.'

7.34. Gert Markeland was 'n reus van iets oor die sewe voet lengte; <u>daarbij</u> was hij eweredig gebouwd, met 'n gemiddelde gewig van 325 pond; was so regop as 'n kers en so sterk as 'n os. (corpus #1, Fiction)

Gert Markeland be.PST a giant of something over the seven feet length; <u>there.with</u> be.PST he evenly built with an average weight of 325 pound be.PST so upright as a candle and so strong as an ox

'Gert Markeland was a giant of something over seven feet in length; <u>in addition</u> he was built evenly, with an average weight of 325 pounds; was as upright as a candle and as strong as an ox.'

7.35. <u>Daarnaas</u> stel Keller die teleurstelling wat groot wereld-bewe-ginge soos die Volkebond, die Marxistiese beweging en die demokrasie gebring het. (corpus #2, Religious text)

<u>there.next</u> put Keller the disappointment that large world-movements like the Nation.League, the Marxist movement and the democracy brought have

'<u>Next</u> Keller reports the disappointment that large world-move-ments like the League of Nations, the Marxist movement and democracy brought.'

7.36. Hulle was persone om na op te sien, terwyl diegene met minder of geen grond nie, <u>dienooreenkomstig</u> minder invloed uitgeoefen het. (corpus #4, Humanities)

they be.PST persons PTCL.INF to up to see while those with less or no ground PTCL.NEG <u>accordingly</u> less influence exerted have

'They were people to look up to, while those with little or no property, exerted less influence <u>accordingly</u>.'

7.37. . . . Afrika oor drie eeue heen geboorte gegee (het) aan twee unieke etniese groepe, <u>naamlik</u> die Afrikaners en die Kleurl-inge, albei basies inheems in Suid-Afrika. (corpus #3, Informa-tional text)

Africa over three centuries over birth gave (have) to two unique ethnic groups <u>namely</u> the Afrikaners and the Coloreds both basically native in South-Africa

'. . . Africa gave birth through three centuries to two unique ethnic groups, <u>namely</u> the Afrikaners and the Coloreds, both basically native to South Africa.'

7.38. <u>Naas</u> die hoogste eis stel ons dat die onderwyser 'n persoon-
likheid moet wees wat in alles ons op skool kan vervang en
wat in alles ons kinders tot goeie voorbeeld dien. (corpus #2,
Informational text)
*<u>next</u> the highest demand put we that the teacher a personality
must be that in everything us on school can replace and that in
everything our children to good example serve*
'<u>Next</u> to the highest requirement we expect that the teacher
has to have a personality that can replace everything we are in
school and that can serve our children as a good example in
everything.'

7.39. Hulle het nie die rijkdom van uitdrukking, <u>ook</u> nie die sagtheid
van woorde, en lang nie so'n presiese spraakkunst gehad als
die Zulu's. (corpus #1, Humanities)
*they have not the riches of expression <u>also</u> not the softness of
words and long not such.a precise grammar had as the Zulus*
'They did not have such riches in expression, <u>also</u> not the soft-
ness of words, and not at all such a precise grammar as the
Zulus.'

7.40. <u>Verder</u> is u bewus van toestande in die Fort tussen die 52
aangeklaagdes. (corpus #2, Manuscript)
*<u>further</u> be.PRS you.HON aware of conditions in the Fort between
the 52 accused*
'<u>Further</u>, you are aware of the conditions in the Fort between
the 52 accused.'

7.41. Die verspreiding van die gronde in elke gebied, wat hierbo ge-
groepeer is, sal <u>vervolgens</u> bespreek word. (corpus #3, Natural
sciences)
*the spread of the grounds in every area that here.above
grouped be.AUX.PASS.PST will <u>furthermore</u> discussed become.AUX.
PASS.PRS*
'The spread of the turf in each area, grouped above, will <u>fur-
thermore</u> be discussed.'

7.42. Daar word <u>voorts</u> gekyk na die oorsprong van die swartwattel
en na gebiede waar kommersiële bosaanplanting plaasvind.
(corpus #4, Natural sciences)
*there become.AUX.PASS.PRS <u>further</u> looked at the origin of the
black.wattle and at areas where commercial forest.on.planting
place.take*

'<u>Further</u>, the origin of the black wattle will be investigated, and the areas where commercial forest plantation is taking place.'

Not all of these connectives are used as traditional conjunctions, but all function to create or maintain the coherence of the text in some way. The frequencies per 100,000 words of each are given in table 7.6.

Table 7.6.   Frequencies of additive connectives per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
boonop 'moreover'	0	4	6	7
bowendien 'moreover'	2	3	1	3
buitendien 'besides'	3	6	2	1
byvoorbeeld 'for example'	2	12	28	36
daarbenewens 'in addition'	0	2	4	4
daarby 'in addition'	23	20	11	4
daarnaas 'next'	1	1	0	<1
dienooreenkomstig 'accordingly'	0	0	1	2
naamlik 'namely'	11	13	22	17
naas 'next'	1	2	1	2
ook 'also'	483	434	445	393
verder 'further'	13	24	15	15
vervolgens 'furthermore'	<1	2	11	5
voorts 'further'	2	2	4	5

Some of the connectives have a consistently low frequency, such as *bowendien*, *daarnaas*, and *naas*. Others seem to appear (in the data at least) during the course of the century, such as *boonop*, *daarbenewens*, and *dienooreenkomstig*. Their frequencies remain low, but it does rise from being too low to occur in the data to a few or even several occurrences per 100,000 words. The item *voorts* is present in every corpus and increases a bit over the course of the century. One connective, *byvoorbeeld* (log likelihood of 104.97), increases significantly, and *daarby* declines significantly (log likelihood of 18.88). The only frequently used connective in this category is *ook*, and it shows some instability but with a general downward trajectory.

The next category is adversative connectives, those that indicate a contrast or a change in direction of the conversation. The one-word items from the sources that are present in the data are *alhoewel* 'although,' *hoewel* 'although,' *alternatiewelik* 'alternatively,' *asof* 'like,' *hetsy* 'whether,' *ofskoon* 'although,' *tensy* 'unless,' and *maar* 'but.' Once again, some of the items are also used in other functions, like *maar* 'but,' which can also be an adverb, and

these instances were removed from the data. Examples 7.43 to 7.50 illustrate each of the connectives in use:

7.43. <u>Alhoewel</u> daar 'n neerslag van ongeveer 1″ gedurende die voorafgaande 3 dae was, was die water besonder helder. (corpus #2, Natural sciences)
*although there a precipitation of about 1″ during the previous 3 days be.*PST *be.*PST *the water particularly clear*
'<u>Although</u> there was precipitation of about 1″ during the past 3 days, the water was particularly clear.'

7.44. En daar ver, noord van die skoolgebou, staan nog die rye ou kleinhuisies, <u>hoewel</u> daar nou spoele in die lobbies geïnstalleer is. (corpus #3, Fiction)
*and there far north of the school.building stand still the rows old small.houses although there now flushes in the lobbies installed be.*PST
'And far away, north of the school building, are the rows of outhouses, <u>although</u> there are flush toilets installed in the lobbies now.'

7.45. <u>Alternatiewelik</u> kan die krat van 'n spesiale selfvoerder voorsien word om ad lib-voeding moontlik te maak. (corpus #4, Informational text)
*alternatively can the crate of a special self.feeder provide become.*AUX.PASS.PRS PTCL.INF *ad lib-feeding possible to make*
'<u>Alternatively</u>, the crate can be provided with a self feeder to enable ad lib feeding.'

7.46. Dis vir altwee <u>asof</u> die swye vertroosting bring. (corpus #2, Fiction)
it.s for all.two as.if the silence comfort bring
'For both, it's <u>like</u> the silence brings comfort.'

7.47. In gesprekke, <u>hetsy</u> op kantoor of op sosiale gebied, is die deursnee Amerikaner heel gemaklik en uitgesproke oor haar geloof. (corpus #4, Religious text)
*in conversations whether on office or on social territory be.*PRS *the average American rather comfortable and outspoken about her religion*
'In conversations, <u>whether</u> in office or social situations, the average American is fairly comfortable and outspoken about her religion.'

7.48. Maar daar moet tog 'n rede wees dat die arme delfer so volhard in sij arbeid: <u>ofskoon</u> hij selde een van die skitterblinke

steentjes krij, waarnaar hij soek. (corpus #1, News report)
but there must now a reason be that the poor miner so per-
severe in his work: <u>*although*</u> *he rarely one of the glistening*
*stones.*DIM *get where.for he seek*
'But there must be a reason that the poor miner perseveres in
his work: <u>although</u> he rarely find one of the glistening stones
he's looking for.'

7.49.   Uit die aard van die saak kan vreemdelinge nie aan verkiesings
        deelneem <u>tensy</u> hulle genaturaliseer is of volle burgerskap ver-
        kry het nie. (corpus #2, Informational text)
        out the nature of the matter can strangers not in elections par-
        ticipate <u>*unless*</u> *they naturalized be.*AUX.PASS.PST *or full citizen-*
        ship gained have PTCL.NEG
        'From the nature of the matter, foreigners cannot participate in
        elections <u>unless</u> they have been naturalized or have gained full
        citizenship.'

7.50.   Ons vee het hulle afgeneem, <u>maar</u> mij oom het dit die aand
        weer teruggeneem. (corpus #1, Biographical text)
        our livestock have they taken <u>*but*</u> *my uncle have it the evening*
        again back.taken
        'They took our livestock, <u>but</u> my uncle took it back that night.'

The frequencies of the different adversative connectives per 100,000 words
are given in table 7.7.

Table 7.7.   Frequencies of adversative connectives per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
alhoewel 'although'	7	10	8	12
hoewel 'although'	18	28	33	22
alternatiewelik 'alternatively'	0	0	<1	1
asof 'like / as if'	23	28	24	42
hetsy 'whether'	2	2	2	4
ofskoon 'although'	5	3	2	<1
tensy 'unless'	4	9	3	2
maar 'but'	527	436	399	376

There is one connective that emerges in the course of the century: *alterna-*
tiewelik; in corpus #1 and corpus #2, it is too infrequent to occur in the data,
and there is only one occurrence in corpus #3 and a few in corpus #4. Another
low-frequency item is *hetsy*, which is fairly stable from corpus #1 to corpus
#3 and then rises in corpus #4. The connective *ofskoon* is not particularly

frequent but still shows a consistent decline throughout the century (log likelihood of 14.64). The two other connectives with similar meaning, *alhoewel* and *hoewel*, are somewhat more frequently used and show some variability but nothing pointing toward change. The only frequently used connective is *maar*, and it shows a significant decline in usage throughout the century (log likelihood of 69.96).

The third category of connectives is causal, to indicate causal relationships or give reasons for claims. The list of one-word causal connectives in the data is quite substantial: *aangesien* 'because,' *afgesien* 'apart,' *as* 'if,' *daarom* 'therefore,' *derhalwe* 'therefore,' *dus* 'therefore,' *gevolglik* 'consequently,' *hierdeur* 'here by,' *indien* 'if,' *mits* 'provided that,' *namate* 'as,' *weens* 'due to,' and *want* 'because.' Examples with each of the relevant connectives are given in examples 7.51 to 7.63 below:

7.51. Ek noem taalfeeste <u>aangesien</u> die plaaslike Taalfeeskomitee 'n taalfees te Glencoe sowel as te Dundee moes reël. (corpus #3, News report)
I mention language.festivals <u>because</u> the local Language.festival.committee a language.festival at Glencoe as.well as at Dundee must.PST arranged
'I mention language festivals <u>because</u> the local Language Festival Committee had to arrange a language festival at Glencoe as well as Dundee.'

7.52. <u>Afgesien</u> van bogenoemde aanpassings is die bossie ook daartoe instaat om sy lewensiklus as dit nodig is in 'n baie kort tydjie te voltooi. (corpus #2, Natural sciences)
<u>apart</u> from abovementioned adjustments be.PRS the bush.DIM also there.to capable PTCL.INF his life.cycle if it necessary be.PRS in a very short time.DIM to complete
'<u>Apart</u> from the abovementioned adjustments, the little bush is also capable of completing its life cycle in a very short time if necessary.'

7.53. . . . tien of honderd maak blijkbaar geen onderskeid nie, <u>as</u> hulle in die verre vreemde gaan verdwyn. (corpus #1, Fiction)
ten or hundred make apparently no difference PTCL.NEG <u>if</u> they in the far strangeness go disappear
'. . . ten or hundred apparently makes no difference, <u>if</u> they are going to disappear in the distance.'

7.54. <u>Daarom</u> het hulle hul plase verkoop vir spotprijse. (corpus #1, Informational text)

therefore have they their farms sold for mock.prices

'<u>Therefore</u> they sold their farms for ridiculous prices.'

7.55.   <u>Derhalwe</u> is dit gebiedend noodsaaklik vir offisiere om soveel laertrekkings as moontlik te reël waar die Boerejeug kan leer en oefen om onderstaande reëls na te kom. (corpus #2, Informational text)

therefore be.PTCL *it mandatory necessary for officers* PTCL.INF *so.many encampments as possible to arrange where the Boer. youth can learn and practice* PTCL.INF *below.standing rules* V-PTCL *to come*

'<u>Therefore</u> it is necessary for officers to arrange as many encampment events as possible where the Boer youth can learn and practice obeying the rules below.'

7.56.   . . . hij het mij vir 5 jaar laas gezien, <u>dus</u> kan hij mij nie ken nie. (corpus #1, Manuscript)

he have me for 5 years ago saw <u>*therefore*</u> *can he me not know* PTCL.NEG

'. . . he has not seen me for 5 years, <u>therefore</u> he cannot know me.'

7.57.   <u>Gevolglik</u> was matriksalgebra 'n hele aantal jare lank maar net 'n nuttelose aardigheid wat selfs in Wiskunde feitlik geen belangstelling gaande gemaak het nie. (corpus #3, Natural sciences)

<u>*consequently*</u> be.PST *matrix.algebra a whole number years long but just a useless oddity that even in Mathematics almost no interest excite made have* PTCL.NEG

'<u>Consequently</u> matrix algebra was for many years just a useless oddity that received almost no interest even in Mathematics.'

7.58.   <u>Hierdeur</u> is dit duidelik dat die tekspraktisyn tegelyk as ontvanger én sender optree en dus as 't ware twéé kommunikasiestelsels beheers. (corpus #4, Humanities)

<u>*here.through*</u> be.PRS *it clear that the text.practitioner simultaneously as receiver and sender act and so as 't true two communication.systems control*

'<u>Here by</u> it is clear that the text practitioner acts as receiver and sender simultaneously and so in effect controls two communication systems.'

7.59.   <u>Indien</u> die grond nie benat word nie, sal die plante dag en nag verwelk bly en doodgaan. (corpus #3, Natural sciences)

<u>*if*</u> *the ground not irrigated* become.AUX.PASS.PRS PTCL.NEG *will the plants day and night wither stay and die*

'<u>If</u> the ground is not irrigated, the plants will remain withered day and night and continue to die.'

7.60. By Afrikaans en Nederlands is daar byvoorbeeld 'n aantal beurse van die Nederlaandse Taalunie beskikbaar—<u>mits</u> jy gou reageer. (corpus #4, Manuscript)

at Afrikaans and Dutch be.PRS there for.example a number bursaries of the Dutch Language.union available <u>provided</u> you.SG quickly reply

'With Afrikaans and Dutch there is for instance a number of bursaries from the Dutch Language Union—<u>provided that</u> you reply quickly.'

7.61. <u>Namate</u> daardie mynhoop groter en groter geword het, het iets hier binne hom leër en leër geword . . . (corpus #2, Fiction)

as that mine.heap bigger and bigger became have have something here inside him emptier and emptier became

'<u>As</u> the mine heap grew bigger and bigger, something inside of him grew more and more empty . . .'

7.62. <u>Weens</u> gebrek aan voldoende bronne kan hierdie aanname nie gestaaf word nie. (corpus #4, Informational text)

due lack of sufficient sources can this assumption not substantiated become.AUX.PASS.PRS PTCL.NEG

'<u>Due to</u> lack of sufficient sources this assumption cannot be substantiated.'

7.63. Laat mij dit maar direk sê, dat daar in die taalwetenskap g'n absolute verskil tussen taal en dialek gemaak word nie, <u>want</u> die begrippe is relatief. (corpus #1, Humanities)

let me it but directly say that there in the language.science no absolute difference between language and dialect made become.AUX.PASS.PRS PTCL.NEG <u>because</u> the concepts be.PRS relative

'Let me be direct, in the language sciences there is no absolute difference between language and dialect, <u>because</u> the concepts are relative.'

The frequencies per 100,000 words for each of the above connectives are provided in table 7.8.

There are some infrequent connectives in this category as well—*hierdeur* and *namate* fluctuate to an extent, and *afgesien* and *derhalwe* show an increase followed by a decline. The item *mits* shows a slight increase (log likelihood of 3.91, low-level significance), and the slightly more frequent *weens* increases as well (log likelihood of 10.92). The conditional *indien* increases

Table 7.8. Frequencies of causal connectives per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
as 'if'	287	206	161	193
aangesien 'because'	7	20	28	21
afgesien 'apart'	2	11	6	5
daarom 'therefore'	51	44	33	38
derhalwe 'therefore'	5	9	7	0
dus 'therefore'	107	89	87	62
gevolglik 'consequently'	4	11	19	7
hierdeur 'here by'	2	5	1	3
indien 'if'	8	23	39	38
mits 'provided that'	1	2	3	3
namate 'as'	0	6	2	4
want 'because'	180	102	77	81
weens 'due to'	6	12	11	16

quite substantially from corpus #1 to corpus #3 (log likelihood of 57.02) and then remains stable. *Daarom* is more frequent but also shows fluctuation. One connective decreases consistently: *dus* (log likelihood of 30.23); two others also decline from corpus #1 to corpus #3: *as* and *want*.

The final category of connectives are temporals that serve the double function of linking clauses or other items and expressing a temporal relationship simultaneously. One-word connectives in this category are *alvorens* 'before,' *dan* 'then,' *intussen* 'in the meantime,' *sedert* 'since,' *sedertdien* 'since then,' *sodra* 'as soon as,' *tans* 'currently,' *terselfdertyd* 'at the same time,' *terwyl* 'while,' *toe* 'then/when,' *vantevore* 'before,' *volgende* 'next,' *voorheen* 'before,' and *wanneer* 'when.' These connectives are exemplified in examples 7.64 to 7.77:

7.64. Nog drie sake het Leyds se kommentaar uitgelok <u>alvorens</u> hy
 na Suid-Afrika teruggekeer het. (corpus #3, Humanities)
 another three matters have Leyds PTCL.GEN comments provoked
 <u>before</u> he to South-Africa returned have
 'Another three matters provoked Leyds's comments <u>before</u> he
 returned to South Africa.'

7.65. As jij wil saam eet, <u>dan</u> moet jij saam werk. (corpus #1, Bio-
 graphical text)
 if you.SG want together eat <u>then</u> must you.SG together work
 'If you want to eat with us, <u>then</u> you have to work with us.'

7.66. <u>Intussen</u> is my raad: Bly maar liewer binne die raamwerk wat
 die wet vir jou skep. (corpus #4, News report)
 <u>in.between</u> be.PRS my advice stay but rather inside the frame-

       work that the law for you.SG create
       '<u>In the meantime</u> my advice is: Rather stay within the frame-
       work that the law provides.'

7.67. <u>Sedert</u> 2004 help hy 'n swart bemagtigingskatoenprojek
       (Makhathini Cotton Farming) in die noorde van KwaZulu-Na-
       tal as konsultant en deel-bestuurder. (corpus #4, News report)
       <u>since</u> 2004 help he a black empowerment.cotton.project
       (Makhathini Cotton Farming) in the north of KwaZulu-Natal
       as consultant and part-manager.
       '<u>Since</u> 2004 he's been helping a black empowerment cotton
       project (Makhathini Cotton Farming) in the north of Kwa-Zulu
       Natal as consultant and section manager.'

7.68. <u>Sedertdien</u> het hulle die wêreld platgereis. (corpus #3, News
       report)
       <u>since.then</u> have they the world flat.traveled
       '<u>Since then</u> they have traveled all over the world.'

7.69. <u>Sodra</u> die predikant sy preekstoel verlaat, praat hy Afrikaans.
       (corpus #2, Informational text)
       <u>as.soon.as</u> the minister his pulpit leave speak he Afrikaans
       '<u>As soon as</u> the minister leaves his pulpit, he speaks Afrikaans.

7.70. <u>Tans</u> is sy gemiddelde fooi R117. (corpus #4, Manuscript)
       <u>currently</u> be.PRS his average fee R117
       '<u>Currently</u> his average fee is R117.'

7.71. <u>Terselfdertyd</u> word klein bewegings in die wand van die maag
       aan die gang gesit. (corpus #2, Humanities)
       <u>at.same.time</u> become.AUX.PASS.PRS small movements in the wall
       of the stomach on the go put
       '<u>At the same time</u> small movements in the stomach wall are
       launched.'

7.72. Sij sit hulle kos stil-stil op tafel, <u>terwijl</u> hij deurpraat om die
       twis uit te lok. (corpus #1, Biographical text)
       she put their food silent-silent on table <u>while</u> he through.speak
       PTCL.INF the fight out to attract
       'She puts their food on the table quietly, <u>while</u> he's talking to
       solicit controversy.'

7.73. Die nadraai het egter ewe skielik ingetree <u>toe</u> die eerste klom-
       pie kinders na 'n paar maande die skool as volleerd verlaat het.
       (corpus #2, Humanities)
       the after.turn have however equally suddenly came <u>when</u> the
       first group.DIM children after a couple months the school as
       full.learned left have

'The aftermath came equally suddenly <u>when</u> the first group of children left the school a few months later as fully learned.'

7.74. Die resonansies in die energiegebied Ep = 1,2 tot 2 MeV wat bestudeer is, is <u>vantevore</u> deur Wolff e.a. (Wo 68) ondersoek. (corpus #3, Natural sciences)
the resonances in the energy.range Ep = 1,2 to 2 MeV that studied be.AUX.PASS.PST be.AUX.PASS.PST <u>before</u> by Wolff et al (Wo 68) studied
'The resonances in the energy range Ep = 1,2 to 2 MeV that were studied, have been investigated by Wolff et al (Wo 68) <u>before</u>.'

7.75. <u>Volgende</u> ding, iemand het dit goedgedink om daai Italianer se plek op te blaas. (corpus #4, Fiction)
<u>next</u> thing somebody have it good.thought PTCL.INF that Italian PTCL.GEN place up to blow
'<u>Next</u> thing, somebody saw fit to blow up that Italian's place.'

7.76. Dit strook dan ook met die beginsels van verantwoorde-likheidskoste, wat <u>voorheen</u> gedefinieer is. (corpus #3, Humanities)
it consistent then also with the principles of responsibility.costs that <u>before</u> defined be.AUX.PASS.PST
'It is consistent with the principles of responsibility costs, defined <u>before</u>.'

7.77. <u>Wanneer</u> dit afgekoel is, gooi omtrent 4 eetlepels terpentijn daarin en 2 lepels ammonia. (corpus #1, Informational text)
<u>when</u> it down.cooled be.PRS add about 4 eat.spoons turpentine there.in and 2 spoons ammonia
'<u>When</u> it has cooled down, add about 4 table spoons of turpentine and 2 spoons of ammonia.'

The frequencies per 100,000 words of the temporal connectives above are provided in table 7.9.

There are a few low-frequency connectives that show fluctuation without change: *alvorens*, *terselfdertyd*, and *vantevore*. Two low-frequency items show increase (in as far as such low frequencies can be taken at face value): *sedertdien* (log likelihood of 8.65) and *volgende* (log likelihood of 8.22). Other connectives showing signs of increase include *intussen* (log likelihood of 5.33, low-level significance), *sedert* (log likelihood of 6.20, low-level significance), *tans* (log likelihood of 34.74), *terwyl* (log likelihood of 13.09), *voorheen* (log likelihood of 9.86), and *wanneer* (log likelihood of 17.19). The high-frequency *dan* shows a stark decline (log likelihood of 238.10),

Table 7.9.   **Frequencies of temporal connectives per 100,000 words**

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
alvorens 'before'	1	4	5	2
dan 'then'	220	111	82	79
intussen 'in the meantime'	5	8	6	11
sedert 'since'	10	13	12	18
sedertdien 'since then'	0	1	2	2
sodra 'as soon as'	8	15	6	9
tans 'currently'	5	9	10	22
terselfdertyd 'at the same time'	1	7	9	6
terwyl 'while'	61	63	65	84
toe 'then/when'	122	117	118	140
vantevore 'before'	<1	1	2	1
volgende 'next'	0	0	0	1
voorheen 'before'	5	5	7	11
wanneer 'when'	35	64	67	83

especially from corpus #1 to corpus #3, and then another small decrease between corpus #3 and corpus #4. Importantly, the fairly informal *dan* and *toe* are either declining or do not show a clear increase, while the temporal connectives that show a clear rise in frequency are all more formal and low frequency, such as *wanneer* and *tans*. I will return to this topic in the next section on colloquialization and formalization.

One general trend emerges from the data in this section: the high-frequency items are typically on the decline. These items are fairly simple and informal, with general meanings within their categories: *ook* 'also,' *maar* 'but,' *dus* 'therefore, and *dan* 'then' continue to decrease, and *as* 'if' and *want* 'because' decline for most of the century with small increases between corpus #3 and corpus #4. It would seem, then, that the more general, informal connectives are making way for more specific, formal connectives, at least in written Afrikaans. It links directly to trends I explore in the next section on colloquialization and formalization.

COLLOQUIALIZATION AND FORMALIZATION

Background

The process of *colloquialization* is sometimes defined as written language becoming more like spoken language (Collins 2013, 290). Collins (2013, 290) points out certain problems with this description, however, as it leads to a circular definition and fails to account for features shifting across a range of

different registers. Instead, we can define the process in terms of the decline in forms that signify distance and formality, toward more involved and informal options (Baker 2009, 324), or markers of power and asymmetry making way for indications of equality and familiarity (Leech et al. 2009, 259). The process typically includes the spread of a variant from informal, colloquial speech to more formal (spoken and written) contexts (Collins and Yao 2013, 480).

In this section, I use the linguistic and prescriptive literature to determine a number of phenomena that are regarded as particularly formal or informal, as indications of distance and power or of familiarity and equality. For example, the second-person pronoun *u* is the honorific (singular or plural), while *jy/jou* (singular) is regarded as quite familiar, and *julle* (plural) is unmarked for either formality or informality, distance or familiarity.

The sociocultural context of written Afrikaans suggests that extralinguistic factors like colloquialization (and its opposite, formalization) might be particularly relevant for twentieth- and early twenty-first-century language use. I attend to the history of Afrikaans standardization at length in chapter 3, but I briefly highlight the most relevant points in the history again for contextualization.

The first formal prescriptive sources on written, Standard Afrikaans appeared in 1917—the first *Afrikaanse Woordelys en Spelreëls* (Afrikaans wordlist and spelling rules) (Le Roux, Malherbe, and Smith 1917) as well as Malherbe's *Afrikaanse Taalboek* (Afrikaans language book) (Malherbe 1917). Some similar efforts were made in the late nineteenth century, particularly by the *Genootskap van Regte Afrikaanders* (Society of true Afrikaners), such as their *Eerste beginsels van die Afrikaanse taal* (First principles of the Afrikaans language) (1876) and *Fergelijkende taalkunde fan Afrikaans en Engels* (Comparative grammar of English and Cape Dutch) (Du Toit 1897). However, these books were not widely accepted, and the efforts to formally standardize written Afrikaans were arrested by the Second War of Freedom (also called the Anglo-Boer War) from 1899 to 1902, during which there was little time and resources to attend to cultural matters, such as language norms.

In the second decade of the twentieth century, standardizing efforts were revived, and by 1925, Afrikaans was formally recognized as an official language of what was then the Union of South Africa. In the following decades, a plethora of grammars, dictionaries, and other prescriptive guidelines to the "proper" use of Standard Afrikaans were published and republished. It would not be unexpected, then, to see signs of increasing formality, distance, and sophistication in written Afrikaans during the first several decades of the twentieth century.

At the turn of the century, the sociocultural context in South Africa had changed quite dramatically. In 1994, the first democratic general election was held, and Apartheid was formally over. Since then, the political democracy gave further incentive for language to become equally democratized, as Leech et al. (2009, 259) refer to the reduction of markers of power and asymmetry. Some effort has been made to make written Afrikaans more accessible and more representative of broad language use (including other varieties, not only the Standard of the Apartheid era). Other forces, such as mass media and social media, are also able to contribute to colloquialization, as is the case with English. In contrast to the early part of the twentieth century, the last part of the twentieth and the first decade of the twenty-first (represented by corpus #4) could very well have indications of increasing informality and familiarity, where markers of formality and distance are avoided even in more formal contexts and text types.

In the following subsections, I explore trends of formalization and colloquialization in the data according to three broad categories: pronouns, passives, and adverbs, adverbials, and conjunctions.

Pronouns

Second-Person and Generic Pronouns

In the introduction to the section, I already mentioned the honorific second-person pronoun *u*, used as both singular (7.78) and plural (7.79); the plural second-person *julle* (7.80) and its shortened form *jul* (7.81); and the familiar singular second-person *jy* (subject form, 7.82) and *jou* (object form, 7.83):

7.78. Wanneer wil u̲ met die werk begin, meneer Du Preez? (corpus #3, Fiction)
when want you̲.HON with the work begin mister Du Preez
'When do you̲ want to start with the work, mister Du Preez?'

7.79. "Die uurwerk wat u̲ hier sien, dames en here," vervolg hij op salwende toon . . . (corpus #1, Fiction)
the hour.work that you̲.HON here see ladies and gentlemen continue he in soothing tone
'"The clock that you̲ see here, ladies and gentlemen," he continues in a soothing tone . . .'

7.80. Ek vertrou dat bogemelde inligting vir ju̲lle van waarde sal wees. (corpus #4, Manuscript)
I trust that abovementioned information for you̲.PL of value will be
'I trust that the above information will be of value to you̲.'

7.81.   Jul hou maar vir jul wild, dit kamtig beter lyk. (corpus #1, Bio-
        graphical text)
        you.PL keep but for you.PL wild it supposedly better look
        'You act so wild, because it supposedly looks better.'

7.82.   Jy kan jou swemklere saambring, as jy tussen die kurpers en
        die paddaslyk wil gaan afkoel! (corpus #4, Manuscript)
        *you.SG can your.SG swim.clothes along.bring if you.SG between
        the breams and the frog.slime want go down.cool*
        'You can bring along your swimsuit, if you want to cool down
        with the breams and the frog slime!'

7.83.   Dit laat ek aan jou oor. (corpus #2, Fiction)
        it leave I to you.SG over
        'I leave it to you.'

The honorific *u*, plural *julle*, and singular object form *jou* are also used as
attributive possessive pronouns, as illustrated in examples 7.84 to 7.86:

7.84.   Ons vertrou dat u die aangehegte inligting waardevol sal vind
        met u finansiële beplanning. (corpus #4, Manuscript)
        *we trust that you.HON the attached information useful will find
        with your.HON financial planning*
        'We trust that you will find the attached information useful for
        your financial planning.'

7.85.   Met plesier stem ek in om mij naam op julle lijs te laat plaas.
        (corpus #1, News report)
        *with pleasure agree I in PTCL.INF my name on your.PL list to let
        place*
        'With pleasure I agree to have my name added to your list.'

7.86.   Het jy jou eetgoed gekry? (corpus #3, Manuscript)
        have you.SG your.SG eat.things get
        'Did you get your snacks?'

Finally, the honorific *u* and the singular forms *jy* and *jou* are also used as
a generic (human indefinite) pronoun, as in examples 7.87 to 7.89, including
in the possessive uses. In example 7.88, the common use in combination with
('n) mens as generic pronoun is illustrated:

7.87.   As u in die koerante lees hoedat die wetteloosheid alreeds sy
        opmars in ons vaderland loods . . . (corpus #3, Religious text)

if *you.HON* in the newspapers read how.that the lawlessness al-
ready his up.march in our fatherland launch
'If <u>you</u> read in the papers how the lawlessness already marches
in our fatherland . . .'

7.88. 'n Mens voel jou die grootste dwaas als <u>jij</u> uit die Kollege
kom. (corpus #1, Informational text)
a human feel you.SG.REFL. the greatest fool when <u>you.SG</u> out the
College come
'One feels like such a fool when <u>you</u> finish at the College.'

7.89. Wees daarom dankbaar vir hekke. God gebruik dit om <u>jou</u>
terug te wen. (corpus #4, Religious text)
be therefore thankful for hurdles God use it PTCL.INF <u>you.SG</u>
back to win
'Be thankful for the hurdles. God uses it to win <u>you</u> back.'

According to Ponelis (1979, 67), the choice between *jy/jou*, *u*, and *julle* is
determined by the formality of the situation. Donaldson (1986, 4–5) men-
tions that *u* is used much less frequently in Afrikaans than in Dutch; Scholtz
(1980, 67) identifies something un-Afrikaans in *u*, describing it as a cultural
loan from Dutch; and Ponelis (1993, 207) claims that the use of *u* is very lim-
ited in contemporary Afrikaans, almost exclusive to formal written and very
formal spoken language. A measure of social deixis was transferred from
Dutch, then, but adjusted in Afrikaans for much higher levels of formality
and distance associated with *u* even early in the twentieth century (Booysen
and Jansen 1922, 84; Botha and Burger 1926, 67; 1927, 64; Botha and Van
Aardt 1987, 177; Bouman and Pienaar 1924, 73; Prinsloo and Odendal 1995,
90). To show respect in Afrikaans, without the stifling formality of *u*, speak-
ers use a number of forms of address, especially *oom* 'uncle' for older men
and *tannie* 'aunt(y)' for older women, regardless of family relation. When *u* is
regarded as too formal but *jy/jou* as too familiar, speakers also use antecedent
repetition (Botha and Burger 1926, 67; 1927, 64; Bouman and Pienaar 1924,
73; Du Toit 1902, 18; Van der Merwe 1951, 46), such as "*Ma, hier is Ma se*
hoed" 'Mom, here is Mom's hat' or "*Meneer, ek het meneer se hoed gesien*"
'Mister, I saw mister's hat' (Booysen and Jansen 1922, 84). Even though the
sources I cite here are quite old, a recent article indicates that both the prac-
tices of using *oom* and *tannie* as well as antecedent repetition are still used by
speakers today (Bosman and Otto 2015).

Before I attempt to determine possible effects of colloquialization and
formalization, I will provide descriptive statistics briefly for each of the three
variants (singular, plural, and honorific).

Table 7.10.   Frequencies of *jy/jou* per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
second-person singular	131	231	259	266
generic pronoun	112	56	61	288
possessive second-person singular	34	51	53	58
possessive generic	25	14	25	147
reflexive	7	6	3	5

The frequencies of *jy* and *jou* are reported together in table 7.10, with their different contexts of use.

It is clear from the table that the use of *jy/jou* rises subtly from corpus #1 to corpus #3 but then increases quite dramatically to corpus #4 (log likelihood of 671.56). The main contribution to corpus #4 is from generic uses (personal and possessive pronouns), even though the second-person uses continue to increase. Table 7.11 shows the frequencies of the plural form *julle* (including the shortened *jul*).

Table 7.11.   Frequencies of *julle/jul* per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
second-person plural	35	13	35	59
possessive second-person plural	7	4	5	16
reflexive	1	1	<1	3

The second-person plural shows a different pattern from the singular—first, it is not used as a generic pronoun, and, second, it does not consistently increase throughout the century but first declines and then increases from corpus #2 to corpus #4. The usage of *u* is indicated in table 7.12.

Table 7.12.   Frequencies of *u* per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
second-person singular	75	68	45	43
second-person plural	32	5	3	4
generic pronoun	6	40	28	<1
possessive second person	47	29	22	34
possessive generic	5	14	8	0
reflexive	1	1	<1	<1

The second-person uses show clear signs of decline, especially when taken together (log likelihood of 81.84). The other uses are a little more erratic, and reflexive uses remain consistently low. The overall frequency does show a consistent decline (log likelihood of 105.21). However, there is a complicating factor in the frequency data of *u*. In each of the corpora, one or a few texts are included in the corpus with an unusually high frequency of *u* compared to the average. In corpus #1, there is one newspaper report on a student league meeting. In this text, *u* is used while reporting on discussions during the meeting (where attendees addressed one another as *u*), and correspondents refer to the league as *uw Bond* 'your League.' In corpus #2, there are two texts with an unusually high frequency of *u*: a university textbook on botany (in the category of natural sciences) that addresses the reader as *u* while explaining certain procedures and an informational text on procedures in the *Ossewabrandwag* 'Ox wagon sentries' that contains a large number of imperatives where the reader is addressed as *u* as well. In corpus #3, two texts have particularly high frequencies of *u*: one is an informational text on making mosaics, where the reader is regularly addressed as *u* while giving direct instructions, and the other text is about fish in the kitchen, also with a high number of uses of generic *u* to address the reader while giving instructions. Finally, in corpus #4, there is one extract from a historical novel in the fiction category, where *u* is used to create a historical atmosphere, along with some archaic word choices, such as *elkander*, which is an older form of *mekaar*, the reciprocal pronoun in Afrikaans.

When I exclude the texts with a disproportionally high frequency of *u*, the decline in its usage becomes even clearer:

- Corpus #1: 177 per 100,000 words
- Corpus #2: 121 per 100,000 words
- Corpus #3: 96 per 100,000 words
- Corpus #4: 76 per 100,000 words

The frequencies of the honorific second person *u*, the familiar second-person singular *jy/jou*, and the neutral second-person plural *julle* are compared in the chart in figure 7.1.

The trends show that *jy/jou* and *u* take the opposite direction of each other, the former increasing consistently and the latter declining consistently. The plural *julle* also increases from corpus #2 on.

The fact that the honorific, signaling formality and distance declines while the familiar and neutral variants increase points toward colloquialization.
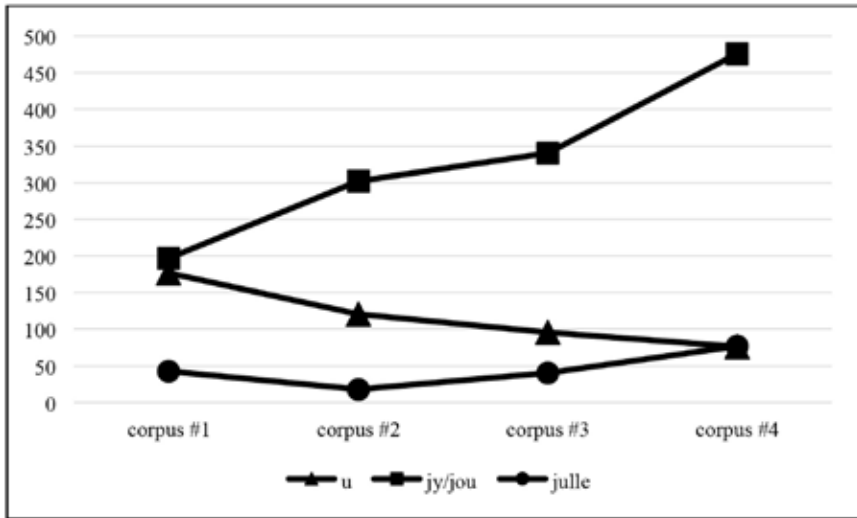
Figure 7.1.   Use of *u, jy/jou*, and *julle*

Demonstratives

The next category of pronouns under investigation is demonstrative pronouns. The demonstrative pronouns in Afrikaans are *hierdie* 'this/these,' *daardie* 'that/those,' and *dié* 'this/these/that/those' (Den Besten 2012b, 61–62; De Stadler 1989, 138; Kirsner 2014, 44; Ponelis 1979, 89; 1993, 168; Roberge 2001, 128). Traditionally, there is a distinction between, on the one hand, *hierdie* and *dié* as signaling that which is relatively close to the speaker and, on the other, *daardie* as signaling that which is far from or not close to the speaker (De Stadler 1989, 138). However, Kirsner (2014, 45–46) argues for a different model—according to him, each of these demonstratives indicates different types of deixis: *hierdie* signals central (or proximate) deixis, *daardie* signals peripheral (or distal) deixis, and *dié* signals deixis that is unspecified for distance or proximity (Kirsner 2014, 48). The reason that *hierdie* and *dié* are often used as synonyms, according to Kirsner (2014, 52–55), is because of people's egocentricity. People tend to refer to objects and matters that are close by, so *dié* as unspecified for distance is used more often than not to refer to proximate objects and matters. This still does not mean that *hierdie* and *dié* are identical.

In Afrikaans, the Dutch demonstratives *deze*, *dit*, *die*, and *dat* have been replaced with the neologisms *hierdie* and *daardie* and accentuated *dié* (Den Besten 2012b, 61–62; Roberge 2001, 124). To trace the origins of the neologisms is quite complex—after simplification typical during creole develop-

ment, the semantic relationship between locatives and demonstratives, as well as Khoikhoi and Malay influence, led to the expansion of the system to include *hierdie* and *daardie*. These forms emerge only later in the varieties of the former slaves using Arabic-Afrikaans than those of the colonists (Den Besten 2012b, 70; Roberge 2001, 127–28). Roberge (2001, 132) summarizes the development as follows:

> We can be reasonably certain that *die* was the sole determiner in the Cape Dutch Pidgin, as Den Besten (1988) argues, serving as both definite article and demonstrative pronoun, with proximal and distal interpretations governed by context. Locally born language learners drew on the resources of a fully developed superstrate language (Cape Dutch) alongside a co-territorial Cape Dutch Pidgin, during which time distinctions of proximity came to be systematically encoded by means of locative adverbs. This syntactic pattern could well represent the extension and re-analysis of the metropolitan *hier/daar* + DP structure, abetted perhaps by universal semantic relations underlying the location and identification of referents in a spatio-temporal context, and by substrate influences, as Den Besten suggests.

Seen in the context of emergent grammar, the items *hierdie* and *daardie* are consequences of grammaticalization of the general demonstrative *dié* in combination with general locatives that signal either proximity (*hier* 'here') or distance (*daar* 'there'). The originally separate locatives and demonstrative were used adjacently so frequently that it eventually merged into a new construction, a new exemplar analyzed and recalled as a single unit. The late emergence (Raidt 1991, 209; Scholtz 1980, 74) and slow spread of *hierdie* and *daardie* are attributed to the semantically similar Dutch forms *deze* and *die* that were still used occasionally in Cape Dutch (Roberge 2001, 132).

The short and deictically neutral *dié* can be used both attributively (7.90) and independently (7.91), with a broad semantic base similar to English *this*, *these* and *that*, *those* (Den Besten 2012b, 75). The independent use of *hierdie* (7.92) and *daardie* (7.93) is, however, a little controversial—while De Stadler (1989, 418) only mentions it, Carstens (2018, 78) disapproves of it, and both Carstens (2018, 78) and Ponelis (1979, 90) attribute this usage to English influence. Kirsner (2014, 70) does mention, however, that adjectival demonstratives tend to develop into independent pronouns—this means that the independent use of *hierdie* and *daardie* could just as well have happened without English influence. More recently, however, Van Rooy (in press) makes a convincing argument, supported with extensive analyses, that this development can indeed be attributed to English influence. The focus of this section is not on the English influence but, rather, on the distinction in formality between *hierdie* and *daardie* on the one hand and *dié* on the other. An

example of the attributive use of *hierdie* and of *daardie* is given in examples 7.94 and 7.95:

7.90.   D<u>ié</u> aanblik van 'n ritmies harmonies bewegende liggaam, merk hy verder op, werk "erotiserend" (verliewend). (corpus #1, Religious text)
<u>*this*</u> *on.look of a rhythmic harmonic moving body remark he further v-ptcl work erotically in.loving*
'<u>This</u> viewing of a rhythmically, harmoniously moving body, he further remarks, works "erotically."'

7.91.   Hy het die kudde by sy pa oorgeneem toe <u>dié</u> afgetree het. (corpus #4, News report)
he have the herd at his father over.took when <u>this</u> retired have
'He took over the herd from his father when <u>he</u> retired.'

7.92.   Om só te verloor is nooit lekker nie. Ek glo egter <u>hierdie</u> was 'n uitsondering en dat ons in Kaapstad sal terugveg. (corpus #4, News report)
ptcl.inf such to lose be.prs never nice ptcl.neg I believe however <u>this</u> be.pst an exception and that we in Cape.Town will back.fight
'To lose like this is never nice. However, I believe <u>this</u> was an exception and that we will fight back in Cape Town.'

7.93.   <u>Daardie</u>, waarin Ds. Du Toit sijn eerste opleiding ontvang het, was 'n sogenoemde 3de klas goewernementskool aan die Dal . . . (corpus #1, Biographical text)
<u>that</u> where.in Ds. Du Toit his first training received have be.pst a so.called 3rd class government.school at the Valley
'<u>That</u>, where Ds. Du Toit received his first training, was a so-called 3rd rate government school in the Valley . . .'

7.94.   Die meeste van <u>hierdie</u> voedings het 'n intensiteit van kleiner as 2% gehad. (corpus #3, Natural sciences)
the most of <u>these</u> feedings have an intensity of smaller than 2% had
'Most of <u>these</u> feedings had an intensity of smaller than 2%.'

7.95.   En <u>daardie</u> deurleefde siening wil die kunstenaar aan ander oordra. (corpus #3, Religious text)
and <u>that</u> through.lived view want the artist to other over.carry
'And <u>that</u> lived through viewpoint, the artist wants to convey to others.'

Regarding the relative formality and informality of the different demonstratives, Kirsner (2014, 56) indicates that *hierdie* is more objective and less involved than *dié*. It is also more specific, which is another indication of formality rather than familiarity. The same distinction can be made between *daardie* and *dié*, which means that an increase in the use of *dié* and a concurrent decrease in the use of *hierdie* and *daardie* could point toward colloquialization.

Before I continue to give descriptive statistics of the demonstratives, I have to attend to a methodological difficulty in extracting all the instances of the deictically neutral *dié* from the corpus data. While in the majority of cases *dié* is indeed spelled with the accented é, this is not always the case. Without the accent, it is spelled exactly the same as the definite article *die*, which is the single most common word in each of the corpora. In order to find an average of occurrences of demonstrative *dié* written as *die*, I extracted six random samples of 500 occurrences of *die* from each corpus and determined the amount of demonstratives in each. Using these numbers, I extrapolated the total number of demonstratives from the total occurrences of *die* in each corpus.

One criterion for distinguishing the articles from the demonstratives is when *die* is used independently, in which case it cannot be an article (e.g., 7.96). Most of the instances of *die* that were classified as demonstratives are of this kind. However, there are some attributive uses that can be classified as demonstratives as well, specifically when it is clear from the context that the *die* is separating one or a collection of something from a larger group, such as example 7.97, which sets two of the festival days apart from the rest, or example 7.98, where the then South African Republic was set apart as the last part of the country that surrendered its independence to the British:

> 7.96. Die besoedeling wat veroorsaak word kan ook vergelyk word met <u>die</u> van steenkoolkragstasies waarvan daar 21 in Suid-Afrika in bedryf is. (corpus #4, Natural sciences)
> *the pollution that caused become.*AUX.PASS.PRS *can also compared become.*AUX.PASS.PRS *with <u>that</u> of coal.power.plants where.of there 21 in South-Africa in operation be.*PRS
> 'The pollution that is caused can be compared to <u>that</u> of the coal power plants of which there are 21 in operation in South Africa.'

> 7.97. <u>Die</u> twee, onafskeidbaar saam verbind, het vir ons Dingaansdag gegee. (corpus #2, Informational text)
> <u>*these*</u> *two inseparably together bound have for us Dingaans. day gave*
> '<u>These</u> two, inseparably linked, gave us Dingaan's day.'

7.98.   Die verlies van die onafhanklikheid van <u>die</u> deel was laaste
bittere voorwaarde vir die bewuswording van die volkseenheid
van Kaap tot Limpopo daarna. (corpus #2, Humanities)
*the loss of the independence of <u>this</u> part be.PST last bitter con-
dition for the aware.becoming of the nation.unity of Cape to
Limpopo thereafter*
'The loss of the independence of <u>this</u> part was the last bitter
condition for the awareness of the national unity from Cape to
Limpopo afterwards.'

In table 7.13, I give the approximated frequencies of the different uses of *dié* and demonstrative *die*.

Table 7.13.   Approximated frequencies of the different uses of *dié* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
demonstrative determiner	158	73	119	156
independent demonstrative	28	49	39	44
TOTAL	186	122	158	200

A pattern emerges from the table—*dié* declines quite substantially from corpus #1 to corpus #2 (log likelihood of 32.69), but then it rather consistently increases until corpus #4 (log likelihood of 51.59). In table 7.14, the frequencies of *hierdie* are given.

Table 7.14.   Frequencies of the different uses of *hierdie* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
demonstrative determiner	117	406	362	284
independent demonstrative	<1	2	3	5
TOTAL	117	408	365	289

The pattern in table 7.14 is to some extent a mirror image of that in table 7.13—the use of *hierdie* increases drastically from corpus #1 to corpus #2 (log likelihood of 428.21), and then it declines from corpus #2 to corpus #4 (log likelihood of 53.59). It is noteworthy that the independent usage remains low frequency but does seem to increase throughout the century. It would not be unexpected to see a pattern similar to *hierdie* for *daardie*. In addition to the full form, *daardie* has a very informal variant *daai*, illustrated in examples 7.99 and 7.100:

7.99. Ek wil maar net noem, <u>daai</u> klein Sizwe gee vir ons probleme, en dit gaan erger word. (corpus #4, Fiction)
I want but just say <u>that.INFORM</u> small Sizwe give for us problems and it go worse become
'I just want to say, <u>that</u> Sizwe is giving us problems, and it's going to become worse.'

7.100. Ek het dringend so 'n handboek nodig, van <u>daai's</u> wat sy self gebruik hoe bekom ek dit? (corpus #4, Manuscript)
I have urgent such a textbook need of <u>those.INFORM</u> that she self use how acquire I it
'I have an urgent need for such a textbook, of <u>those</u> that she uses herself, how do I acquire one?'

The frequencies of *daardie* and *daai* are given in table 7.15.

As expected, *daardie* shows a similar frequency pattern to *hierdie*: a significant increase from corpus #1 to corpus #2 (log likelihood of 67.23) and a decrease from corpus #2 to corpus #4 (log likelihood of 18.04). The use of *daai*, whether attributive or independent, is absent in corpus #1 and corpus #2, with only a single instance in corpus #3 and a more visible frequency in corpus #4. It is not used frequently enough to really influence the bigger picture yet.

From the data reported in tables 7.13 to 7.15, it becomes clear that both *hierdie* and *daardie* are not really in competition with *dié* regarding the independent uses, at least not in written Afrikaans. The category of use where change presents itself is that of demonstrative determiner, where the demonstratives are used attributively. A comparison of the frequencies are given in figure 7.2.

In the graph in figure 7.2, the frequencies of *dié* and *daardie* are mirror images of one another. Regarding *hierdie*, in corpus #1, it is still less frequent than the older *dié*, but its frequency rises drastically in corpus #2. Like *daardie*, it then declines for the rest of the period. It would seem, then, that there

Table 7.15. **Frequencies of the different uses of *daardie* and *daai* per 100,000 words**

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
daardie demonstrative determiner	42	103	74	60
daardie independent demonstrative	<1	0	0	1
daai demonstrative determiner	0	0	<1	9
daai independent demonstrative	0	0	0	1
TOTAL	42	103	74	71
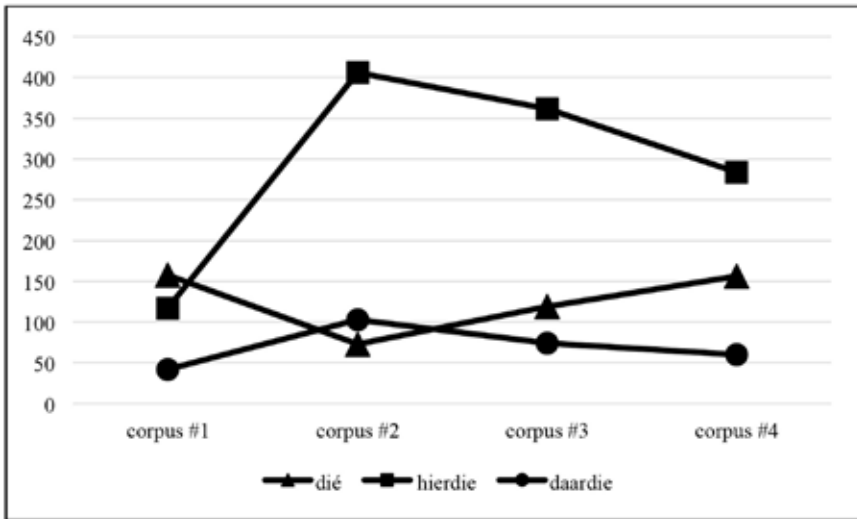
Figure 7.2.   Frequencies of demonstrative determiners

are signs of colloquialization from corpus #2 on, where the two formal vari-
ants decline, while the informal variant *dié*, unmarked for distance, increases.

Passive Constructions

Descriptions of, comments on, and recommendations regarding Afrikaans
passive constructions have been present in prescriptive and descriptive litera-
ture as well as the Afrikaans media for almost a century.

Regarding the passive in general, there are subtle differences in opinion.
On the one hand, in several different editions of *Norme vir Afrikaans* (Norms
for Afrikaans), Carstens admits that passives could be useful but still recom-
mends that it rather be avoided in favor of the active voice, unless there are
specific reasons to use the passive (Carstens 1991, 325; 1994, 325; 2003,
371; 2011, 366; 2018, 400). On the other hand, Du Toit and Smith-Müller
(2003, 85–86) are aware of the often repeated recommendation to avoid the
passive and give a more nuanced description and recommendation in this
regard. They admit that the passive voice could add or increase formality and
distance but also indicate its pragmatic value and refrain from an outright
negative evaluation. Whatever the recommendations may entail, the passive
is prominent enough in the literature to justify a small discourse on the mat-
ter. Authors generally agree that the auxiliary *word* 'become' is used in the
present-tense passives (7.101) and that the auxiliary *is* 'is/are' is used in past-

tense passives (7.102) (Carstens 2018, 401; Du Toit 1986, 137; Ehlers 1990, 17). An issue that draws almost as many comments as passive use in general is the use of the auxiliary *was* 'was/were' in past-tense passives (7.103):

> 7.101. Misskien dat daardeur die hand van die Departement geforseer <u>word</u> om alle kwekelinge te dwing om hulle hele kursus in Opleiding-skole deur te maak. (corpus #1, Informational text)
> *maybe that there.through the hand of the Department forced become.AUX.PASS.PRS PTCL.INF all pupil.teachers to coerce PTCL.INF their whole course in Training-schools through to make*
> 'Maybe through this the hand of the Department <u>is</u> forced to coerce all the pupil teachers to complete the whole course in Training schools.'
> 7.102. In 1895 <u>is</u> daar ook voorsiening gemaak vir 'n universiteit-seksamen vir die "First Class Teachers' Certificate." (corpus #2, Humanities)
> *in 1895 be.AUX.PASS.PST there also provision made for a university.exam for the "First Class Teachers' Certificate"*
> 'In 1895 there <u>was</u> made provision for a university exam for the "First Class Teachers' Certificate."'
> 7.103. Mohammed se leringe <u>was</u> gebaseer op bekende geloofsbeginsels. (corpus #4, Informational text)
> *Mohammed PTCL.GEN teachings be.AUX.PASS.PST based on familiar faith.principles*
> 'Mohammed's teachings <u>were</u> based on familiar religious principles.'

The use of *was* in Afrikaans passives is, however, not the focus of this section. I explore this trend in more detail in Kirsten (2015). The focus here is the use of the passive voice in general in Afrikaans as a sign of formality and distance.

In table 7.16, I give the frequencies per 100,000 words of Afrikaans passive constructions in the present tense and in the past tense as it is used in my corpora.

Table 7.16. Frequencies of the passive voice per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
present-tense passives	464	714	770	722
past-tense passives	291	484	516	450
TOTAL	755	1,198	1,286	1,172

The jump in frequency from corpus #1 to corpus #2 is quite drastic (log likelihood of 258.22), with a further small increase to corpus #3. The use of passive constructions decreases, then, from corpus #3 to corpus #4 (log likelihood of 13.85) to just below the frequency of corpus #2. To get a clearer picture of the spread of passive constructions, table 7.17 gives the frequencies per 100,000 of passives in each of the text types in the corpora.

Table 7.17.   Frequencies of passives in different text types per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
biographical text	754	782	585	1,093
fiction	344	338	259	357
humanities	943	1,398	2,218	1,742
informational texts	1,070	1,240	1,717	1,150
manuscripts	416	1,460	862	1,059
natural science	886	2,361	2,753	2,541
news reports	974	1,816	1,559	1,322
religious texts	859	944	1,277	770

Some text types show an erratic behavior of frequency, but there are a few visible trends. Informational texts show a pattern similar to the overall frequency, increasing from corpus #1 to corpus #3 and then declining in corpus #4. Humanities and natural sciences show a similar pattern but with more pronounced increases and decreases. Biographical texts and fiction show a different pattern than the overall frequency, being at their lowest in corpus #3 and highest in corpus #4. Not surprising for the only unpublished (and presumably unedited) category, manuscripts show no discernible pattern.

The clearest indication of formalization followed by colloquialization comes from informational texts, the only other category with a word count as high as the fiction category. Different from fiction, however, it is not averse to passive constructions, and, different from scientific texts, it is not marked for formality in particular. The increase in use of the passive voice from corpus #1 to corpus #3, then, can be viewed as part of the general pattern of formalization after the initial standardization of Afrikaans. The decline from corpus #3 to corpus #4, on the other hand, is probably a consequence of colloquialization, also indicated by the decline in the two scientific categories.

Adverbs, Adverbials, and Conjunctions

Adverbs, adverbials, and temporal conjunctions are often used for temporal reference in Afrikaans, sometimes together with grammatical tense but also

sometimes instead of it. The available items are quite diverse in terms of specificity as well as formality and distance. Temporal connectives in general are regarded as more typical of narratives and spoken registers (Van Rooy and Esterhuizen 2011, 80), but certain temporal expressions are clearly more formal and specific than others.

According to Van der Merwe (1996, 91) the items *nou* 'now, *toe* 'when/then,' and *dan* 'then' are some of the most frequently used lexical items in spoken Afrikaans, and in the corpora of this study, they fall under the 70 most frequent words (of between 22,000 and 25,000). These items are used for other purposes than temporal reference, however (Van der Merwe 1996, 92). All three these items are used as discourse markers (7.104–7.106) to varying degrees, where the temporal meaning has faded to the background or disappeared completely:

7.104. <u>Nou</u> ja, kom—dis laat. (corpus #3, Biographical text)
now yes come it.s late
'Come <u>now</u>—it's late.'

7.105. <u>Toe</u>, aanstoot, manne, maak klaar dat julle kan pad vat. (corpus #4, Fiction)
then on.push men make finish that you.PL can road take
'Come on <u>then</u> men, finish so you can hit the road.'

7.106. Wat is <u>dan</u> die rusting vir die mens? (corpus #2, Informational text)
what be.PRS <u>then</u> the resting for the human
'What is rest for humanity <u>then</u>?'

Regarding temporal reference, *nou* is used as an adverb (7.107), while both *toe* and *dan* are used as adverbs (7.108–7.109) as well as subordinators (7.110–7.111). For the purpose of this section, the distinction between adverbs and (adverbial) subordinators is not important—the focus is on expression of temporal reference for narrative *and* discourse purposes:

7.107. Nicht Lenie het <u>nou</u> vertien kinders bij mekaar. (corpus #1, Manuscript)
cousin.F Lenie have <u>now</u> fourteen children with each.other
'Cousin Lenie <u>now</u> has fourteen children altogether.'

7.108. Kort daarna val weer twee of drie pistoolskote, en <u>toe</u> is dit stil. (corpus #2, Fiction)
short there.after fall again two or three pistol.shots and <u>then</u> be.PRS it quiet

'Shortly afterwards two or three pistol shots sounded, and <u>then</u> it was quiet.'

7.109. Die leiers doen na elke biduur verslag . . . Annamarie kan <u>dan</u> daaraan aandag skenk. (corpus #3, Informational text)
the leaders do after every pray.hour report Annamarie can <u>then</u> there.to attention pay
'The leaders report after every prayer hour . . . Annamarie can attend to it <u>then</u>.'

7.110. Hulle is maar "simpel," maar ek was nie myself <u>toe</u> ek hulle geskryf het nie. (corpus #3, Manuscript)
they be.PRS but silly but I be.PST not myself <u>when</u> I them wrote have PTCL.NEG
'They are a little "silly," but I wasn't myself <u>when</u> I wrote them.'

7.111 <u>Dan</u> sit ons weer langs mekaar in die kerk . . . (corpus #2, Fiction)
<u>then</u> sit we again beside each.other in the church
'<u>Then</u> we can sit beside each other in church again . . .'

The temporal and discourse marker frequencies per 100,000 words of the three items are given in table 7.18.

Table 7.18.   Frequencies of temporal and discourse *nou*, *toe*, and *dan* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
nou temporal	277	172	168	140
nou discourse particle	86	41	43	46
toe temporal	277	188	182	212
toe discourse particle	6	5	15	18
dan temporal	124	100	95	79
dan discourse particle	304	166	102	72

A few trends emerge from the data. Both *nou* and *dan* decline in their temporal uses (log likelihood of 130.05 for *nou* and 25.90 for *dan*), but while *dan*'s discourse marker use declines more rapidly than any other category (log likelihood of 467.40), *nou*'s discourse marker use remains fairly stable from corpus #2 to corpus #4. The use of *toe* for temporal reference declines from corpus #1 to corpus #3 and then rises again in corpus #4. Its discourse marker use is fairly infrequent in the first two corpora and then starts to rise slowly.

There are other temporal adverbs and conjunctions, most of which are more specific and formal than *nou*, *toe*, and *dan*. Some examples include the temporal *-dat* compounds discussed earlier and others as illustrated in examples 7.112 to 7.116:

> 7.112. Prys haar <u>wanneer</u> sy suksesvol is. (corpus #4, Informational text)
> *praise her <u>when</u> she successful be.PRS*
> 'Praise her <u>when</u> she is successful.'
>
> 7.113. <u>Terwyl</u> die res van die klas stil lees, kan swak lesers af-sonderlik geneem word om hard-op te lees. (corpus #2, Humanities)
> *<u>while</u> the rest of the class quiet read can poor readers sepa-rately taken become.AUX.PASS.PRS PTCL.INF loud-out to read*
> '<u>While</u> the rest of the class reads silently, poor readers can be taken separately for loud reading.'
>
> 7.114. Heelwat meer eierprodukte is die Verenigde Koninkryk uitgevoer as <u>voorheen</u>, en vir die eerste keer is daar ook 'n klein bietjie na Amerika uitgevoer. (corpus #3, News report)
> *much more egg.products be.AUX.PASS.PST to the United King-dom exported as <u>before</u> and for the first time be.AUX.PASS.PST there also a small bit to America exported*
> 'Much more egg products were exported to the United King-dom than <u>before</u>, and for the first time there were exports to America.'
>
> 7.115. <u>Sedert</u> ons laaste Kongres het ons 'n nuwe leerplan gekry vir ons la'ere skole. (corpus #1, News report)
> *<u>since</u> our last Conference have we a new learn.plan got for our lower schools*
> '<u>Since</u> our last Conference, we have received a new learning plan for our primary schools.'
>
> 7.116. Dis <u>tans</u> die heersende klimaat in die Suid-Afrikaanse speel-goedindustrie. (corpus #4, Informational text)
> *it.s <u>currently</u> the prevailing climate in the South-African toy.things.industry*
> 'It's <u>currently</u> the prevailing climate in the South African toy industry.'

Working with possible candidates from Carstens (1997), Fouché, Van den Berg, and Olivier (2017), and Van Rooy and Esterhuizen (2011), I compiled a list of temporal adverbs and connectives and extracted their occurrences from

the data. I focused on those items that occur 10 times per 100,000 words or more in one or more of the corpora, and the frequencies of these items (used as temporal adverbs and/or connectives) are given in table 7.19.

Table 7.19. Frequencies of formal, specific temporal markers per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
deesdae 'nowadays'	0	1	3	11
intussen 'in the meantime'	5	8	6	11
sedert 'since'	10	13	12	18
sodra 'as soon as'	8	15	6	9
tans 'currently'	5	9	10	22
terwyl 'while'	61	63	65	84
tydens 'during'	2	12	37	37
voorheen 'before'	5	5	7	11
wanneer 'when'	35	64	67	83

Different from the informal items *nou*, *toe*, and *dan*, the more formal and specific items show a general pattern of increase. Comparing the total frequencies of the informal items on the one hand and the formal items on the other, the graph in figure 7.3 shows how formal temporal adverbs and connectives increase proportionally while informal ones decline.
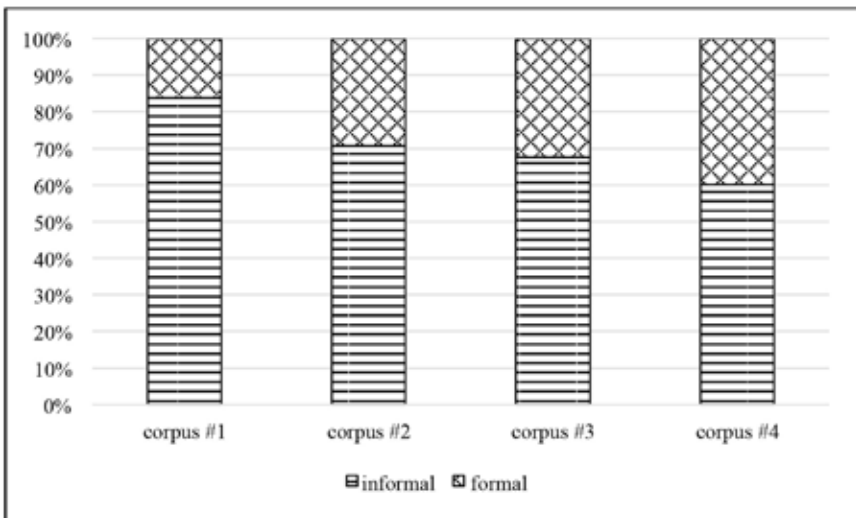
Figure 7.3.   Formal vs. informal temporal adverbs and connectives

In this category, there is a fairly consistent increase in items that denote greater specificity, formality, and distance while informal, general items decline. This would speak to continuing formalization since initial standardization in the early twentieth century.

Other than temporal adverbs and connectives, there are other conjunctions that could point toward formalization and colloquialization. The conjunction *as* 'if' can be used as a conditional (7.117) as well as a temporal connective (7.118). The former is very similar to the more formal *indien* 'if' (7.119), and the latter is similar to more formal *wanneer* 'when' (7.120). In many cases, however, *as* can be interpreted as expressing either a conditional meaning or a temporal one or, possibly, both at the same time (7.121):

> 7.117. <u>As</u> hulle 'n volk wil bly, was dit nou die tyd om te volhard. (corpus #4, Biographical text)
> *if they a nation want stay be.PST it now the time PTCL.INF to persevere*
> '<u>If</u> they wanted to remain a nation, now was the time to persevere.'

> 7.118. Het u al ooit ondersoek gestel in die boeke wat u kinders lees, <u>as</u> hul die skool verlaat? (corpus #1, Humanities)
> *have you.HON already ever investigation put in the books that your.HON children read <u>when</u> they the school leave*
> 'Have you ever investigated the books that your children read <u>when</u> they leave school?'

> 7.119. <u>Indien</u> daar vir meer arbeid begroot word, moet die redes in 'n aparte verklaring genoem word. (corpus #3, Humanities)
> *if there for more work budgeted become.AUX.PASS.PRS must the reasons in a separate declaration named become.AUX.PASS.PRS*
> '<u>If</u> there is budgeted for more work, the reasons should be given in a separate report.'

> 7.120. Ek is bang dat <u>wanneer</u> ek tyd kry om jou verhaal neer te skryf, ek die helfte vergeet het. (corpus #4, Religious text)
> *I be.PRS afraid that <u>when</u> I time get PTCL.INF your.SG story down to write I the half forgot have*
> 'I am afraid that <u>when</u> I have time to write down your story, I will have forgotten half of it.'

> 7.121. Waarom sal ons nog kinders voortbring <u>as</u> daar soveel miljoene werkloses rondloop? (corpus #2, Informational text)
> *why will we still children forth.bring <u>if</u> there so.many millions unemployeds around.walk*
> 'Why would we still bear children <u>if/when</u> there are so many unemployed people?'

A large proportion of the instances of connective *as* is ambiguous in each corpus: between 40 and 60 percent. The conditional and temporal uses are still substantial, however, and in direct competition with *indien* and *wanneer*, respectively. The frequencies per 100,000 words of both conditional *as* and *indien*, as well as temporal *as* and *wanneer*, are given in table 7.20.

Table 7.20. Frequencies of *as*, *indien*, and *wanneer* per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
as conditional	81	63	64	72
indien	8	23	39	38
as temporal	33	21	27	16
wanneer	35	64	67	83

While the trends are not clear cut, there are indications that, broadly speaking, temporal *as* is making way for *wanneer* throughout the century, similar to the situation regarding temporal adverbs in general. With conditional *as* and *indien*, it is slightly more complicated—a substantial decline in *as* and an increase in *indien* from corpus #1 to corpus #2, but *as* remains stable in corpus #3 and rises again in corpus #4, while *indien* increases still in corpus #3 but then remains stable. In very broad strokes, the use of *as* and *indien* as conditionals seems to point toward formalization between corpus #1 and corpus #2 and colloquialization between corpus #3 and corpus #4.

Summary Remarks: Colloquialization and Formalization

In this subsection, I investigated the possibility of colloquialization and formalization in several different grammatical and pragmatic categories. Regarding second-person pronouns, there are signs of colloquialization from corpus #1 through corpus #4, with formal *u* declining throughout and familiar *jy/jou* increasing at the same time. In the use of demonstrative pronouns, the more general and informal *dié* declines between corpus #1 and corpus #2 and then increases until corpus #4, with the more specific and formal *hierdie* and *daardie* showing a mirror image of increase (corpus #1 to corpus #2), followed by decline. Passive constructions show a pattern of increasing usage from corpus #1 to corpus #3 and then a slight decline in corpus #4. Temporal adverbs and connectives have a pattern in direct opposition to second-person pronouns, with a consistent decline in the use of informal variants and an increase in use of formal variants. Finally, another group of related conjunctions—*as* 'if/when,' *indien* 'if,' and *wanneer* 'when'—show a consistent increase of the formal temporal variant *wanneer* and a decline of the informal

as, while the conditional pair point toward first formalization and then colloquialization at the end of the century.

To get a clearer picture of the collective development, I plotted a chart with the formal variants and the informal variants, respectively. The trajectories of the formal variants are presented in figure 7.4.

From the figure, two broad trends emerge. First, all the formal variants apart from honorific *u* show in increase between corpus #1 and corpus #2, and all the formal variants apart from temporal adverbs and connectives show a decline between corpus #3 and corpus #4. The differences in frequency between corpus #2 and corpus #3 go up for some and down for others, but the
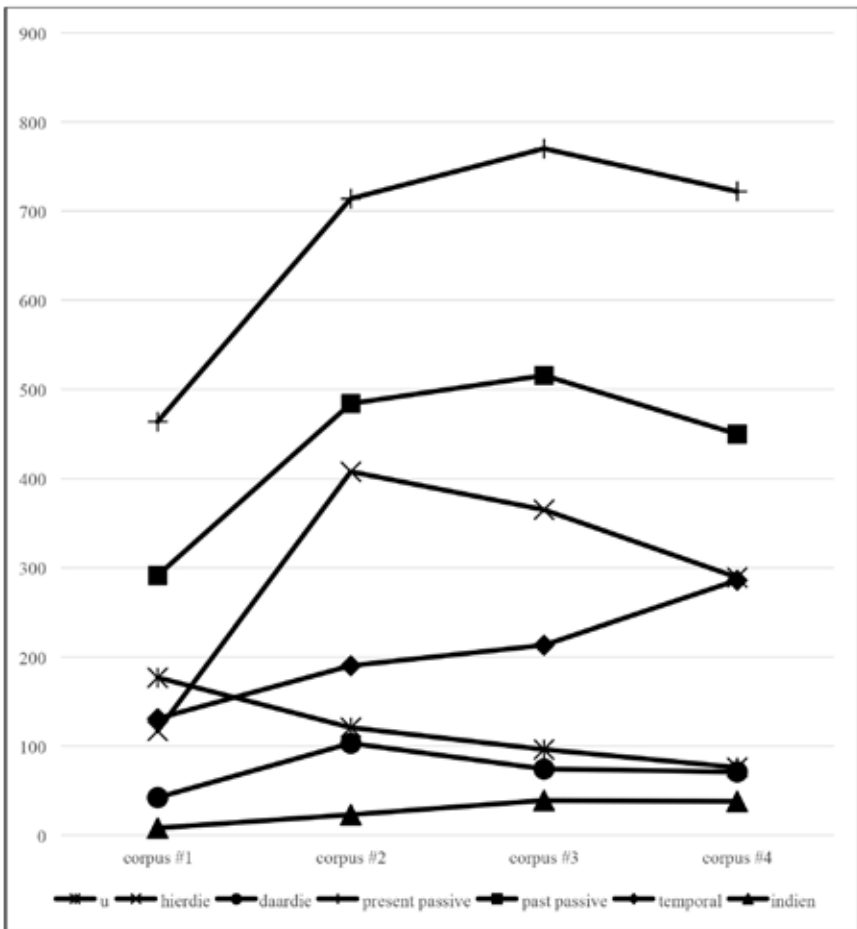
Figure 7.4.   **Patterns of use of formal variants**
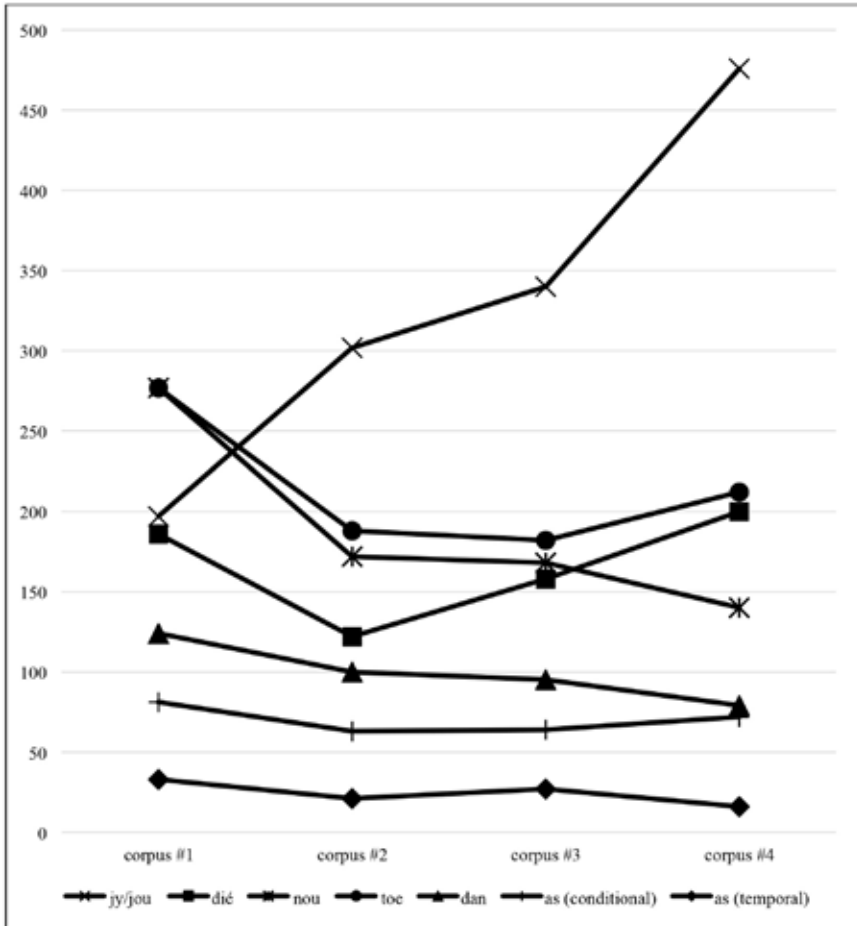
Figure 7.5.   Patterns of use of informal variants

differences are typically not as big as those between the first two or the last
two corpora. The trajectories of the informal variants are given in figure 7.5.

Once again, there are certain trends that emerge from the graph. All the
informal variants apart from the second-person *jy/jou* show a decline from
corpus #1 to corpus #2. Many of the variants show an increase from corpus
#3 to corpus #4—second-person *jy/jou*, demonstrative *dié*, temporal *toe*, and
conditional *as*—even if some variants continue to decline. As with the formal
variants, the differences between corpus #2 and corpus #3 are mostly not very
drastic.

While there are certainly exceptions, a general pattern across different categories of variants emerge—formalization in the early to mid-twentieth century, stability in the first part of the second half of the twentieth century, and colloquialization in the late twentieth and early twenty-first centuries. Against the backdrop of the sociocultural development of Afrikaans, this is not surprising. Formal standardization of Afrikaans commenced in the early twentieth century, and it took a few decades to establish the standard variety. For several decades, there were no macro sociocultural changes that would affect formality and distance in language use. Near the end of the twentieth century, South Africa entered an era of democracy, with a new appreciation for diversity and equality, relaxing some of the formal(ity) principles that used to guide written language use.

DECLINE IN SEXISM: PRONOUNS

A somewhat controversial practice in Afrikaans (as in some other languages, including English) is the use of the male singular third-person pronouns *hy* 'he' and *hom* 'him' for supposedly "gender-neutral" reference. The same practice in English has received attention for several decades now. A number of authors describe the practice of using *male* pronouns for general *human* reference as sexist, as it equates *man* and *human* with each other, implying that *woman* is a *marked human* (Earp 2012, 6; Gastil 1990, 630–31; Mühl-häusler and Harré 1990, 325; Van Rooy 1996, 35). Mühlhäusler and Harré (1990, 235) take the position that "the individual chosen to instantiate or exemplify the species is usually a member of the male sex. That is just bad biology as well as being a genuine instance of sexism." Gastil (1990, 360) points out that some authors acknowledge the sexist origins of this prac-tice but even in the face of research exposing its issues deny its negative consequences in contemporary language use. An Afrikaans example is an endnote after chapter 1 in the subsequent editions of Carstens's *Norme vir Afrikaans* (Norms for Afrikaans) where the author explains that male pro-nouns are used with gender-neutral intent throughout the book for the sake of convenience (Carstens 1989, 54; 1991, 31; 2003, 32; 2011, 34; 2018, 38). In his investigation of this phenomenon in Afrikaans, Van Rooy (1996, 35) draws on Foucault's concept of discursive practice, a method to give voice to certain discourses while silencing others through procedures that determine which discourses are "acceptable." He positions the "generic" male pronouns within a sexist discursive practice that denies women their rightful place in the discourse (Van Rooy 1996, 35). The sexism underlying the practice of using male forms for gender-neutral reference comes to the surface through

research that demonstrates how this practice is *not* gender neutral but, rather, favors a male interpretation (Everett 2011; Gastil 1990; Van Rooy 1996). In contrast, pronouns that are really gender neutral lead to a more truly gender-neutral interpretation (Everett 2011, 149). Furthermore, gender-neutral pronouns also make room for the often erased nonbinary people who do not fit into or identify with either the male or the female gender.

Van Rooy (1996, 35) indicates that the use of male pronouns for gender-neutral reference creates the archetype that humans are by default male, which in turn creates a negative female (and nonbinary) space where women (and others) are always expressed in terms of maleness, becoming "nonmales."

In English, the option to use *they* as singular, generic gender-neutral pronoun has existed for several centuries and continues to be utilized as such today, even if it has been discouraged by prescriptivists for its supposed grammatical incongruence since the eighteenth century (Everett 2011, 134; Mülhäusler and Harré 1990, 232).

The situation for Afrikaans is, however, slightly different. The plural third-person *hulle* 'they' does not fulfill the same function in Afrikaans as *they* in English to the same extent. This leaves Afrikaans users with limited options regarding gender-neutral anaphoric reference to singular referents. The use of *hy/hom* in such contexts is sometimes regarded as unproblematic or even recommended above combination forms (*hy/sy* 'he/she')—by two authors from the early twentieth century (Du Toit 1902, 9; Malherbe 1917, 78) as well as others from the later twentieth century (Carstens 1989, 54; Combrink and Spies 1986, 69; Van der Merwe and Ponelis 1982, 85). This means that *sy/haar* 'she/her' is used only when referents are marked as female (Ponelis 1979, 68; Van der Merwe 1997, 32), and people who do not fit into the binary are completely ignored.

A number of studies point toward general sexism in Afrikaans language use, at least until recently. Van Rooy (1996) does an experiment that shows how Afrikaans-speaking students typically interpret male "generic" pronouns as referring to males, not as gender neutral, in spite of the knowledge that the male form is used for gender-neutral reference. Beylefeld and Van Jaarsveld (1991, 1994) investigate sexism in Afrikaans dictionaries, and they find that meaning and usage guidelines continue to erase women and do not provide the linguistic tools to communicate about women as the equals of men (Beylefeld and Van Jaarsveld 1991, 12). They also show that the world of Afrikaans speakers appears to be predominantly male in this context (Beylefeld and Van Jaarsveld 1994, 47). Atkinson (1995) and Schwerdtfeger (1989) investigate example texts and sentences from second- and first-language Afrikaans textbooks, and both find an abundance of sexist language use, an overwhelming preference for anything male, and negative stereotyping of

women. Schwherdtfeger (1989, 34) strikingly claims at the end of the article that if someone from a different planet would visit the earth and immerse themselves in a library for a time to learn about our species, history, and cultures, we should not be surprised if they deduce that the planet is populated by 95 percent men (who lead interesting, active lives) with only 5% women (who are busy with insignificant things in the background). In addition, non-binary people remain completely invisible.

To illustrate the social prominence of men even in contemporary Afrikaans language use, I give the overall frequencies of male, female, and plural third-person pronouns in table 7.21. (As I mentioned before, Afrikaans has no gender-neutral singular third-person pronouns that refer to *people*, instead of *things* like *dit* 'it.')

Table 7.21.   Frequencies of third-person pronouns referring to specific human referents per 100,000 words

	corpus #1	*corpus #2*	*corpus #3*	*corpus #4*
hy/sy (male singular)	899	1,149	954	1,004
sy/haar (female singular)	259	357	567	380
hulle (plural)	424	401	400	377

In all the corpora except corpus #3, authors refer to men up to three times as often using third-person pronouns than to women, and even in the atypical corpus #3, pronominal reference to men is almost twice as high as to women. The corpora were compiled from a wide variety of texts and text types on a wide variety of topics. The frequencies in table 7.21 are at least some indication of the continuing social prominence of men—people write about (singular) men much more often than about women or groups when using third-person pronouns. Van Rooy's (1996, 34) study indicates that this is the case for spoken Afrikaans as well (the study is fairly old, but if the corpus data are any indication, this is pretty much still the case).

The underlying sexism pointed out in the studies I mention above is probably still relevant today, at least to some extent. There are other indications, however, that at least in certain ways, the sexist use of pronouns is declining. The category of pronouns where this decline is most visible is generic pronouns. Table 7.22 gives the frequencies per 100,000 words of all the variants that are used for generic pronominal reference in some way or another.

The first obvious change in generic pronoun use is the significant drop in use of the male forms *hy* and *hom*. In addition, corpus #4 contains substantial use of the female *sy* and *haar* for gender-neutral reference, which is completely absent in corpus #1 and corpus #2 and has a single occurrence in

Table 7.22.   Frequencies of generic pronouns per 100,000 words

	corpus #1	corpus #2	corpus #3	corpus #4
('n) mens 'one'	95	60	56	73
jy/jou 'you' singular	112	56	61	288
u 'you' honorific	6	40	28	<1
hulle 'they'	5	2	3	2
hy/hom 'he/him'	178	91	77	18
sy/haar 'she/her'	0	0	<1	36
hy/sy and *hom/haar* 'he/she' and 'him/her'	1	1	1	9

corpus #3. Finally, the combination forms *hy/sy* and *hom/haar* that combine the male and female pronouns are used only sporadically in the first three corpora but rise in frequency in corpus #4. These trends point toward a decline in sexist generic pronoun use, but the situation is slightly more complicated than it seems at first glance.

All of the female pronouns used for gender-neutral reference in corpus #4 come from three texts in the corpus, two articles from magazines aimed at women, and a book aimed at the parents of toddlers. The female forms are used throughout to refer generically to *children*: babies, toddlers, and occasionally teenagers and typically in relation to their mother (the presumed reader of the texts). In the magazine *Sarie* (a traditional Afrikaans woman's name), there is an article about parenting and managing anger toward your children, and *sy/haar* is used to refer to children in general, as exemplified in example 7.122:

> 7.122.   Bly weg van stellings wat jou kind skuldig laat voel of wat
> <u>haar</u> onveilig en bedreig laat voel. (corpus #4, Informational text)
> *stay away from statements that your.ꜱɢ child guilty let feel or that <u>her</u> unsafe and threatened let feel*
> 'Stay away from statements that will let your child feel guilty, or that will let <u>her</u> feel unsafe and threatened.'

In another article, also in *Sarie*, there are tips for potty training a toddler, and the female form is also used to refer to the toddler (7.123):

> 7.123.   As <u>sy</u> weet dat jy sekere verwagtinge koester, kan <u>sy</u> gestres
> raak. (corpus #4, Informational text)
> *if <u>she</u> know that you.ꜱɢ certain expectations harbor can <u>she</u> stressed become*

'If <u>she</u> knows that you harbor certain expectations, <u>she</u> can become stressed.'

The third text that employ female pronouns for gender-neutral reference is a book, *Slaapgids vir babas en kleuters* (Sleeping guide for babies and toddlers), where the author refers to babies and toddlers in general with the female pronouns (7.124):

> 7.124. Masseer <u>haar</u> met spesiaal gemengde olie met laventel en/of kamille, en streel <u>haar</u> met lang, ferm hale van kop tot tone. (corpus #4, Informational)
> *massage <u>her</u> with special blended oil with lavender and/or chamomile and caress <u>her</u> with long firm strokes from head to toe*
> 'Massage <u>her</u> with specially blended oil with lavender and/or chamomile, and caress <u>her</u> with long, firm strokes from head to toe.'

The high concentration of female gender-neutral pronouns in a limited number of texts suggests that the practice is limited; the fact that it consistently refers to children is also an indication that it might be used only in a very specific context at this time. The number of texts that use male gender-neutral pronouns in each of the corpora are the following:

- Corpus #1: 55 texts
- Corpus #2: 34 texts
- Corpus #3: 27 texts
- Corpus #4: 16 texts

In contrast, no texts use female gender-neutral pronouns in corpus #1 and corpus #2, with only one in corpus #3 and three in corpus #4. This shows that, even if the raw frequency of male gender-neutral pronouns is lower in the last corpus than female gender-neutral pronouns, its use is still more widespread.

There are also some complexities regarding combination forms like *hy/sy* 'he/she' and *hom of haar* 'him or her.' This practice still erases anyone who does not identify as male or female, although this issue has not received any public attention that I could find regarding Afrikaans. Furthermore, while the use of combination forms has not been widely and explicitly criticized since the nineteenth century, as is the case for English (Everett 2011, 134), it is sometimes portrayed in a negative light. Ponelis (1979, 59), in a descriptive syntax of Afrikaans, describes it as nothing less than pedantic, and prescrip-

tivists echo this sentiment (Combrink and Spies 1986, 69; Van der Merwe and Ponelis 1991, 97). The more recent usage guide by Müller (2003, 106) also advises against using combination forms, but instead of suggesting the male forms, she points toward indefinite forms, such as *elkeen* 'each one.' Still, table 7.22 shows that the combination forms are used—and more so recently than before.

One development toward true gender neutrality is the use of the second-person singular *jy/jou* 'you' for generic reference, as I briefly discussed in the section on colloquialization and formalization. A similar development (in some respects) occurred in Dutch, with an increasing use of *je* 'you' instead of *men* (literally 'people,' also translated as 'one' and 'they') (De Hoop and Tarenskeen 2015, 164). The appropriation of second-person pronouns for generic reference is widespread typologically (Gast et al. 2015, 161). Still, it is possible that the increasing preference for *jy/jou* instead of male forms for gender-neutral reference has been hastened and/or intensified by feminism and the general strive toward equality in society at large and that the development will continue for the foreseeable future. At the very least, it suggests a decline in linguistic sexism regarding pronoun use, even if authors still write about men much more than women and others.

CONCLUDING REMARKS: DISCURSIVE AND SOCIOCULTURAL CHANGES

In this chapter, I focused on language change that exceeds the boundaries of what is traditionally regarded as grammar. First, I investigated connectives and changes in how they are used to structure texts in different ways. The two subordinators for signaling projected complement clauses *dat* 'that' and *of* 'if' both decline throughout the century. This implies that users of Afrikaans increasingly prefer bare complement clauses, where independent word sequence remains preserved, above the use of subordinate complementizers, which trigger subordinate word sequence. In contrast, subordinators that include *-dat* as the second part of compounds, such as *sodat* 'so.that,' *voordat* 'before.that,' and so on, show a general trend of an increase in frequency even as certain archaic forms disappear. While new compound forms are not added in written Afrikaans during the century, some established forms rise in frequency, such as *sodat* 'so.that,' *voordat* 'before.that,' and *nadat* 'after.that.'

The next section focused on general trends of formalization and colloquialization. Focusing on variants in different grammatical categories that are regarded as either formal or informal in the literature, I traced general patterns of increase and decline in frequency. A general trend of formalization

between corpus #1 and corpus #2 emerged from the data, with a measure of stability between corpus #2 and corpus #3, followed by a general trend of colloquialization between corpus #3 and corpus #4. This is not unexpected in light of the initial standardization of written Afrikaans in the early twentieth century as well as democracy and increasing equality at the end of the twentieth and early twenty-first centuries.

In the final section, I focused on linguistic sexism. While the content of the texts included in the corpora continue to favor males and maleness, there are signs of declining sexism regarding pronoun use, especially generic pronouns, where an increase in the second-person singular *jy/jou* 'you' continues while the male third-person *hy/hom* 'he/him' declines. This means that male preference makes way not for female preference but, rather, for genuinely gender-neutral generic pronouns, also leaving room for people who fall outside the traditional (Western) gender binary.

8

Conclusion

RETURNING TO THE RESEARCH QUESTIONS

In chapter 1, I posited two related research questions:

1. What is the nature and extent of linguistic changes in written Afrikaans from 1911 to 2010, with a focus on temporal reference, pronouns, the genitive, and connectives?
2. What are the differences (if any) between internal and external language change, and, based on that, what can we learn from the data regarding language contact and its influence on language use?

In this final chapter, I return to these questions more explicitly. In chapters 5 to 7, I attended to the first question in detail. In the following section, I summarize the relevant findings to answer question 1. Subsequently, I return to question 2 and to particular issues that emerged from theoretical discussions in chapter 2 especially. Finally, I reflect on certain shortcomings of this study and make some recommendations for future research.

TYPES OF LANGUAGE CHANGE

Paradigmatic Changes

There are several instances of paradigmatic change in Afrikaans in the course of the past century. First, there are some changes in the verbal paradigm of Afrikaans, specifically regarding the preterit. The high-frequency verbs,

191

mainly modal auxiliaries, retained the productive use of the preterit (irregular past-tense forms), while lower-frequency verbs, such as *weet* 'know' with *wis* 'knew' and *mag* 'may' with *mog* 'might,' lost the productive use of the preterit. The number of verbs with an irregular preterit form declined, and with it the paradigm of verbal past-tense formation became slightly more regular than before.

Regarding the pronominal system, there were also some minor changes. The reduced third-person plural form *hul* is now used mostly as an attributive possessive pronoun, while the full form *hulle* is used in most other functions, mainly third-person plural anaphoric reference. Reflexive pronouns, or reflexive uses of personal pronouns, increasingly take the extended *-self* form, even when it is not obligatory; this is a development in the direction of the English reflexive paradigm. The third-person inanimate form *dit* sees a decline in use as a dummy subject and an increase in use for sentence-anaphoric reference rather than sentence-cataphoric reference.

A further paradigmatic change is the loss of the Dutch morphological genitive. In the course of the twentieth century, the Dutch genitive went from productive (albeit rather infrequent) to almost completely absent apart from a few fixed expressions. A whole morphological paradigm all but disappeared from Afrikaans language use, with only the periphrastic genitive with *van* 'of' and *se* ''s' remaining.

Grammaticalization

Several instances of grammaticalization can be observed in the Afrikaans usage data. In the first instance, there is the use of *gaan* 'go' for future reference. While *sal* 'will' is still the preferred future auxiliary in written Afrikaans, *gaan* underwent significant context expansion and a proportional increase of grammatical uses compared to lexical uses.

Regarding pronouns, there are some grammaticalization developments as well. The third-person inanimate *dit* 'it,' followed by *is* 'is/are,' is often contracted to the form *dis* 'it's.' This contracted variant increases in frequency in the course of the century. The generic pronoun *('n) mens* '(a) person' has two variants—the one with the indefinite article and the one without—and the shorter version without the article also increases in frequency at the expense of the longer version. The quantifier *enige* 'any' is sometimes merged with indefinite pronouns, such as *iemand* 'somebody' and *iets* 'something,' and the merged variants, such as *enigiemand* 'anybody' and *enigiets* 'anything,' show an increase in frequency.

Finally, the genitive particle *se* ''s' grammaticalizes throughout the century, with a consistent and significant increase in frequency as well as gram-

matical contexts. The most important factors contributing to the increasing preference for *se* over *van* are possessor animacy and the semantic relation expressed by the genitive construction.

Discursive and Sociocultural Changes

Turning to discourse-level changes, the data show a decrease in the use of two complementizers: *of* 'if/whether' and *dat* 'that.' This could be an indication of an increasing preference for bare complement clauses. Some other connectives also show signs of change, but these link to trends of formalization and colloquialization.

In terms of formalization and colloquialization, a trend emerged from the data. For the first half of the twentieth century, there are clear signs of formalization following the initial standardization of Afrikaans. In the late twentieth and early twenty-first centuries, there is a turn in the trend, with signs of colloquialization in several categories. Items involved in these trends include honorific second-person *u* and familiar second-person singular *jy* and *jou* 'you,' also used as generic pronouns; the demonstrative pronouns, with more formal *hierdie* 'this/these' and *daardie* 'that/those' and informal *dié* 'this/these/that/those'; passive constructions, typically associated with formality and distance; and certain adverbs, adverbials, and conjunctions, such as the informal *nou* 'now,' *toe* 'then/when' and *dan* 'then,' and the formal *sedert* 'since,' *terwyl* 'while' and *wanneer* 'when.'

The final type of sociocultural change under investigation is the use of sexist language, with a clear decline in sexism observed in pronoun use regarding straightforward reference to females with pronouns as well as the choices of generic pronouns (male, female, combination forms, and gender-neutral forms).

INTERNALLY VERSUS EXTERNALLY MOTIVATED CHANGE

What Is the Difference?

A simple difference between internally and externally motivated linguistic change is that internally motivated change originates from within a given language system and externally motivated change from outside that system. However, the differentiation of separate languages, as well as the concept of a language as a self-contained system, is problematic. If language is conceived as a dynamic process (or in terms of languaging), with the primary goal of communication, then the distinction between internal and external becomes not only difficult but also, in some ways, redundant.

The first important question is why this distinction should be made in the first place. If language cannot be easily and clearly delineated from the rest of reality and if languages cannot be easily differentiated from one another, why bother with it to begin with? In short, the answer is that just because the differences are difficult to determine does not mean that they do not exist and still have some value.

Written language use is typically contained in a text, and a text is seemingly contained and bounded. Intertextuality and reference to extralinguistic reality blur the line, however, and remind us that language use is always anchored in extralinguistic reality. Even if the boundaries are fluid and variable, there is some difference between the linguistic and the nonlinguistic, and internally motivated change is when a change originates from the language itself, while externally motivated change originates from outside the language.

Furthermore, different language are traditionally distinguished from each other, and while these distinctions are at least partially based on extralinguistic factors, it can be important if the relevant speech communities treat it as important. Externally motivated change is, then, when the origin or motivation for a change can be found outside of language as such or outside of the particular language at hand. So, while the boundaries between the linguistic and the nonlinguistic and between different languages are variable and fluid, there are cases where the origin or motivation of a change can be clearly traced to one or the other.

In cognitive-linguistic terms, the next question is whether processes of linguistic change, such as grammaticalization, analogy, and reanalysis, are internal or external. The process proceeds not within the language system but, rather, in the minds of the speakers who use the language. While it is indeed the case that the process is not internal to the language itself, the source of the item undergoing change is the language itself, and the reasons behind the change are tied to the (use of) the language. For this reason, it is regarded as internal change.

In this study, there are two types of externally motivated change: influence from other languages and extralinguistic factors, such as standardization and sociocultural circumstances. Grammaticalization is a mainly internally motivated change, even though it can be influenced or hastened by external factors, as was the case during the development of Afrikaans. Analogy is also, within the parameters of this study, an internally motivated change, even though the initial overwhelming regularization took place during the original development of Afrikaans in intense contact situations. However, in the time period covered by this study, the vernacular varieties of Afrikaans have already been established to a large extent, and the small-scale boundary shifts observed in the study are mainly internally motivated.

Influence from Language Contact

In this whole study, only one instance of influence from English on written Afrikaans was confirmed: the increasing use of reflexive pronouns with -*self*, whether it is necessary for disambiguation or emphasis or not. While other studies have focused on the (sometimes extensive) influence of English on Afrikaans lexicon, pronunciation, and idioms (so-called Anglicisms), there is little evidence for significant English influence in the grammatical categories investigated in this study. The use of reflexive pronouns with -*self* increases proportionally, but it is not particularly frequent compared to many other pronouns, and it does not cause further paradigmatic changes in the pronoun system.

Nicolle (2012, 388) indicates that tense and aspect resist influence from other languages to some extent because it is part of a paradigmatic system. While the use of reflexive pronouns with -*self* obviously has nothing to do with tense and aspect in itself, pronominal systems are also paradigmatic, and a change in one category has the potential to affect the rest of the system. This is not so much the case for using -*self* with reflexive pronouns, however, but the principle could explain why English influence is not more prominent in the grammatical categories investigated in this study. There are few contexts where "simple" borrowing could take place without further paradigmatic implications.

Interestingly, there are cases of Afrikaans influence on South African English, among others in terms of the typical modality that *must* conveys (Wasserman and Van Rooy 2014) and certain usage patterns of *now* that show influence from Afrikaans *nou* (Jeffery and Van Rooy 2004). This goes contrary to the assumption that only "bigger," international languages affect "smaller" local languages and not the other way around, especially in situations with high levels of individual bi- or multilingualism. What is clear, however, is that continued contact between speakers of different languages and multilingual speakers results in mutual influence (or transfer) in one way or another.

Standardization, Prescriptivism, and Sociocultural Circumstances

In contrast to the effects of influence from other languages, the influence from extralinguistic factors in language change in written Afrikaans is more prominent.

The first extralinguistic factor behind language change, with a potentially very direct influence on language use, is standardization. Time and again, with many of the phenomena under investigation, there are obvious changes from corpus #1 (the first stages of standardization) and corpus #2 (after standard norms and conventions have been widely accepted in the written

language). There are especially signs of formalization between corpus #1 and corpus #2 that do not continue at the same pace for the rest of the century or that even change course later in the century. Another clear sign of standardization is the radical reduction in variation. Many of the phenomena still have several variants in corpus #1, with only one or very few that remained in corpus #2—temporal auxiliaries, many pronouns, variants of the *se*-genitive, and so on.

As can be expected, a tradition of prescriptivism was established with the standard variety. In some cases, the corpus data reflect the prescriptions and proscriptions put forward by normative sources. This is not always the case, however: the use of reflexive pronouns with *-self* continues to increase in spite of consistent proscriptions and the decline in the use of male pronouns as "gender-neutral" pronouns in spite of several negative comments on the alternatives. While male gender-neutral pronouns are not strictly prescribed, it is often put in a positive light, especially in comparison to combination forms, such as *hy/sy* 'he/she.' Still, the use of male pronouns in gender-neutral contexts continues to decline, but it is replaced with the gender neutral *jy/jou* 'you' rather than combination forms or the plural.

This trend also involves the final extralinguistic factor explored in this study: sociocultural circumstances. The use of male and female pronouns, as well as the profile of generic pronouns, changes in accordance with sociocultural circumstances, especially under the influence of feminism. There is a decline in female reference from corpus #3 to corpus #4, but it can be at least partially ascribed to the topics that happened to be included in the corpora—a large proportion of texts in corpus #3 are aimed at women and are on topics traditionally associated with female roles, such as mosaic, cooking, child rearing, and prominent female figures in society. The profile of generic pronouns, however, confirms the decline in linguistic sexism and that people other than men are gaining more visibility in language use and topics of popular texts. The increasing use of familiar *jy/jou* 'you' also links with colloquialization. The most obvious signs of colloquialization can be seen between corpus #3 and corpus #4, reflecting the extensive sociopolitical changes in South African society and the use of Afrikaans in the public sphere.

To conclude, then, it would seem that the externally motivated changes in written Afrikaans observed in this study are due mostly to extralinguistic factors rather than influence from other languages, such as English.

THE ISSUE OF INNOVATION VERSUS SELECTION

One particular issue regarding language contact and influence relates to the distinction between the origin of a new variant, or innovation, and the spread

of a new variant, or selection. While the distinction is important in principle, the limited time frame covered in this study, as well as the focus on grammatical and discourse phenomena rather than on, for example, lexicon, do not create the ideal context to explore this matter in detail.

Few instances of innovation are recorded in this study—the changes I reported are typically frequency changes that show the spread of a new variant, not its first appearance. In the one case of confirmed English influence, English is also not the source of the variant. Rather, it is a case of covert transfer, where exposure to and the use of English could be contributing factors in the increasing use of reflexive pronouns with *-self*. The *-self* forms are present in all the corpora. This is similar to the type of findings that Leech et al. (2009) and Mair (2006) report in their studies.

There are instances where a new usage surfaces only in the latter part of the twentieth century, which does not necessarily mean it was completely absent from language use before that, only that is was not common enough in written language to be recorded in the corpus data. One variant that emerges later is the use of the female third-person singular pronouns *sy* 'she' and *haar* 'her' as gender-neutral pronouns. This usage is absent from the corpus data early in the century, with a single occurrence in corpus #3 and several occurrences in corpus #4. This usage occurs in only three texts in corpus #4, so while it is still not widespread, it is completely absent in the early twentieth century and could be an innovation.

Another new variant recorded later in the corpus data is the shortened form *daai* of the distal demonstrative pronoun *daardie* 'that/those.' This shortened form is used once in corpus #3 and more than two dozen times in corpus #4. It is regarded as informal and probably much more commonly used in spoken language—it could have been established in spoken language much earlier but accepted in written language only late in the century. Because *daardie* itself is a fairly new innovation in Afrikaans (nineteenth century), however, it is probable that the shortened *daai*, based on *daardie*, is an even newer innovation. The data do confirm, though, that it is not yet an established, widely accepted form in written Afrikaans.

Other variants that emerge in the data only later in the century include indefinite pronouns *enigeen* 'anyone' (corpus #2) and *enigiemand* (corpus #3) and a number of connectives: in corpus #2, we find the first instances (as connectives) of *boonop* 'moreover,' *daarbenewens* 'in addition,' *namate* 'as,' and *sedertdien* 'since then'; in corpus #3, *dienooreenkomstig* 'accordingly' and *alternatiewelik* 'alternatively' emerge as connectives; and in corpus #4, we find the first instance of *volgende* 'next' as discourse connective.

Apart from these instances, all the other changes reported in the study are cases of spread and selection and not innovation. It is not unexpected in a

study that covers only one century of language change, especially when fo-
cusing on grammatical and discourse-level changes.

It is important to note that in the majority of cases where Afrikaans
linguists and prescriptivists refer to English influence on the grammar of
Afrikaans, they are not referring to innovations or direct borrowings from
English but, rather, instances of covert transfer. With the high levels of
Afrikaans–English bilingualism in South Africa, covert transfer is not a
surprising development. There are many instances of cognate forms, such
as Afrikaans *moet* and English *must*, where the one has subtle influences on
the other (Wasserman and Van Rooy 2014). Surface similarities can lead
speakers to intuitively group Afrikaans and English constructions together
in their feature pools regardless of certain subtle or even obvious differences
between the constructions as they are used in the different languages. When
such constructions are grouped together by speakers, interlingual transfer can
easily happen in any direction, causing certain shifts or changes in meaning
or usage to occur in either or both of the languages. Unlike English, which
is spoken all over the world, with many different varieties, most of which
are *not* in contact with Afrikaans, there are no varieties of Afrikaans where
speakers have not experienced prolonged contact with English and high lev-
els of bilingualism. This means that, even in cases where English influence
seems obvious, we cannot with certainty determine what the trajectory of a
change would be without that influence, especially if it involves an already
existing Afrikaans construction.

FURTHER THEORETICAL REFLECTIONS

There are several theoretical matters that emerge from the earlier theoretical
discussions and results of this study.

The first matter regards the way in which language change proceeds when
in contact. In principle, language changes when it is used, whether speakers
are multilingual or in contact with speakers of other languages or not. The
presence of language contact is relevant in two ways: first, language contact
could result in innovations that would not have emerged otherwise, and,
second, language contact could result in covert transfer, where a similar con-
struction is used in the different languages, mutually reinforcing each other.
The former is a qualitative change and the latter a quantitative one.

If there is a similar construction in two or more languages in contact,
whether through borrowing, shared origins or coincidence, bi- or multilin-
gualism could lead to the particular construction gaining a high frequency
for some speakers, as it is (collectively) used more frequently than it would

be in the case of exposure to only one of the languages. Because exposure to high frequency leads to familiarity and a construction being cognitively entrenched, speakers would find it easier to use the construction that is similar in the different languages because it would be easier to recognize and recall. For bi- or multilingual speakers, especially those who frequently use or are exposed to the different languages, the shared construction would be activated more frequently than it would be with only one language in use. Frequency-based changes would then be absent or proceed at a slower pace in contexts without contact than those where speakers are exposed to shared constructions. For this principle to hold, it is necessary that a substantial portion of the language community be bi- or multilingual; otherwise, the number of idiolects that facilitate the contact influence might be too small for a significant change to develop.

This conceptualization links with Mufwene's view that contact-induced changes do not proceed differently from other types of change but that the difference is quantitative rather than qualitative. It also reminds us that mutual influence in contact situations is normal, as the linguistic systems of different languages are not strictly separate in the individual's mind and function according to the same cognitive mechanisms and principles. The distinction between different languages is not necessarily applicable in an individual's mind and is maintained only by external factors that an individual has to apply at their discretion. In the pressures of day-to-day language use and attempts at communicating effectively, mutual influence between languages can be expected, even when speakers deliberately attempt to keep the systems or codes apart. In this sense, the osmotic nature of boundaries between languages that Mufwene proposes becomes more apparent than many linguistic theories would have. He conceptualizes the feature pool to which speakers have access in terms of *language*, not in terms of *a language*, facilitating the reality of mutual influence in contact situations. The feature pool can also be linked with Bybee's exemplars and exemplar models—speakers mentally store exemplars of all the language use they are exposed to and actively use, and this collection of exemplars is the feature pool that speakers access during their own language production.

This topic once again reminds us why it is important to be careful of conceptualizations of language as a self-contained system, rather focusing on languaging and language as dynamic processes, continually formed and created through active use, with grammatical structures as abstractions and generalizations from this usage. It also, once again, reminds us how fluid and fuzzy the boundaries between different languages are. In light of this, the sense of superiority and negative and dismissive attitudes of many Afrikaans prescriptivists and linguists toward English influence on Afrikaans (e.g.,

Basson et al. 1968, 38; Boshoff 1956, 58–59; 1964, 39; Botha and Burger 1926, 202; Le Roux 1939, 77; Malherbe 1917, 16; Pienaar and Langenhoven 1932, 84; Raidt 1989, 116; Van Schalkwyk and Kroes 1979, 160; Van Schalkwyk and Viviers 1977, 43) are unsustainable, and no measure of language pride and/or language nationalism (e.g., Botha and Van Aardt 1987, 29; Bouman and Pienaar 1924, 17; Grobler 1993, 49; Le Roux, Malherbe, and Smith 1917, iii; Malherbe 1917, 9; Pienaar and Langenhoven 1932, 4–5) will prevent mutual influence between languages in a contact situation as intensive as that between Afrikaans and English in South Africa.

Tying to this, the next theoretical matter at hand regards the fuzzy boundaries between linguistic and nonlinguistic phenomena and how these also mutually influence one another. The sociocultural context within which language is used can have far-reaching implications for the way language is used and cannot be ignored when exploring possible explanations for language change. However, it remains important to be careful when looking at this relationship and not grasp at random straws that might explain some or other linguistic phenomenon. When sociocultural circumstances bring about language change, it would not manifest in one isolated instance. Further corroboration of the effect of particular social conditions should be sought in language use in general. It is also important to keep in mind that there are many different systems functioning within any particular language or variety, any number of which may also be involved in a particular change.

This leads to my next observation: one change or trend should not be highlighted outside of context. This context might include extralinguistic factors as well as the rest of the language structures that function together with a particular phenomenon. Even if languages are not self-contained systems, systematicity and regularity exist in many aspects of language use. For example, if we regard the frequency of the adverbs and connectives *toe* 'then/when,' *'nou* 'now,' and *dan* 'then' in relation to more specific, formal variants, it would seem like formalization continues unabated throughout the twentieth century and into the twenty-first, as I prematurely claimed in Kirsten (2016). However, considering all the other manifestations of formalization and colloquialization in the data, a more complex pattern emerges. While there is a clear trend of formalization between corpus #1 and corpus #2, this trend flattens between corpus #2 and corpus #3 and changes course between corpus #3 and corpus #4 to rather show signs of colloquialization. It is important, then, to not base generalizations on one phenomenon or even one category of phenomena.

It is further important to remember just how messy and seemingly chaotic language use can be. Simplistic explanations and absolute claims rarely capture the whole picture. While it is quite impossible, with our present facilities,

to include every and all possible contributing factors in linguistic explanations, it is necessary to be as thorough as possible and not make claims that cannot be scientifically substantiated.

The full potential of the corpora in this study has also not been explored yet, and it can be used for studying many other topics—other syntactic and morphological changes, specific constructions such as infinitive clauses and constructions, modality, lexical and grammatical aspect, and so on. A more deliberate investigation of genre conventions could also bring important insights to the table and enable us to more accurately contextualize grammatical and other changes in the data.

LOOKING AHEAD

There are some partially resolved and unresolved issues in this study. Some of these could be better addressed by using larger, more comprehensive corpora. Using corpora from more intervals, such as every decade and not every third decade, will also give a finer-grained picture of certain developments.

An unfortunate shortcoming of this study, as with many historical studies, is the complete lack of spoken data. In spite of concerted efforts by many different linguists, spoken data from even early in the twentieth century have remained elusive. More recently, the Ponelis Spoken Corpus from the 1970s has been converted to electronic format and contains data from different registers and geographical areas. This could be quite fruitfully compared to corpus #3 from this study to investigate differences between spoken and written Afrikaans. Recently available online Afrikaans corpora, such as on the corpus portal of the Virtual Institute for Afrikaans (Virtuele Instituut vir Afrikaans 2018), also include informal online language use and recent spoken data from radio broadcasts that could be incorporated into future studies. Another step in expanding the written corpora of this study could be to find published and manuscript data from other varieties than the standard, to be more representative of the whole speech community, not only the educated (often white) speakers.

The findings of this study are thus not representative of all of Afrikaans but merely the written Afrikaans of mostly white speakers. It can, however, serve as a starting point and contextualization of future studies, investigating many of the unresolved issues that remain.

References

Abraham, Werner. 1999. "Preterite Decay as a European Areal Phenomenon." *Folia Linguistica* 33, no. 1: 11–18.

Aijmer, Karin, and Anne-Marie Simon-Vandenbergen. 2011. "Pragmatic Markers." In *Discursive Pragmatics*, edited by Jan Zienkowski, Jan-Ola Östman, and Jef Verschueren, 223–47. Amsterdam: John Benjamins.

Atkinson, A. 1995. "Seksisme in taal en implikasies vir tweedetaalonderrig" [Sexism in language and implications for second language teaching]. *Dubbelfluit* 1, no. 1: 1–15.

Baker, Paul. 2009. "The BE06 Corpus of British English and Recent Language Change." *International Journal of Corpus Linguistics* 14, no. 3: 312–37.

Basson, Michiel A., Jan Kromhout, Petrus G. Nel, and Jan H. Senekal. 1968. *Afrikaans vir die student: Die suiwer woord* [Afrikaans for the student: The pure word]. Johannesburg: Voortrekkerpers.

Beylefeld, Adri, and Gert van Jaarsveld. 1991. "Taalseksisme as 'n vorm van eensydige leksikografie" [Linguistic sexism as a form of one-sided lexicography]. *Lexikos* 1: 1–19.

———. 1994. "Is die wêreld manlik, tensy anders vermeld?" [Is the world male, unless otherwise specified?]. *South African Journal of Linguistics* 12, no. 2: 43–49.

Biber, Douglas, Ulla Connor, and Thomas A. Upton. 2007. "Discourse Analysis and Corpus Linguistics." In *Discourse on the Move. Using Corpus Analysis to Describe Discourse Structure*, edited by Douglas Biber, Ulla Connor, and Thomas A. Upton, 1–20. Amsterdam: John Benjamins.

Biber, Douglas, Susan Conrad, and Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Blommaert, Jan, and Jef Verschuren. 1998. "The Role of Language in European Nationalist Ideologies. In *Language Ideologies: Practice and Theory*, edited by Bambi B. Schieffelin, Kathryn A. Woolard, and Paul V. Kroskrity, 189–210. Oxford: Oxford University Press.

Bloomfield, Leonard. 1914. *Introduction to the Study of Language*. New York: Henry Holt.

———. 1933. *Language*. New York: Henry Holt.

Booysen, C. Murray, and Martha Jansen. 1922. *Praktiese grammatika van die Afrikaanse taal: spesiaal vir gebruik in skole tot en met matriek* [Practical grammar of the Afrikaans language: Specially for use in schools until matric]. 3rd impr. Pietermaritzburg: Nasionale Pers.

Börjars, Kersti. 2003. "Morphological Status and (De)grammaticalisation: The Swedish Possessive." *Nordic Journal of Linguistics* 26, no. 2: 133–63.

Bosman, Nerina, and Anna N. Otto. 2012. "Die hantering van neweskikkers en onderskikkers in Afrikaanse woordeboeke" [The handling of coordinators and subordinators in Afrikaans dictionaries]. *Lexikos* 22: 69–85.

———. 2015. "'Moenie my 'jy' en 'jou' nie'—Die gebruik van *u* in die 21ste eeu. Resultate van 'n loodsondersoek" ["Moenie my 'jy' en 'jou' nie"—The use of *u* in the 21st century. Results from a pilot study]. *Litnet Akademies* 12, no. 3: 359–93.

Boshoff, Stephanus P. E. 1956. *Radiopraatjies oor Afrikaans* [Radio talks about Afrikaans]. Johannesburg: Voortrekkerpers.

———. 1964. "Prof S.P.E. Boshoff." In *Kruim en Kors* [Crumb and crust], edited by Stephanus P. E. Boshoff, Benedictus Kok, and Frederick J. Snijman, 13–115. Johannesburg: Voortrekkerpers.

Botha, Johannes P., and Johannes M. H. van Aardt. 1987. *Afrikaans vir die praktyk* [Afrikaanse for practical use]. 3rd ed. Johannesburg: Lexicon.

Botha, Marthinus C., and Johannes F. Burger. 1926. *Maskew Miller se Afrikaanse grammatika* [Maskew Miller's Afrikaans grammar]. 5th impr. Cape Town: Maskew Miller.

———. 1927. *Maskew Miller's Grammar of Afrikaans.* 5th ed. Cape Town: Maskew Miller.

Bouman, Arie C., and Eduard C. Pienaar. 1924. *Afrikaanse spraakkuns* [Afrikaans speech art]. Stellenbosch: Pro Ecclesia.

Bresnan, Joan, Anette Rosenbach, Benedikt Szmrecsanyi, Sali A. Tagliamonte, and Simon Todd. 2017. "Syntactic Alternations Data: Datives and Genitives in Four Varieties of English. Stanford Digital Repository. Available at http://purl.stanford.edu/qj187zs3852

Bybee, Joan. 2010a. "Diachronic Linguistics." In *The Oxford Handbook of Cognitive Linguistics*, edited by Dirk Geeraerts and Hubert Cuyckens, 945–87. Oxford: Oxford University Press.

———. 2010b. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, Joan, and Paul J. Hopper. 2001. "Introduction to Frequency and the Emergence of Linguistic Structure." In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 1–24. Amsterdam: John Benjamins.

Carstens, Adelia. 1998. "Die Afrikaanse medewerkwoorde, familieverwantskappe en grammatikalisering" [The Afrikaans auxiliaries, family resemblences and grammaticalization]. *South African Journal of Linguistics* 16, no. 3: 73–81.

Carstens, Adelia, and Heinrich Grebe, eds. 2001. *Taallandskap: Huldigingsbundel vir Christo van Rensburg* [Linguistic landscape: Festschrift for Christo van Rensburg]. Pretoria: Van Schaik.

Carstens, Wannie. 1989. *Norme vir Afrikaans: Enkele riglyne by die gebruik van Afrikaans* [Norms for Afrikaans: A few guidelines for the use of Afrikaans]. Pretoria: Van Schaik.

———. 1991. *Norme vir Afrikaans: Enkele riglyne by die gebruik van Afrikaans* [Norms for Afrikaans: A few guidelines for the use of Afrikaans]. 2nd ed. Pretoria: Van Schaik.

———. 1994. *Norme vir Afrikaans: Enkele riglyne by die gebruik van Afrikaans.* [Norms for Afrikaans: A few guidelines for the use of Afrikaans.] 3rd ed. Pretoria: Van Schaik.

———. 1997. *Afrikaanse Tekslinguistiek. 'n Inleiding* [Afrikaans text linguistics. An introduction]. Pretoria: Van Schaik.

———. 2003. *Norme vir Afrikaans: Enkele riglyne by die gebruik van Afrikaans* [Norms for Afrikaans: A few guidelines for the use of Afrikaans]. 4th ed. Pretoria: Van Schaik.

———. 2011. *Norme vir Afrikaans: Enkele riglyne by die gebruik van Afrikaans* [Norms for Afrikaans: A few guidelines for the use of Afrikaans]. 5th ed. Pretoria: Van Schaik.

———. 2018. *Norme vir Afrikaans: Moderne Standaardafrikaans* [Norms for Afrikaans: Modern Standard Afrikaans]. 6th ed. Pretoria: Van Schaik.

Coetzee, Anna E. 1990. "r-Weglating: 'n Tipe leksikale verandering" [r-Deletion: A type of lexical change]. *Tydskrif vir Geesteswetenskappe* 30, no. 2: 81–99.

———. 2001. "Van morfologie tot sintaksis: Taalvorme van die Bolandse wingerdwerkers" [From morphology to syntax: Linguistic forms of the Boland vineyard workers]. In *Taallandskap: Huldigingsbundel vir Christo van Rensburg* [Linguistic landscape: Festschrift for Christo van Rensburg], edited by Adelia Carstens and Heinrich Grebe, 21–29. Pretoria: Van Schaik.

Collins, Peter. 2013. "Grammatical Variation in English Worldwide: The Role of Colloquialization." *Linguistics and the Human Sciences* 8, no. 3: 289–306.

Collins, Peter, and Xinyue Yao. 2013. "Colloquial Features in World Englishes." *International Journal of Corpus Linguistics* 18, no. 4: 479–505.

Combrink, Johan G. H. 1995. *Hoe om paragrawe te skryf* [How to write paragraphs]. Cape Town: Tafelberg.

Combrink, Johan G. H., and Johan Spies. 1986. *SARA: Sakboek van Regte Afrikaans* [Pocketbook of correct Afrikaans.] Cape Town: Tafelberg.

Conrad, Susan. 2010. "What Can a Corpus Tell Us about Grammar?" In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 227–40. London: Routledge.

Conradie, C. Jac. 1986. *Taalgeskiedenis* [Language history]. Pretoria: Academica.

———. 1992. "Tempus in Afrikaanse Bybelvertalings" [Tense in Afrikaans Bible translations]. *South African Journal of Linguistics* 10, no. 2: 60–67.

———. 1998. "Tempusgebruik in Afrikaanse narratiewe" [Tense usage in Afrikaans narratives]. *South African Journal of Linguistics* 16, no. 2: 37–43.

———. 1999. "Preterite Loss in Early Afrikaans." *Folia Linguistica* 33, no. 1: 19–37.

———. 2001. "Hond se gedagte en nog 'n paar ingewings oor *se*" [Dog's thoughts and another few ideas about *se*]. In *Taallandskap: Huldigingsbundel vir Christo van Rensburg* [Linguistic landscape: Festschrift for Christo van Rensburg], edited by Adelia Carstens and Heinrich Grebe, 30–40. Pretoria: Van Schaik.

———. 2006. "Defleksie en grammatikalisasie: Die Afrikaanse hulpwerkwoord *het*" [Deflection and gramaticalization: The Afrikaans auxiliary verb *het* 'have']. *Journal for Language Teaching* 40, no. 1: 86–98.

———. 2010. "Het Nederlands en de standaardisering van het Afrikaanse werkwoordssysteem" [Dutch and the standardization of the Afrikaans verbal system]. In *Standaardtalen in beweging* [Standard languages in motion], edited by Marijke van der Wal and Eep Francken, 67–84. Amsterdam: Stichting Neerlandistiek VU.

Coussé, Evie, and Johan van der Auwera. 2012. "Human Impersonal Pronouns in Swedish and Dutch: A Contrastive Study of *Man* and *Men*." *Languages in Contrast* 12, no. 2: 121–138.

Croft, William. 1995. "Autonomy and Functionalist Linguistics." *Language* 71, no. 3: 490–532.

———. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow: Pearson Education.

———. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

———. 2010. "Construction Grammar." In *Oxford Handbook of Cognitive Linguistics*, edited by Dirk Geeraerts and Hubert Cuyckens, 463–508. Oxford: Oxford University Press.

Crowley, Tony. 1990. "That Obscure Object of Desire: A Science of Language." In *Ideologies of Language*, edited by John E. Joseph and Talbot J. Taylor, 27–50. London: Routledge.

Curzan, Anne. 2014. *Fixing English: Prescriptivism and Language History*. Cambridge: Cambridge University Press.

Davids, Achmat. 1987a. "Arabic-Afrikaans. A View of the Written Afrikaans of the Cape Muslims during the Nineteenth and Early Twentieth Centuries." *South African Journal of Linguistics* 5, no. 1: 24–40.

———. 1987b. "The Role of Afrikaans in the History of the Cape Muslim Community." In *Afrikaans en die taalpolitiek* [Afrikaans and language politics], edited by Hans du Plessis and Theo du Plessis, 37–60. Pretoria: HAUM.

———. 1990. "Words the Cape Slaves Made: A Socio-Historical Study." *South African Journal of Linguistics* 8, no. 1: 1–24.

———. 2011. *The Afrikaans of the Cape Muslims from 1815 to 1915*. Pretoria: Protea Book House.

Den Besten, Hans. 2012a. "The Complex Ancestry of the Afrikaans Associative Constructions." In *Roots of Afrikaans: Selected Writings of Hans den Besten*, edited by Ton van der Wouden, 25–33. Amsterdam: John Benjamins.

———. 2012b. "Demonstratives in Afrikaans and Cape Dutch Pidgin." In *Roots of Afrikaans: Selected Writings of Hans den Besten*, edited by Ton van der Wouden, 61–77. Amsterdam: John Benjamins.

———. 2012c. "From Khoekhoe Foreigner Talk via Hottentot Dutch to Afrikaans: The Creation of a Novel Grammar." In *Roots of Afrikaans: Selected Writings of Hans den Besten*, edited by Ton van der Wouden, 257–87. Amsterdam: John Benjamins.

———. 2012d. "The Origins of the Afrikaans Pre-Nominal Possessive System(s)." In *Roots of Afrikaans: Selected Writings of Hans den Besten*, edited by Ton van der Wouden, 7–23. Amsterdam: John Benjamins.

De Hoop, Helen, and Sammie Tarenskeen. 2015. "It's All about *You* in Dutch." *Journal of Pragmatics* 88: 163–75.

De Saussure, Ferdinand. [1916] 1966. *Kursus in algemene taalkunde* [Course in general linguistics.] Translated by Alewyn Lee. Pretoria: Van Schaik.

De Stadler, Leon G. 1989. *Afrikaanse semantiek* [Afrikaans semantics]. Johannesburg: Southern Book Publishers.

———. 1992. "Die werkwoorde *kom* en *gaan* as deiktiese medewerkwoorde" [The verbs *kom* and *gaan* as deictic auxiliaries]. *South African Journal of Linguistics* 10, no. 2: 91–98.

Deumert, Ana. 2002. "Language Variation and Standardization at the Cape (1880–1922): A Contribution to Afrikaans Sociohistorical Linguistics." *Journal of Germanic Linguistics* 13, no. 4: 301–52.

———. 2004. *Language Standardization and Language Change: The Dynamics of Cape Dutch*. Amsterdam: John Benjamins.

Deutscher, Guy. 2005. *The Unfolding of Language: An Evolutionary Tour of Mankind's Greatest Invention.* New York: Henry Holt.

De Villiers, Meyer. 1971. *Die grammatika van tyd en modaliteit* [The grammar of tense and modality]. 2nd impr. Cape Town: Balkema.

———. 1983. *Afrikaanse grammatika vir volwassenes* [Afrikaans grammar for adults]. Goodwood: Nasou.

Donaldson, Bruce. 1986. "The Second Person Pronouns in Germanic Languages: A Comparative Diachronic Study with Special Reference to Dutch and Afrikaans." *South African Journal of Linguistics* 4, no. 3: 1–12.

———. 1988. *The influence of English on Afrikaans.* Pretoria: Serva.

———. 1995. "Language Contact and Linguistic Change: The Influence of English on Afrikaans." In *Language and Social History: Studies in South African Sociolinguistics*, edited by Rajend Mesthrie, 222–29. Cape Town: David Philip.

Du Toit, Petrus J. 1986. *Taalleer vir onderwyser en student* [Language learning for teacher and student]. 2nd ed. Pretoria: Academica.

Du Toit, Piet, and Wanda Smith-Müller. 2003. *Stylboek: Riglyne vir paslike skryf* [Stylebook: Guidelines for appropriate writing]. Pretoria: Van Schaik.

Du Toit, Stephanus J. 1897. *Fergelijkende taalkunde fan Afrikaans en Engels* [Comparative grammar of English and Cape Dutch]. Paarl: D.F. du Toit.

———. 1902. *Fergelijkende taalkunde fan Afrikaans en Engels* [Comparative grammar of English and Cape Dutch]. 2nd ed. Paarl: D.F. du Toit.

Earp, Brian D. 2012. "The Extinction of Masculine Generics." *Journal for Communication and Culture* 2, no. 1: 4–19.

Eddington, David. 2010. "A Comparison of Two Tools for Analyzing Linguistic Data: Logistic Regression and Decision Trees." *Italian Journal of Linguistics* 22, no. 2: 265–86.

Ehlers, Dineke 1990. "Praktiese Afrikaans: Die gebruik van die passiefkonstruksie" [Practical Afrikaans: The use of the passive construction]. *Die Staatsamptenaar* 70, no. 11: 17.

Ehret, Katharina, Christoph Wolk, and Benedikt Szmrecsanyi. 2014. "Quirky Quadratures: On Rhythm and Weight as Constraints on Genitive Variation in an Unconventional Data Set." *English Language and Linguistics* 18, no. 2: 263–303.

Everett, Caleb. 2011. "Gender, Pronouns and Though: The Ligature between Epicene Pronouns and a More Neutral Gender Perception." *Gender and Language* 5, no. 1: 133–52.

Evison, Jane. 2010. "What Are the Basics of Analysing a Corpus?" In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 122–35. London: Routledge.

Fenk-Oczlon, Gertraud. 2001. "Familiarity, Information Flow, and Linguistic Form." In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 431–48. Amsterdam: John Benjamins.

Fischer, Olga, Muriel Norde, and Harry Perridon. 2004. "Introduction: In Search of Grammaticalization." In *Up and Down the Cline—The Nature of Grammaticalization*, edited by Olga Fischer, Muriel Norde, and Harry Perridon, 1–16. Amsterdam: John Benjamins.

Fouché, Nadine, Ria van den Berg, and Jako Olivier. 2017. "Carstens se raamwerk van Afrikaanse konjunksiemerkers: 'n Akkurater en vollediger weergawe" [Carstens's framework of Afrikaans conjunction markers: A more accurate and complete version]. *Tydskrif vir Geesteswetenskappe* 57, no. 3: 814–36.

Gast, Volker, Lisa Deringer, Florian Haas, and Olga Rudolf. 2015. "Impersonal Uses of the Second Person Singular: A Pragmatic Analysis of Generalization and Empathy Effects." *Journal of Pragmatics* 88: 148–62.

Gastil, John. 1990. "Generic Pronouns and Sexist Language: The Oxymoronic Character of Masculine Generics." *Sex Roles* 23, no. 11–12: 629–43.

Genootskap van Regte Afrikaanders. 1876. *Eerste beginsels van die Afrikaanse taal* [First principles of the Afrikaans language]. Cape Colony.

Gisborne, Nikolas, and Willem B. Hollmann. 2014. "Theory and Data in Cognitive Linguistics." In *Theory and Data in Cognitive Linguistics*, edited by Nikolas Gisborne and Willem B. Hollmann, 1–14. Amsterdam: John Benjamins.

Grafmiller, Jason. 2014. "Variation in English Genitives across Modality and Genres." *English Language and Linguistics* 18, no. 3: 471–96.

Grebe, Heinrich. 2012. *Op die keper beskou: Oor die ontstaan van Afrikaans* [Upon close examination: On the origin of Afrikaans]. Pretoria: Van Schaik.

Gries, Stefan Th. 2014. "Frequencies, Probabilities, and Association Measures in Usage-/Examplar-Based Linguistics: Some Necessary Clarifications." In *Theory and Data in Cognitive Linguistics*, edited by Nikolas Gisborne and Willem B. Hollman, 15–48. Amsterdam: John Benjamins.

Grobler, Hilda. 1993. *Praktiese werkboek vir Afrikaans* [Practical workbook for Afrikaans]. Pretoria: Academica.

Haegeman, Liliane. 2013. "Two Prenominal Possessors in West Flemish." In *Morphosyntactic Categories and the Expression of Possession*, edited by Kersti Börjars, David Denison, and Alan Scott, 219–51. Amsterdam: John Benjamins.

Hallan, Naomi. 2001. "Paths to Prepositions? A Corpus-Based Study of the Acquisition of a Lexico-Grammatical Category." In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 91–120. Amsterdam: John Benjamins.

Hantson, André. 2001. "Casusaanduiding in het Afrikaans en andere (Indo-) Germaanse talen: Een vergelijkende schets" [Case marking in Afrikaans and other (Indo-) Germanic languages: A comparative sketch]. *Tydskrif vir Geesteswetenskappe* 41, no. 1: 1–20.

Hare, Mary L., Michael Ford, and William D. Marslen-Wilson. 2001. "Ambiguity and Frequency Effects in Regular Verb Inflection." In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 181–200. Amsterdam: John Benjamins.

Haspelmath, Martin. 2004. "On Directionality in Language Change with Particular Reference to Grammaticalization." In *Up and Down the Cline—The Nature of Grammaticalization*, edited by Olga Fischer, Muriel Norde, and Harry Perridon, 17–44. Amsterdam: John Benjamins.

Heine, Bernd. 2003. "Grammaticalization." In *The Handbook of Historical Linguistics*, edited by Brian D. Joseph and Richard D. Janda, 575–601. Malden, MA: Blackwell.

Heller, Benedikt, Benedikt Szmrecsanyi, and Jason Grafmiller. 2017. "Stability and Fluidity in Syntactic Variation World-Wide: The Genitive Alternation across Varieties of English." *Journal of English Linguistics* 45, no. 1: 3–27.

Heltoft, Lars. 2010. "Paradigmatic Structure and Reanalysis: From NPs to DPs in Scandinavian." *Acta Linguistica Hafniensia* 42, no. 1: 11–25.

Hendriks, Jennifer. 2012. "Re-Examining the 'Origins' of the Prenominal Periphrastic Possessive Construction *Jan z'n boek* in Dutch: An Empirical Approach." *Diachronica* 29, no. 1: 28–71.

Hilpert, Martin. 2008. *Germanic Future Constructions. A Usage-Based Approach to Language Change*. Amsterdam: John Benjamins.

Hinrichs, Lars, and Benedikt Szmrecsanyi. 2007. "Recent Changes in the Function and Frequency of Standard English Genitive Constructions: A Multivariate Analysis of Tagged Corpora." *English Language and Linguistics* 11, no. 3: 437–74.

Hoeksema, Jack. 2010. "Dutch *ENIG*: From Nonveridicality to Downward Entailment." *Natural Language and Linguistic Theory* 28: 837–59.

Hoekstra, Jarich. 2010. "On the Impersonal Pronoun *Men* in Modern West Frisian." *Journal of Comparative Germanic Linguistics* 13: 31–59.

Hunston, Susan. 2010. "How Can a Corpus Be Used to Explore Patterns?" In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 152–66. London: Routledge.

Jankowski, Bridget L. 2013. "A Variationist Approach to Cross-Register Language Variation and Change." PhD thesis, University of Toronto.

Jankowski, Bridget L., and Sali A. Tagliamonte. 2014. "On the Genitive's Trail: Data and Method from a Sociolinguistic Perspective." *English Language and Linguistics* 18, no. 2: 305–29.

Jeffery, Chris, and Bertus van Rooy. 2004. "Emphasizer *Now* in Colloquial South African English." *World Englishes* 23, no. 2: 269–80.

Joseph, Brian D. 2004. "Rescuing Traditional (Historical) Linguistics from Grammaticalization Theory." In *Up and Down the Cline—The Nature of Grammaticalization*, edited by Olga Fischer, Muriel Norde, and Harry Perridon, 45–71. Amsterdam: John Benjamins.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. "Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 229–54. Amsterdam: John Benjamins.

Kennedy, Graeme. 2002. "Variation in the Distribution of Modal Verbs in the British National Corpus." In *Using Corpora to Explore Linguistic Variation*, edited by Randi Reppen, Susan Fitzmaurice, and Douglas Biber, 73–90. Amsterdam: John Benjamins.

Kirsner, Robert S. 2014. *Qualitative-Quantitative Analyses of Dutch and Afrikaans Grammar and Lexicon*. Amsterdam: John Benjamins.

Kirsten, Johanita. 2013. "Totius en die standaardisering van vroeg-moderne Afrikaanse werkwoorde" [Totius and the standardization of early-modern Afrikaans verbs]. *Tydskrif vir Geesteswetenskappe* 53, no. 1: 60–75.

———. 2015. "The Use of *Was* in Afrikaans Passive Constructions: A Diachronic Corpus Study." *Southern African Linguistics and Applied Language Studies* 33, no. 2: 159–70.

———. 2016. "Veranderinge in adverbiale tydsverwysing in Afrikaans van 1911 tot 2010" [Changes in adverbial temporal reference in Afrikaans from 1911 to 2010]. *Tydskrif vir Geesteswetenskappe* 56, no. 1: 45–61.

———. 2017. "What Is in a Language: Essentialism in Macro-Sociolinguistic Research on Afrikaans." *International Journal of the Sociology of Language* 248: 159–95.

———. 2018. "'Dit gaan nog belangrik word': Veranderinge in grammatikale toekomsverwysing in Afrikaans" ["Dit gaan nog belangrik word": Changes in grammatical future reference in Afrikaans]. *Tydskrif vir Geesteswetenskappe* 58, no. 2: 277–92.

Koester, Almut. 2010. "Building Small Specialised Corpora." In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 66–79. London: Routledge.

Kotzé, Ernst. 2011. "Demokratisering en standaardisering—Teenstrydige doelwitte vir inheemse tale?" [Democratization and standardization—Conflicting objectives for indigenous languages?]. *Tydskrif vir Geesteswetenskappe* 51, no. 4: 746–57.

———. 2014. "Afrikaans as besitting, en die vraagstuk van herstandaardisering" [Afrikaans as property, and the question of restandardization]. *Tydskrif vir Geesteswetenskappe* 54, no. 4: 635–55.

Kotzé, Ian, and Johanita Kirsten. 2016. "The Heritage of a Language: Discourses of Purism in Afrikaans Historical Linguistics." *Language Matters* 47, no. 3: 349–71.

Kriel, Mariana. 2006. "Fools, Philologists and Philosophers: Afrikaans and the Politics of Cultural Nationalism." *Politikon: South African Journal of Political Studies* 33, no. 1: 45–70.

———. 2013. "Loose Continuity: The Post-Apartheid Afrikaans Language Movement in Historical Perspective." PhD thesis, London School of Economics and Political Science.

Kroon, Caroline. 1997. "Discourse Markers, Discourse Structure and Functional Grammar." In *Discourse and Pragmatics in Functional Grammar*, edited by John H. Connolly, Roel M. Vismans, Christopher S. Butler, and Richard A. Gatward, 17–32. Berlin: Mouton de Gruyter.

Kruger, Haidee, and Bertus van Rooy. 2016. "Syntactic and Pragmatic Transfer Effects in Reported-Speech Constructions in Three Contact Varieties of English Influenced by Afrikaans." *Language Sciences* 56: 118–31.

Kuhn, Thomas. 1996. *The Structure of Scientific Revolutions*. 3rd ed. Chicago: University of Chicago Press.

Kunz, Kerstin, and Ekaterina Lapshinova-Koltunski. 2014. "Cohesive Conjunctions in English and German: Systemic Contrasts and Textual Differences." In *Recent Advances in Corpus Linguistics: Developing and Exploiting Corpora*, edited by Lieven Vandelanotte, Kristin Davidse, Caroline Gentens, and Ditte Kimps, 229–62. Amsterdam: Rodopi.

Labov, William. 2010. *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors.* Chichester: Wiley-Blackwell.

Laitinen, Lea. 2004. "Grammaticalization and Standardization." In *Up and Down the Cline—The Nature of Grammaticalization*, edited by Olga Fischer, Muriel Norde, and Harry Perridon, 247–62. Amsterdam: John Benjamins.

Lakoff, George, and Mark Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Langacker, Ronald W. 2010. "Cognitive Grammar." In *Oxford Handbook of Cognitive Linguistics*, edited by Dirk Geeraerts and Hubert Cuyckens, 421–62. Oxford: Oxford University Press.

Langer, Nils, and Agnete Nesse. 2012. "Linguistic Purism." In *The Handbook of Historical Sociolinguistics*, edited by Juan Manual Hernández-Campoy and Juan Camilo Conde-Silvestre, 607–25. Chichester: Wiley-Blackwell.

Lass, Roger. 1990. "How to Do Things with Junk: Exaptation in Language Evolution." *Journal of Linguistics* 26, no. 1: 79–102.

———. 1997. *Historical Linguistics and Language Change.* Cambridge: Cambridge University Press.

Le Cordeur, Michael. 2011. "Die variëteite van Afrikaans as draers van identiteit: 'n sosiokulturele perspektief" [The varieties of Afrikaans as carriers of identity: A socio-cultural perspective]. *Tydskrif vir Geesteswetenskappe* 51, no. 4: 758–77.

———. 2015. "Die kwessie van Kaaps: Afrikaansonderrig op skool benodig 'n meer inklusiewe benadering" [The issue of Kaaps: Afrikaans teaching at school needs a more inclusive approach]. *Tydskrif vir Geesteswetenskappe* 55, no. 4: 712–28.

Leech, Geoffrey, Marianne Hundt, Christian Mair, and Nicholas H. Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.

Le Roux, Jac J. 1939. *Praatjies oor ons taal* [Talks about our language]. Cape Town: Nasionale Pers.

Le Roux, Thomas H. 1946. *Afrikaanse taalstudies* [Afrikaans language studies]. 4th impr. Pretoria: Van Schaik.

Le Roux, Thomas H., Daniël F. Malherbe, and Johannes J. Smith. 1917. *Afrikaanse Woordelijs en Spelreëls* [Afrikaans word list and spelling rules]. Bloemfontein: Het Volksblad-Drukkerij.

MacWhinney, Brian. 2001. "Emergentist Approaches to Language." In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 449–70. Amsterdam: John Benjamins.

Mair, Christian. 2006. *Twentieth-Century English: History, Variation and Standardization*. Cambridge: Cambridge University Press.

———. 2011. "Grammaticalization and Corpus Linguistics." In *The Oxford Handbook of Grammaticalization*, edited by Bernd Heine and Heiko Narrog. Oxford: Oxford Handbooks Online.

Malherbe, Daniël F. 1917. *Afrikaanse taalboek: Praktiese wegwijser bij die vernaamste moeilikhede in verband met die Afrikaanse grammatika* [Afrikaans language book: Practical guidance for the most important difficulties with the Afrikaans grammar]. Bloemfontein: Nationale Pers.

McCarthy, Michael, and Anne O'Keeffe. 2010. "Historical Perspective: What Are Corpora and How Have They Evolved?" In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 3–13. London: Routledge.

McMahon, April M. S. 1994. *Understanding Language Change*. Cambridge: Cambridge University Press.

Messerschmidt, Johanna J. E., and Hans J. Messerschmidt. 2006. "Die kategorisering van onderskikkers: Enkele kenmerke van 'dat'-samestellings" [The categorization of subordinators: Some characteristics of "dat" compounds]. *Southern African Linguistics and Applied Language Studies* 24, no. 1: 7–22.

———. 2011. "Voorwaardelike konstruksies met *indien* soos gebruik in wetenskaplike tydskrifte (Deel 1: Basiese konstruksies)" [Conditional constructions with *indien* as used in scientific journals (Part 1: Basic constructions)]. *Tydskrif vir Geesteswetenskappe* 51, no. 2: 124–41.

———. 2012. "Voorwaardelike konstruksies met *indien* soos gebruik in wetenskaplike tydskrifte (Deel 2: Gemerkte konstruksies)" [Conditional constructions with *indien* as used in scientific journals (Part 2: Marked constructions)]. *Tydskrif vir Geesteswetenskappe* 52, no. 2: 271–89.

Mesthrie, Rajend, ed. 1995. *Language and Social History: Studies in South African Sociolinguistics*. Cape Town: David Philip.

Milroy, James. 2001. "Language Ideologies and the Consequences of Standardization." *Journal of Sociolinguistics* 5, no. 4: 530–55.

———. 2012. "Sociolinguistics and Ideologies in Language History." In *The Handbook of Historical Sociolinguistics*, edited by Juan Manual Hernández-Campoy and Juan Camilo Conde-Silvestre, 571–84. Chichester: Wiley-Blackwell.

Mufwene, Salikoko S. 2003. "Genetic Linguistics and Genetic Creolistics: A Response to Sarah G. Thomason's 'Creoles and Genetic Relationships.'" *Journal of Pidgin and Creole Languages* 18, no. 2: 273–88.

———. 2008. *Language Evolution: Contact, Competition and Change*. London: Continuum.

Mühlhäusler, Peter, and Rom Harré. 1990. *Pronouns & People: The Linguistic Construction of Social and Personal Identity*. Oxford: Basil Blackwell.

Müller, Dalene. 2003. *Skryf Afrikaans van A tot Z: Die essensiële gids vir taalgebruikers* [Write Afrikaans from A to Z: The essential guide for language users]. Cape Town: Pharos Dictionaries.

Müller, Dalene, and Sebastian Pistor. 2011. *Skryf Afrikaans van A tot Z: Die essensiële gids vir taalgebruikers* [Write Afrikaans from A to Z: The essential guide for language users]. 2nd ed. Cape Town: Pharos Dictionaries.

Nikiforidou, Kiki. 1991. "The Meanings of the Genitive: A Case Study in Semantic Structure and Semantic Change." *Cognitive Linguistics* 2, no. 2: 149–205.

O'Connor, Catherine, Joan Maling, and Barbora Skarabela. 2013. "Nominal Categories and the Expression of Possession: A Cross-Linguistic Study of Probabilistic Tendencies and Categorical Constraints." In *Morphosyntactic Categories and the Expression of Possession*, edited by Kersti Börjars, David Denison, and Alan Scott, 89–121. Amsterdam: John Benjamins.

Odendaal, Gerda. 2014. "Moet Afrikaans geherstandaardiseer word?" [Should Afrikaans be restandardized?]. *Tydskrif vir Geesteswetenskappe* 54, no. 4: 656–74.

Patard, Adeline. 2014. "When Tense and Aspect Convey Modality: Reflections on the Modal Uses of Past Tenses in Romance and Germanic Languages." *Journal of Pragmatics* 71: 69–97.

Patten, Amanda. 2014. "The Historical Development of the *It*-Cleft: A Comparison of Two Different Approaches." In *Theory and Data in Cognitive Linguistics*, edited by Nikolas Gisborne and Willem B. Hollmann, 87–114. Amsterdam: John Benjamins.

Perridon, Harry. 2013. "The Emergence of the *S*-Genitive in Danish." *Language Sciences* 36: 134–46.

Phillips, Betty S. 2001. "Lexical Diffusion, Lexical Frequency, and Lexical Analysis. In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 123–36. Amsterdam: John Benjamins.

Pienaar, Eduard C., and Cornelis J. Langenhoven. 1932. *Hoe om te skrywe: Hulpboekie by die studie van die stelkuns* [How to write: Little help book for the study of writing]. Cape Town: Nasionale Pers.

Pierrehumbert, Janet. 2001. "Exemplar Dynamics: Word Frequency, Lenition and Contrast." In *Frequency and the Emergence of Linguistic Structure*, edited by Joan Bybee and Paul J. Hopper, 137–57. Amsterdam: John Benjamins.

Ponelis, Fritz A. 1979. *Afrikaanse sintaksis* [Afrikaans syntax]. Pretoria: Van Schaik.

———. 1989. "Sintaksis" [Syntax]. In *Inleiding tot die Afrikaanse taalkunde* [Introduction to Afrikaans linguistics]. 2nd ed., edited by Theunis J. R. Botha, 255–357. Pretoria: Academica.

———. 1992. "Standaardafrikaans in oorgang" [Standard Afrikaans in transition]. In *Afrikaans ná Apartheid* [Afrikaans after Apartheid], edited by Vic Webb, 71–89. Pretoria: Van Schaik.

———. 1993. *The Development of Afrikaans*. Frankfurt am Main: Peter Lang.

———. 1994. "Die ontstaan van Afrikaans" [The origin of Afrikaans]. *Tydskrif vir Geesteswetenskappe* 34, no. 4: 218–30.

Poplack, Shana, Lauren Zentz, and Nathalie Dion. 2012a. "Phrasefinal Prepositions in Quebec French: An Empirical Study of Contact, Codeswitching and Resistance to Convergence." *Bilingualism: Language and Cognition* 15, no. 2: 203–25.

———. 2012b. "What Counts as (Contact-Induced) Change." *Bilingualism: Language and Cognition* 15, no. 2: 247–54.

Prinsloo, Anton F. 2004. *Spreekwoorde en waar hulle vandaan kom* [Proverbs and where they come from]. Cape Town: Pharos Dictionaries.

———. 2009a. *Annerlike Afrikaans: Woordeboek van Afrikaanse Kontreitaal* [Different Afrikaans: Dictionary of Afrikaans Regional Language]. Pretoria: Protea Boekhuis.

———. 2009b. *Spreekwoorde en waar hulle vandaan kom* [Proverbs and where they come from]. 2nd ed. Cape Town: Pharos Dictionaries.

Prinsloo, Anton F., and Francois F. Odendal. 1995. *Afrikaans op sy beste: Hulp met moderne taalkwessies* [Afrikaans at its best: Help with modern language issues]. Pretoria: Van Schaik.

Raidt, Edith H. 1989. "Ontwikkeling van vroeë Afrikaans" [Development of early Afrikaans]. In *Inleiding tot die Afrikaanse taalkunde* [Introduction to Afrikaans linguistics]. 2nd ed., edited by Theunis J. R. Botha, 96–126. Pretoria: Academica.

———. 1991. *Afrikaans en sy Europese verlede* [Afrikaans and its European past]. 2nd ed. Cape Town: Nasou.

Rayson, P. 2018. *Log-likelihood and Effect Size Calculator*. http://ucrel.lancs.ac.uk/llwizard.html. Accessed May 22, 2018.

Redeker, Gisela, and Helmut Gruber. 2014. "Introduction: The Pragmatics of Discourse Coherence." In *The Pragmatics of Discourse Coherence: Theories and Applications*, edited by Helmut Gruber and Gisela Redeker, 1–20. Amsterdam: John Benjamins.

Reppen, Randi. 2010. "Building a Corpus: What Are the Key Considerations?" In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 31–37. London: Routledge.

Reppen, Randi, Susan Fitzmaurice, and Douglas Biber. 2002. *Using Corpora to Explore Linguistic Variation*. Amsterdam: John Benjamins.

Roberge, Paul T. 1990. "The Ideological Profile of Afrikaans Historical Linguistics." In *Ideologies of Language*, edited by John E. Joseph and Talbot J. Taylor, 131–49. London: Routledge.

———. 1994. "On Detecting a Prior Linguistic Continuum in Cape Dutch. In *Nuwe Perspektiewe op die Geskiedenis van Afrikaans: Opgedra aan Edith H. Raidt* [New

perspectives on the history of Afrikaans: Dedicated to Edith H. Raidt], edited by Gerrit Olivier and Anna Coetzee, 153–65. Halfweghuis: Southern Boekuitgewers.

———. 1996. "Multilevel Syncretism and the Evolution of Afrikaans Periphrastic Possessives with *se*." *Stellenbosch Papers in Linguistics* 30: 123–53.

———. 2001. "Diachronic Notes on the Afrikaans Demonstrative Pronouns." In *Taal-landskap: Huldigingsbundel vir Christo van Rensburg* [Linguistic landscape: Festschrift for Christo van Rensburg], edited by Adelia Carstens and Heinrich Grebe, 124–36. Pretoria: Van Schaik.

———. 2002. "Afrikaans: Considering Origins." In *Language in South Africa*, edited by Rajend Mesthrie, 79–103. Cambridge: Cambridge University Press.

———. 2003. "Afrikaans." In *Germanic Standardizations. Past to Present*, edited by Ana Deumert and Wim Vandenbussche, 15–40. Amsterdam: John Benjamins.

———. 2012. "Afrikaans: Might It Be a Little More 'South Africa'?" In *Roots of Afrikaans: Selected Writings of Hans den Besten*, edited by Ton van der Wouden, 389–97. Amsterdam: John Benjamins.

Rosenbach, Anette. 2002. *Genitive Variation in English*. Berlin: Mouton de Gruyter.

———. 2005. "Animacy versus Weight as Determinants of Grammatical Variation in English." *Language* 81, no. 3: 613–44.

———. 2008. "Animacy and Grammatical Variation—Findings from English Genitive Variation." *Lingua* 118: 151–71.

———. 2010. "How Synchronic Gradience Makes Sense in the Light of Language Change (and *Vice Versa*). In *Gradience, Gradualness and Grammaticalization*, edited by Graeme Trousdale and Elizabeth Closs Traugott, 149–79. Amsterdam: John Benjamins.

———. 2014. "English Genitive Variation—The State of the Art." *English Language and Linguistics* 18, no. 2: 215–62.

Rossouw, Ronel, and Bertus van Rooy. 2012. "Diachronic Changes in Modality in South African English." *English World-Wide: A Journal of Varieties of English* 33, no. 1: 1–26.

Roux, Susanna P. 2001. "'n Kritiese bespreking van taalverandering met verwysing na Afrikaans en sy variëteite" [A critical discussion on language change with reference to Afrikaans and its varieties]. MA diss., University of the Free State.

RStudio Team. 2016. *RStudio: Integrated Development for R*. (Ver. 1.0.143) [computer software]. Boston: RStudio, Inc. http://www.rstudio.com.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Co.

Scholtz, Johannes du P. 1950. *Oor die herkoms van Afrikaans* [On the origin of Afrikaans]. Cape Town.

———. 1963. "Aanvulling by: 'Die geskiedenis van die persoonlike en besitlike voornaamwoorde in Afrikaans.'" [Addition to: "The history of the personal and possessive pronouns in Afrikaans"]. *Tydskrif vir Geesteswetenskappe* 3, no. 2: 67–74.

———. 1980. *Wording en ontwikkeling van Afrikaans* [Origin and development of Afrikaans]. Cape Town: Tafelberg.

Scholtz, Johannes J. J. 1990. "'Is' en 'was,' 'wat' en 'wie.'" ['Is' and 'was,' 'what' and 'who']. *Beeld* 13 (August 14).

Schøsler, Lene. 2010. "A Paradigmatic Approach to Language and Language Change." In *Gradience, Gradualness and Grammaticalization*, edited by Graeme Trousdale and Elizabeth Closs Traugott, 203–20. Amsterdam: John Benjamins.

Schwerdtfeger, Anna-Mart. 1989. "Seksisme en stereotipes: Kriteria by die skryf en voorskryf van skoolhandboeke" [Sexism and stereotypes: Criteria for writing and prescribing school textbooks]. *Journal for Language Teaching* 23, no. 4: 20–35.

Scott, Alan K. 2011. "The Position of the Genitive in Present-Day Dutch." *Word Structure* 4, no. 1: 104–35.

Scott, Mike. 2010. "What Can Corpus Software Do?" In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 136–51. London: Routledge.

Smith, Mark. 2004. "Light and Heavy Reflexives." *Linguistics* 42, no. 3: 573–615.

Speelman, Dirk. 2014. "Logistic Regression: A Confirmatory Technique for Comparisons in Corpus Linguistics." In *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*, edited by Dylan Glynn and Justyna A. Robinson, 487–533. Amsterdam: John Benjamins.

Spitulnik, Debra. 1998. "Mediating Unity and Diversity: The Production of Language Ideologies in Zambian Broadcasting." In *Language Ideologies: Practice and Theory*, edited by Bambi B. Schieffelin, Kathryn A. Woolard, and Paul V. Kroskrity, 163–88. Oxford: Oxford University Press.

Stede, Manfred. 2014. "Resolving Connective Ambiguity: A Prerequisite for Discourse Parsing." In *The Pragmatics of Discourse Coherence: Theories and Applications*, edited by Helmut Gruber and Gisela Redeker, 121–40. Amsterdam: John Benjamins.

Stoops, Yvette H. L. 1972. "Die taalgebruik in 'Die Huisgenoot': Bydrae tot die ondersoek na die ontwikkeling van geskrewe Afrikaans in die laaste 50 jaar" [The language use in "Die Huisgenoot": Contribution to the investigation of the development of written Afrikaans in the last 50 years]. MA diss., University of Port Elizabeth.

Swart, Corlietha. 1995. "'n Leksikaal-semantiese ondersoek na verandering in die Afrikaanse woordeskat soos weerspieël in die hedendaagse nuusmedia" [A lexical-semantic investigation into change in the Afrikaans lexicon as reflected in current news media]. MA diss., University of the Orange Free State.

Szmrecsanyi, Benedikt. 2013. "The Great Regression: Genitive Variability in Late Modern English News Texts." In *Morphosyntactic Categories and the Expression of Possession*, edited by Kersti Börjars, David Denison, and Alan Scott, 59–88. Amsterdam: John Benjamins.

———. 2016. "About Text Frequencies in Historical Linguistics: Disentangling Environmental and Grammatical Change." *Corpus Linguistics and Linguistic Theory* 12, no. 1: 153–71.

Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2017. *Afrikaanse woordelys en spelreëls* [Afrikaans word list and spelling rules]. 11th ed. Cape Town: Pharos Dictionaries.

Thompson, Sandra A., and Paul J. Hopper. 2001. "Transitivity, Clause Structure, and Argument Structure: Evidence from Conversation." In *Frequency and the Emer-*

gence of Linguistic Structure, edited by Joan Bybee and Paul J. Hopper, 27–60. Amsterdam: John Benjamins.

Tognini Bonelli, Elena. 2010. "Theoretical Overview of the Evolution of Corpus Linguistics." In *The Routledge Handbook of Corpus Linguistics*, edited by Michael McCarthy and Anne O'Keeffe, 14–27. London: Routledge.

Tomasello, Michael. 2014. *A Natural History of Human Thinking*. Cambridge, MA: Harvard University Press.

Traugott, Elizabeth Closs, and Graeme Trousdale. 2010. "Gradience, Gradualness and Grammaticalization: How Do They Intersect?" In *Gradience, Gradualness and Grammaticalization*, edited by Graeme Trousdale and Elizabeth Closs Traugott, 19–44. Amsterdam: John Benjamins.

Trousdale, Graeme. 2014. "Theory and Data in Diachronic Construction Grammar: The Case of the *What With* Construction." In *Theory and Data in Cognitive Linguistics*, edited by Nikolas Gisborne and Willem B. Hollmann, 115–40. Amsterdam: John Benjamins.

Trousdale, Graeme, and Elizabeth Closs Traugott. 2010. "Preface." In *Gradience, Gradualness and Grammaticalization*, edited by Graeme Trousdale and Elizabeth Closs Traugott, 1–18. Amsterdam: John Benjamins.

Trudgill, Peter. 2002. *Sociolinguistic Variation and Change*. Washington, DC: Georgetown University Press.

Uys, Mariëtte D. 1983. "Die vernederlandsing van Afrikaans" [The Dutchification of Afrikaans]. DLitt thesis, University of Pretoria.

Van der Auwera, Johan, and Vladimir A. Plungian. 1998. "Modality's Semantic Map." *Linguistic Typology* 2: 79–124.

Van der Merwe, Henderik J. J. M. 1951. *An Introduction to Afrikaans*. Cape Town: Balkema.

Van der Merwe, Henderik J. J. M., and Fritz A. Ponelis. 1982. *Die korrekte woord: Afrikaanse taalkwessies* [The correct word: Afrikaans language issues]. 6th ed. Pretoria: Van Schaik.

———. 1991. *Die korrekte woord: Afrikaanse taalkwessies* [The correct word: Afrikaans language issues]. 7th ed. Pretoria: Van Schaik.

Van der Merwe, Truida. 1996. "*Nou*, *toe* end *dan* as temporele leksikale elemente in Afrikaans" [*Nou*, *toe*, and *dan* as temporal lexical items in Afrikaans]. *Literator* 17, no. 2: 91–104.

———. 1997. "Pronominale verwysing na inanimata in Afrikaans" [Pronominal reference to inanimates in Afrikaans]. *South African Journal of Linguistics* 15, no. 1: 27–34.

———. 2003. "'n Etimologiese perspektief op betekenisverandering in die Afrikaanse leksikon" [An etymological perspective on semantic change in the Afrikaans lexicon]. *Literator* 24, no. 1: 103–23.

Van der Vliet, Nynke, and Gisela Redeker. 2014. "Explicit and Implicit Coherence Relations in Dutch Texts." In *The Pragmatics of Discourse Coherence: Theories and Applications*, edited by Helmut Gruber and Gisela Redeker, 23–52. Amsterdam: John Benjamins.

Van Gelderen, Elly. 2010. "Features in Reanalysis and Grammaticalization." In *Gradience, Gradualness and Grammaticalization*, edited by Graeme Trousdale and Elizabeth Closs Traugott, 129–47. Amsterdam: John Benjamins.

Van Helden, S., and Anna E. Coetzee. 2006. "Die sosio-historiese en sosio-linguistiese aspekte van die leksikale veranderings van die 1936- en 2001-Psalmomdigtings" [The sociohistorical and sociolinguistic aspects of the lexical changes in the 1936 and 2001 Psalm versifications]. *In die Skriflig* 40, no. 2: 199–220.

Van Olmen, Daniël, and Adri Breed. 2018. "Human Impersonal Pronouns in Afrikaans: A Double Questionnaire-Based Study." *Language Sciences* 69: 1–29.

Van Olmen, Daniël, and Tanja Mortelmans. 2009. "Movement Futures in English and Dutch: A Contrastive Analysis of *Be Going To* and *Gaan*." In *Studies on English Modality*, edited by Anastasios Tsangalidēs and Roberta Facchinetti, 357–86. Bern: Peter Lang.

Van Rensburg, Christo. 1998. "Oor skakelwerkwoorde" [On linking verbs]. In *Sintaksis op die voorgrond* [Syntax at the fore], edited by Rufus Gouws and Ilse Feinauer, 164–72. Pretoria: Van Schaik.

———. 2017. "Ná honderd jaar: Die Afrikaans van die eerste taalkommissie" [After hundred years: The Afrikaans of the first language commission]. *Tydskrif vir Geesteswetenskappe* 57, no. 2-1: 249–70.

Van Rensburg, Christo, Achmat Davids, Jeanette Ferreira, Tony Links, and Karel Prinsloo. 1997. *Afrikaans in Afrika* [Afrikaans in Africa]. Pretoria: Van Schaik.

Van Rooy, Bertus. 1996. "Is generiese manlike voornaamwoorde regtig generies?" [Are generic male pronouns really generic?]. *South African Journal of Linguistics* 14, no. 1: 33–41.

———. In press (2019). "The Influence of English on Afrikaans Reconsidered." In *English in Multilingual South Africa*, edited by Raymond Hickey. Cambridge: Cambridge University Press.

Van Rooy, Bertus, and Ia Esterhuizen. 2011. "Die gebruik van konjunksiemerkers by nagraadse studente en gepubliseerde navorsing oor Afrikaanse taal- en letterkunde" [The use of conjunctive markers by postgraduate students and published research on Afrikaans linguistics and literature]. *Journal for Language Teaching* 45, no. 1: 67–86.

Van Rooy, Bertus, and Haidee Kruger. 2015. "The Case for an Emergentist Approach." *Stellenbosch Papers in Linguistics Plus* 48: 41–67.

———. 2016. "Faktore wat die weglating van die Afrikaanse onderskikker *dat* bepaal" [Factors that determine the omission of the Afrikaans subordinator *dat*]. *Tydskrif vir Geesteswetenskappe* 56, no. 1: 102–16.

Van Rooy, Herrie. 1984. "'n Deelsgenitief in Afrikaans?" [A partitive genitive in Afrikaans?]. *South African Journal of Linguistics* 2, no. 4: 27–41.

Van Schalkwyk, Dirk J., and Gabriël Kroes. 1979. *Doeltreffende Afrikaans: 'n Handboek vir skriftelike taalgebruik* [Efficient Afrikaans: A textbook for written language use]. Goodwood: Nasou.

Van Schalkwyk, Helena, and Deidré A. Viviers. 1977. *Taalkommunikasie: Afrikaans* [Language communication: Afrikaans]. Cape Town: Maskew Miller.

Van Schoor, J. L. 1983 *Die grammatika van Standaard-Afrikaans* [The grammar of Standard Afrikaans]. Cape Town: Lex Patria.

Van Wyk, Tia. 2009. *Tempusgebruik in Afrikaanse narratiewe tekste* [Tense usage in Afrikaans narrative texts]. MA mini-diss., University of Johannesburg.

Vincent, Nigel, and Kersti Börjars. 2010. "Grammaticalization and Models of Language." In *Gradience, Gradualness and Grammaticalization*, edited by Graeme Trousdale and Elizabeth Closs Traugott, 279–99. Amsterdam: John Benjamins.

Vink, Marietjie. 1981. "*Van* vs. *se*." In *LVSA-Kongresreferate* [LVSA Conference papers], edited by A. J. L. Sinclair, 431–44. Belville: University of the Western Cape.

Virtuele Instituut vir Afrikaans [Virtual Institute for Afrikaans]. 2018. "Korpusportaal." [Corpus Portal]. http://viva-afrikaans.org.

Visconti, Jacqueline. 2004. "Conditionals and Subjectification: Implications for a Theory of Semantic Change." In *Up and Down the Cline—The Nature of Grammaticalization*, edited by Olga Fischer, Muriel Norde, and Harry Perridon, 169–92. Amsterdam: John Benjamins.

Wasserman, Ronel, and Bertus van Rooy. 2014. "The Development of Modals of Obligation and Necessity in White South African English through Contact with Afrikaans." *Journal of English Linguistics* 42, no. 1: 31–50.

Webb, Vic, and Mariana Kriel. 2000. "Afrikaans and Afrikaner Nationalism." *International Journal for the Sociology of Language* 144: 19–49.

Weerman, Fred, and Petra de Wit. 1999. "The Decline of the Genitive in Dutch." *Linguistics* 37, no. 6: 1155–92.

Weerman, Fred, Mike Olson, and Robert A. Cloutier. 2013. "Synchronic Variation and Loss of Case: Formal and Informal Language in a Dutch Corpus of 17th-Century Amsterdam Texts." *Diachronica* 30, no. 3: 353–81.

Winford, Donald. 2012. "Creole Languages." In *The Oxford Handbook of Tense and Aspect*, edited by Robert I. Binnick, 428–57. Oxford: Oxford University Press.

Woolard, Kathryn A. 1998. "Introduction: Language Ideology as a Field of Inquiry." In *Language Ideologies: Practice and Theory*, edited by Bambi B. Schieffelin, Kathryn A. Woolard, and Paul V. Kroskrity, 3–47. Oxford: Oxford University Press.

Wybenga, Daan M. 1993. "Gewete en waargenome wêreld: 'n Tekstuele perspektief op Afrikaanse tydsuitdrukking" [Known and observed world: A textual perspective on Afrikaans temporal expression]. *South African Journal of Linguistics* 24, no. 2: 15–30.

Index

Page references for tables and figures are italicized.

221

About the Author

Johanita Kirsten, senior lecturer in Afrikaans linguistics at the Vaal Triangle Campus of the North-West University (NWU) in South Africa, is interested in recent and ongoing changes in Afrikaans language use. She completed her MA and PhD in Afrikaans and Dutch (historical linguistics) at the NWU under supervision of Professor Bertus van Rooy and received the ABSA Bronze Medal for her MA dissertation. The research topics she has published on include changes in the verbal system of written Afrikaans during initial standardization of Afrikaans in the early twentieth century, changes regarding temporal reference in Afrikaans during the course of the twentieth and early twenty-first centuries, and how this links to theories of language change; another focus of her work is issues of language purity in the historical descriptions of Afrikaans and ideological issues in published linguistic research. She continues to research different recent and ongoing changes in Afrikaans usage as well as ideological issues in Afrikaans linguistics. She has been teaching undergraduate and graduate Afrikaans linguistics modules at the NWU since 2012 and supervises postgraduate studies in Afrikaans linguistics and general linguistics.