Grammatical change in South African Englishes

Bertus van Rooy¹,²

¹University of Amsterdam, The Netherlands
²North-West University, Vanderbijlpark, South Africa

Abstract
This article investigates whether grammatical change in South African English (SAFE) leads to convergence or divergence in the different varieties within SAFE. Contact between speakers of English and other languages potentially leads to divergence from varieties beyond SAFE, and to convergence within SAFE. Modal auxiliaries show convergence between White SAFE (WSAFE) and Afrikaans, but not between WSAFE and Black SAFE (BSAFE). No convergence is observed between the innovative use of the progressive in BSAFE and WSAFE, except for similar long-term frequency increases. The omission of the complementiser that shows similar long-term frequency changes between WSAFE, Afrikaans and BSAFE. The findings indicate that higher degrees of contact between speakers result in closer convergence, especially as far as constructional semantics are concerned, but convergent frequency changes seem to be possible without the same closeness of contact.

1 | INTRODUCTION

A number of grammatical changes can be observed in the use of English since it was transplanted to South Africa two centuries ago. These changes have been studied among its ancestral native (L1) speakers (Kruger & Van Rooy, 2019; Lass & Wright, 1986; Mesthrie, 2002; Wasserman, 2014), and more recently also among second-language (L2) speakers (Van Rooy & Piotrowska, 2015; Van Rooy & Wasserman, 2014). However, like the case with Australian English (Collins, 2014: 8), the phonological (Bekker, 2009, 2012; Lanham & MacDonald, 1979; Lass, 1995) and lexical properties (Silva, 1996) have received more attention over the years, with Pettman’s (1913) dictionary Africanderisms representing the first systematic documentation of the uniquely South African lexicon. Some changes can be attributed to external factors, such as language contact, while others are likely to represent ongoing internal changes. Among its ancestral speakers, ongoing change has resulted in the disappearance of subject-verb inversion in examples such as (1a), from 19th century South African fiction, where the subject is a pronoun (Kruger & Van Rooy, 2019), and to the almost complete
disappearance of the use of the modal shall outside the context of legal obligations (Wasserman, 2014). These changes cannot be attributed to language contact with Afrikaans, since Afrikaans has obligatory subject-verb inversion in the corresponding cases illustrated by example (1b) – the direct translation of (1a), and Afrikaans continues to use the cognate modal verb sal in contexts where English uses will in the place where shall was historically used, as shown by (2a) and its direct translation (2b).

(1a) ‘Hadn’t we better change the name, Bain,’ said I when he had gone, ‘and call this “Orthis Kloof?”’ (WSAfE narrative, 1871)

(1b) ‘Moes ons nie liewer die naam verander nie, Bain,’ sé ek toe hy weg is, ‘en hierdie plek “Orthiskloof” noem nie?’

(2a) Obviously, we will still retain our peak holiday visitors… (WSAfE reportage, contemporary (ICE-SA, W2C-012))

(2b) Uiteraard sal ons steeds ons hoogseisoenbesoekers behou…

These changes are well-known among Inner Circle varieties of English world-wide, and simply represent ongoing change that had been set in motion prior to the transplantation of English to South Africa, and continued after transplantation.

Other changes in the grammar of South African English can be attributed to language contact, such as the adoption of a number of Afrikaans innovations in informal usage, for example the associative plural and them or reduplication of adverbs (Mesthrie, 2017), as illustrated by example (3) and (4) respectively, taken from websites.

(3) I also wonder of their father and them see their twin status as being more “sellable” performance wise (https://www.tvsa.co.za/user/blogs/viewblogpost.aspx?blogpostid=51214)


In the case of L2 users, contact with indigenous languages is a very productive source of language change, as shown by the extended uses of the progressive, transferred from the persistitive in the Bantu languages (Van Rooy & Piotrowska, 2015), as shown by example (5). Even among L1 speakers, grammatical change due to contact with other South African languages have been identified as source for the innovative uses of now (Jeffery & Van Rooy, 2004).

(5) This happened on June 26. Cycling on the East London road via Mount Coke, then branching off to Berlin, we saw a village on top of a mountain. As the main road was not leading there it was necessary to look for a footpath up the mountain side. (BSAfE reportage, 1947)

Mesthrie (1999) identifies a further source of innovation in the use of unstressed do in Cape Flats English, which is the retention of early input. He notes that while the ancestral speakers of English lost the unstressed do, its use continued among L2 speakers in the Cape Flats areas, as illustrated by example (6) from his paper. There was further possible reinforcement by missionary teachers, themselves often L2 users of English, whose late 18th and early 19th century letters also contain similar instantiations.

(6) Who did throw that? (Mesthrie, 1999)

While the sources of grammatical change in South African English (henceforth SAFE) has been studied in some detail, we know less about the extent to which grammatical change in SAFE leads to divergence or convergence among the Englishes used in the country. It has become established practice over the years to approach the grammatical (and phonological) properties of South African English through the lens of ethnic varieties, labelled White South African English (WSAfE), Black South African English (BSAfE), Indian South African English (ISAfE), Afrikaans English (AfrE) and Coloured (or sometimes Cape Flats) English (CE/CFE). The justification for this classification is historical: the racial separation brought about by the policy of apartheid (and its informal precursors throughout the colonial era) for all but the final years of the 20th century resulted in the formation of ethnolects because individuals mainly communicated
within and seldom across their statutorily defined population groups. In conjunction with the spatial separation, particular contact languages or historical and cultural factors reinforced the development of unique linguistic features within particular ethnolects. However, as Mesthrie (2009) notes, the ethnic varieties are not self-evident fixed categories to which speakers adhere, and fluidity is to be expected as speakers come into contact with a larger variety of Englishes, going back even to the 1980s. Mesthrie (2010) and Mesthrie, Chevalier, and Dunne (2015) point to the convergence in phonological features of middle class South Africans across ethnic groups, who are increasingly in face-to-face contact in the desegregated schools and public spaces that have opened up since the early 1990s. Lexical borrowing between different South African languages (Branford & Claughton, 1995) has likewise given rise to convergent tendencies, with innovations originating in one ethnic contact situation spreading to other speaker groups beyond the innovating community (Van Rooy & Terblanche, 2010).

The question of convergence or divergence in grammatical change has not received the same amount of attention. Previous scholarship focuses mainly on documenting variety features, and explaining underlying factors that play a role in observed changes. This paper is intended to contribute to our understanding of grammatical convergence across the range of Englishes in South Africa, by examining a number of grammatical features where change has been ongoing for the past two centuries, in order to establish where convergence and divergence respectively result from these changes. In the process, the relative contribution of contact, input and endogenous change will be considered within the broader pattern of contact between speakers. It will be shown that contact and interaction between speakers facilitate convergence, while lack of contact results in divergence, even if for some features, especially insofar as the frequency of use of old and new variants is concerned, converging tendencies can be observed despite limited contact.

The next section of the paper considers perspectives on convergence within the broader framework of world Englishes theorising, followed by three sections that consider grammatical features in SAfE: modals, the progressive aspect, and the omission of the complementiser that from complement clauses. These features have been subject to previous research with historical corpus data, where the focus has been on factors contributing to language change rather than the overall development towards convergence or divergence, and therefore a reinterpretation of some previous data is offered. Previous research also usually focused on one subvariety of SAfE at a time, or at best at comparing two varieties. The focus in this paper expands to more comprehensive comparisons, while still not being able to attend to all possible comparisons, due to limitations on the available corpus data. Additional corpus analysis was also undertaken in the case of the modal must and that omission, over and above what has been done before.

The historical data used in previous research being reviewed, as well as new analyses reported in this paper, are mainly taken from two historical corpora of SAfE, representing WSAfE and BSAfE respectively. The most complete descriptions of these two corpora can be found in Wasserman (2014: 196–209) and Piotrowska (2014: 81–88). The historical WSAfE corpus, which also incorporates certain registers of ICE-SA (Jeffery, 2003), consists of letters, fiction, news and published non-fiction, and amounts to 659,770 words, covering the period from 1820 to 2000. The historical BSAfE corpus includes news, letters and fiction, and amounts to 663,695 words, covering the period 1867–2012. In the case of the WSAfE corpus, there is limited coverage of the period 1960–1990, and in the case of the BSAfE corpus, no fiction could be located before the 1920s, while no letters were collected after 1965. These two corpora do have limitations, but they represent the most complete corpus resources available for the study of grammatical change. No historical corpus of AfkE is available, although a corpus of contemporary AfkE has been used by Kruger and Van Rooy (2016, 2019), and for comparative purposes, the historical corpus of Afrikaans (Kirsten, 2019) is available to study ongoing change in one of the contact languages for SAfE.

2 | A FRAMEWORK FOR THE STUDY OF CONVERGENCE

The idea of convergence within a setting such as South Africa is not a central concern of Kachru’s (1985) model of world Englishes. His concern is with the recognition of the linguistic effects of nativisation, which is a divergent current in his
conceptualisation, where a variety of English acquires new features that differentiate it from its colonial parent variety. Convergence is much more central to Schneider’s (2007) model, where he recognises various kinds of convergence and divergence. In the earliest, foundation stage, dialect differences between the transplanted English dialects level out through koineization (an idea that corresponds to Trudgill’s, 2004 analysis of the inevitability of colonial dialect formation), while a degree of divergence is also expected when stage 4, endonormative stabilisation, is reached. Schneider (2007) highlights divergence at both stage 3 and 5: nativisation occurs in stage 3 (and already starts in stage 2), which results in divergence between the new English and its colonial parent; renewed differentiation takes place in stage 5 to overturn the homogenising effects of endonormative stabilisation in stage 4.

Both models highlight nativisation as a process by which English becomes different in a new place due to contact with indigenous languages and non-L1 speakers of English. To Kachru, this phase requires recognition of the nativised forms in an Outer Circle context, but Schneider focuses on the interaction between two strands, the STL and IDG (more fully, the settler and indigenous populations), which correspond to a degree to Kachru’s Inner and Outer Circle contexts. However, where Kachru’s focus is on the prototypical cases of entire countries that mainly experienced one form of colonisation, resulting in a dominant pattern of nativisation, Schneider’s focus is on the interaction between the two strands, even if there is a clear numerical advantage of the STL or the IDG strands. Nevertheless, for progression to stage 4 in his model, recognition of the nativised forms is also required, and in this sense, there is correspondence between Schneider and Kachru’s views of nativisation, followed by recognition of the change.

Models of world Englishes provide expectations that guide the exploration of South African English. Divergence from the input variety results from nativisation due to contact between English and other cultures and/or languages. Cultural contact poses new expressive challenges for English and creates a context for borrowing from local languages, while contact with languages also offer subconscious opportunities for transfer of grammatical and phonological patterns. To the extent that a particular contact situation differs from the next one, divergence will result as the particulars of nativisation and transfer in one context will differ from the particulars in the next. Convergence results from prolonged contact between two (or more) groups of speakers of English, whose Englishes initially differ, but comes to grow more similar. An important idea from Schneider’s (2007) model is that identity alignment between the two groups of speakers functions as condition for the convergence between the two groups of speakers.

Van Rooy (2014b), following a suggestion by Bekker (2012), argues that the contact situation in which convergence potentially occurs can be narrowed down even further, below national level, to intra-national group level. In the specific case of South Africa, this implies that different ethnic contact contexts may lead to different potential convergence points in the country. Mesthrie (2014) goes even further, by pointing to cases where convergence might even be at the level of individuals, whose own situations may be atypical for their presumed group (defined in ethnic, but also economic terms), and who display convergence with the norms of another group due to prolonged contact.

There are a number of unresolved questions about convergence and divergence in South African English, which can be articulated in terms of the dynamic model of Schneider (2007). A first question revolves around the degree to which convergence in South Africa leads to the hybridity that one expects to results from contact, or whether it rather means convergence on the norms of the STL strand. A second question concerns the extent to which identity realignment is in evidence when convergence takes place, or whether, in the alternative, contact is enough to lead to convergence of grammatical forms.

3 | CHANGE IN MODAL AUXILIARIES

Modal auxiliaries are known to undergo changes in frequency in ‘native’ varieties of English around the world. Leech, Hundt, Mair, and Smith (2009) identify the increased frequency of quasi-modals like have to, be able to, and supposed to, and the decreased frequency of core modals, especially modals with the lowest frequencies, like shall and ought (Collins, 2009; Leech, 2011). The modal must has experienced a particularly dramatic drop in frequency, which Myhill
(1995) traces to changing social values in the United States (and later the United Kingdom), in terms of which communica-
tion in terms of authority and hierarchical situations made way for more equal relationships, and hence less use
of the strong, face-threatening modal must, which imposes obligations from the position of authority. In view of these
world-wide developments in English, the question is what has happened in South African English. Beeton and Dorner
(1975: 122) caution against the use of must ‘in a simple inquiry’ or with constructions that ‘do not suggest compulsion,’
raising the suspicion that these authors already noticed divergence between SAFE and other varieties of English. How-
ever, Bowerman (2004: 477) observes that a characteristic feature of SAE must is that it has less social impact than in
other varieties and often overlaps with polite uses of should.

Wasserman’s (2014: 241) results for the individual modals in WSAfE show that there is a very gradual, but statist-
ically non-significant increase in overall frequency of modals from the early 19th century to the middle of the 20th
century, but a significant decline from there to the end of the 20th century. As far as individual modals are concerned,
there is a significant increase in overall frequency (from early 19th to end of the 20th century) of can, no significant fre-
quency change for could, may, must and will, and a significant decrease in frequency for might, need, ought, shall, should
and would. Particularly noticeable is that ought declines from 16.5 to 0.7 per 100,000 words, and shall from 85.5 to 8.5.
The sharp increase in frequency of can (from 162.8 to 251.7 per 100,000) words, and the fact that must does not decline
in frequency (there is a slight, but non-significant, rise from 98.9 to 113.5 per 100,000 words) distinguish WSAfE from
the varieties investigated by other researchers using historical corpora.

Wasserman and Van Rooy (2014: 47) point out that these frequency differences can be attributed to contact with
Afrikaans, where the two cognate forms, kan and moet are even more frequent than in English, and are the first and
third most frequent modals in Afrikaans. Subsequent to the studies of Wasserman (2014) and Wasserman and Van
Rooy (2014), Erasmus (2019) completed a diachronic corpus analysis of all Afrikaans models from the beginning of the
20th to the 21st century. She reports a consistent, sharp increase in the frequency of kan, a slight increase in moet,
and otherwise relatively steady or slightly declining frequencies for the other modals (Erasmus, 2019: 199). Thus, the
notion that Afrikaans and WSAfE follow similar patterns of frequency changes, and that the exceptional frequencies
of WSAfE (compared to other L1 varieties) are closely matched by the frequency changes of the cognate modals in
Afrikaans, further strengthen the conclusion that contact with Afrikaans is an important factor in the development of
the modals of WSAfE. Quasi-modals in WSAfE show a significant increase in frequency, overall, which is consistent with
other varieties and often overlaps with polite uses of the strong, face-threatening modal must. Erasmus (2019) and
Wasserman (2014) also note, there is a degree of reciprocity

Wasserman (2014) also identifies a semantic shift in the use of must. She identifies examples such as (7), which are
quite unlikely from the perspective of the conventional strong obligation encoded by must elsewhere in the world by
‘native’ English varieties:

(7) Some day you must come up with me. You can have a horse. (WSAfE Letter, 1892)

Wasserman (2014) argues that Afrikaans influence is also noticeable behind the semantic shift. Afrikaans moet also
reveals a similar range of strong and median strength obligation than SAFE. Erasmus (2019) offers qualified confirma-
tion for Wasserman’s findings in her diachronic study of the cognate form in Afrikaans, moet. Especially after the middle
of the 20th century, the frequency of median strength deontic and epistemic meanings increases at the expense of the
strong force of the 1940s, which Erasmus (2019: 258–265) attributes to changes in the socio-political and cultural
landscape of Afrikaans. However, as Erasmus (2019) and Wasserman (2014) also note, there is a degree of reciprocity
in the semantic changes of Afrikaans and English during the 20th century, as Afrikaans also adopted senses like the
extended range of epistemic senses characteristic of English, but not of the colonial parent variety of Afrikaans, viz.
Dutch. Both the overall frequencies and the semantic details of must/moet in SAFE and Afrikaans show converging
trends in their ongoing change, which clearly set them apart from their respective colonial parent varieties.

By contrast, in Australian English, a close colonial sister of WSAfE, which had very similar input to its formation
(and is only about 30 years older than WSAfE), there is no similar development. The modal must declines to an even
lower frequency in Australian than in British or American English (Collins, 2014: 15). The divergence from both British
Table 1  Frequencies of modals and quasi-modals in written BSAfE and WSAfE (normalised values per 100,000 words; *significant difference p < 0.05, measured by log-likelihood test)

<table>
<thead>
<tr>
<th></th>
<th>ca 1950</th>
<th></th>
<th>ca 2000</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BSAfE</td>
<td>WSAfE</td>
<td>%diff</td>
<td>BSAfE</td>
</tr>
<tr>
<td>Modals</td>
<td>1128</td>
<td>1159</td>
<td>−3%</td>
<td>1253</td>
</tr>
<tr>
<td>Quasi-modals</td>
<td>247</td>
<td>263</td>
<td>−6%</td>
<td>287</td>
</tr>
<tr>
<td>Total</td>
<td>1375</td>
<td>1422</td>
<td>−3%</td>
<td>1540</td>
</tr>
</tbody>
</table>

Table 2  Normalised frequency per 100,000 words of must in historical BSAfE corpus

<table>
<thead>
<tr>
<th>Period</th>
<th>Fiction</th>
<th>Letter</th>
<th>Reportage</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>19(2)</td>
<td>106</td>
<td>96</td>
<td>110</td>
<td>104</td>
</tr>
<tr>
<td>20(1)</td>
<td>103</td>
<td>190</td>
<td>70</td>
<td>67</td>
</tr>
<tr>
<td>20(2)</td>
<td>70</td>
<td></td>
<td>65</td>
<td></td>
</tr>
</tbody>
</table>

and Australian English, and convergence between WSAfE and Afrikaans, support the thesis of convergence between Afrikaans and WSAfE in the sustained contact situation, consistent with the expectation set out by Van Rooy (2014b: 26–28), and despite identity alignment not taking place (Coetzee-Van Rooy, 2014: 52; Van Rooy, 2014b: 27). This raises the question of what happened in BSAfE. Van Rooy and Wasserman (2014) compiled a historical corpus of BSAfE covering the mid-20th century and the end of the 20th century, and observed that overall, the frequencies of modals and quasi-modals did not change significantly in BSAfE, while the corresponding overall frequencies for WSAfE show a decline. Thus, in the middle of the 20th century, there is no significant difference in the frequencies of the modals, but both modals and quasi-modals show a statistically significant difference of frequency by the end of the century, as shown in Table 1 (from Van Rooy & Wasserman, 2014: 59).

Van Rooy and Wasserman (2014: 60) do note a decline in frequency of must in BSAfE, although the decline fails to research statistical significance. Since the completion of the work by Van Rooy and Wasserman (2014), the historical corpus of BSAfE has expanded, so for the purposes of this paper, an extended investigation was conducted on the letters, fiction and reportage in BSAfE. The results are presented in Table 2.

The data show a steady overall decline in frequency, confirming the tentative trend observed by Van Rooy and Wasserman (2014) on a smaller corpus, except for a fluctuation in the letter register in the second half of the 20th century, which masks the overall trend somewhat on aggregate. This picture is more consistent with the decline of must in varieties of English elsewhere in the world, and less so with the relative stability of the frequency of must in WSAfE reported by Wasserman (2014). In terms of the semantics, Van Rooy and Wasserman (2014: 61) report that, unlike in WSAfE, the strong obligation sense of must remains the dominant one by far, accounting for 68 per cent of all instances of use in both the periods (mid-20th century and present) they analysed, while the median strength obligation meaning, which has seen an increase in WSAfE remains negligible at 4 per cent and 6 per cent respectively in the two periods. Thus, the dominant sense of strong obligation is equally prominent in example (8) from the 1950s as in example (9) from the more recent period; both examples taken from Van Rooy and Wasserman (2014: 62).

(8) But every African on the Reef knows how much more must be done before this dreadful wave of crime is stopped. (BSAfE reportage, 1951)

(9) It will prohibit illegal evictions, but procedures must be followed to evict illegal occupiers. (BSAfE reportage, contemporary)
The divergence between WSAfE and BSAfE is especially prominent from the middle to late 20th century, both in frequency and in semantic change of must. This coincides with the period of apartheid, during which contact between black and white was at its most difficult, whereas in the earlier period, the Black South African elite, who were the writers of fiction, letters and news reportage, were in much closer interaction with L1 speakers, especially through associations built by missionary schools. The centrality of contact between speakers as condition for convergence of language use is therefore supported by the divergence between WSAfE and BSAfE, especially when contrasted to the degree of convergence seen between WSAfE and Afrikaans.

4 | CHANGES IN THE PROGRESSIVE ASPECT

The frequency of the progressive aspect has seen a gradual increase in the 19th and 20th century across all L1 varieties. This is attributed to structural extension across the entire verbal paradigm, including the emergence of passive progressives to replace the passival in the 19th century, but also due to functional extension to new meanings (Kranich, 2010: 32–35; Leech et al., 2009: 122–124; Mair, 2012). In the new Englishes, the extension goes even further, to combine with stative verbs that in contexts where L1 varieties do not typically use the progressive (Mesthrie & Bhatt, 2008: 67). Sharma (2009: 185) offers an explanation for the pervasive extension of the progressive in the new Englishes. She argues that it results from the problem of mapping from the aspect-rich substrate languages, such as Hindi, to the much smaller set of aspectual contrasts in English. The richness of aspectual contrasts in comparison to English is also true for the Bantu languages of Africa (Nurse, 2003), and specifically for the Southern Bantu languages that form the substrate to BSAfE, as shown by Piotrowska (2014: 61–71).

The extension of the progressive aspect has been identified as characteristic of BSAfE by a number of researchers, at least since the prescriptive literature of the 1970s (Scheffler, 1978), and confirmed by corpus research (De Klerk, 2003: 468). The likely explanation is transfer: according to Makalela (2004), there is transfer of the entire ‘Bantu language logic’ that does not draw a conceptual distinction between verbs in terms of stativity, while Van Rooy (2006) points to the more extended semantic possibilities of the persistitive (PERS) aspectual prefix sa- in the Southern Bantu languages, as illustrated by example (10).

(10) Re sa rua dikgomo.

SC 1st pl PERS farm cattle

‘We are still farming with cattle.’

Corpus analyses of the frequency of the progressive (Van Rooy, 2006, 2014a) show that in synchronic data, BSAfE progressives are more frequent than in any other variety for which frequencies have been determined in corpus research (especially by Collins, 2008). This includes including WSAfE (Van Rooy, 2006), but also L1 varieties from elsewhere in Africa or from Asia, both in written and spoken usage (Collins, 2008; Van Rooy, 2014a).

When we turn to the development over time, Piotrowska (2014: 136) reports the most complete results, which are reproduced in Figure 1. For available registers, we mostly observe a small increase in frequency over time in BSAfE, while the results ambiguous for WSAfE, with a number of registers declining in frequency from the middle of the 20th century to the end of the 20th century. The important point that emerges from the comparison is that for most available registers, in earlier periods, the frequency of the progressive was higher in WSAfE, but for the most recent period, the frequency is generally higher in BSAfE.

The question is whether the change in frequency is due to an extension in the meanings of the progressive in BSAfE. The short answer to this question is no, the innovative extensions of the progressive are attested in the historical BSAfE data from the start, and do not show a proportional increase over time (Van Rooy & Piotrowska, 2015). On the one hand, the overall proportion of stative verbs used with the progressive aspect fluctuates in a narrow band between 10 per cent and 20 per cent of all progressive instances over the course of time, without the overall proportion of
FIGURE 1  Normalised frequency per 100,000 words of progressive in BSAfE and WSAfE registers over time

statics showing a clear increase (Van Rooy & Piotrowska, 2015: 478). In addition, the different kinds of meanings that the stative progressive conveys do not show an increase from more prototypical ‘native’-like uses, such as using the progressive to convey temporary states, as in example (11), to less prototypical uses, when judged from the ‘native’-speaker perspective, such as ongoing states in example (12) or states of unlimited duration, as in example (13), all taken from Van Rooy (2014a), where the full classification framework is presented.

(11) He revelled in those moments when he’d be wearing striped suits and sit in the front seat while Laurie sat in the back. (BSAfE fiction, contemporary)
(12) he is still depending on his family (BSAfE spoken, contemporary)
(13) you know even the white people the black the green and the yellow people mm they are having their ancestors (BSAfE spoken, contemporary)

The meanings of an ongoing or unlimited stated are also well-attested in the earlier data, as shown by example (14) as well as (5) above, repeated here as (15), taken from Van Rooy and Piotrowska (2015).

(14) It is not, however, in his relation to the European electors that we propose to notice Mr Innes’s visit to his Peddie constituents; this may be left for journals which are professedly catering for our white neighbours. (BSAfE reportage, 1886)
(15) Cycling on the East London road via Mount Coke, then branching off to Berlin, we saw a village on top of a mountain. As the main road was not leading there it was necessary to look for a footpath up the mountain side. (BSAfE reportage, 1947)

The point about examples like (14) and (15) is not that they are completely unacceptable in ‘native’ usage, just that they do not convey temporary states, but states of extended duration. Thus, in example (14), there is no indication in context that the journals being modified by the relative clause containing the progressive verb ‘catering’ are likely to alter their readership – they will continue to cater for the ‘white neighbours.’ Likewise in example (15), there is no indication in context that the main road will be extended in the foreseeable future to lead to the village that was seen on top of the mountain. Piotrowska (2014: 143–147) reports that the overall proportion of temporary states is consistently higher in WSAfE than in BSAfE, which in turn results in consistently higher proportions of ongoing and unlimited states together for BSAfE, although there is noticeable variability within periods and registers.
The most important finding that emerges from the historical analysis, however, is that the extended temporal sense of the BSAfE progressive with stative verbs is not a recent innovation, or even a trend that grows over time. These extended senses were present from the earliest data, and in higher proportions that the same senses are in the corresponding data from WSAfE. Thus, transfer of the extended meaning took place early in the development of BSAfE and has remained a feature of the variety ever since. By contrast, the proportion of the extended senses has not increased over time in WSAfE; if anything, WSAfE used fewer progressives as time went by, resulting in overall divergence between BSAfE and WSAfE as far as the use of the progressive is concerned.

5 | CHANGES IN COMPLEMENTISER OMISSION

The final construction under consideration is the omission of the complementiser that from object complement clauses, as illustrated by the contrast in (16), taken from Kruger and Van Rooy (2019). The omission of the complementiser is on the increase over time, while the use of the overt that gradually decreases. There are clear register effects, such that speech and reportage lead the change, while formal and academic writing retain a much larger proportion of overt complementisers (Biber, Johansson, Leech, Conrad, & Finegan, 1999: 680–683).

(16a) I just wish hereby to say that if I wanted to remain in Lovedale, I would do so in spite of you. (WSAfE, Letters, 1908)
(16b) I just wish hereby to say ø if I wanted to remain in Lovedale, I would do so in spite of you.

Unlike the previous two phenomena, there is no strong social evaluation attached to this feature, despite associations with formality. There is no prescriptive comment similar to the caution against the extended use of must or the ‘overextension’ or ‘incorrect’ use of the progressive with stative verbs. Thus, changes in the omission rate of the complementiser are below the level of social awareness and offer a different kind of perspective on ongoing language change and the possible factors that contribute to it.

Historically speaking, Afrikaans did not inherit complementiser omission from its Dutch ancestor. Even in 18th century Cape Dutch, the omission of the complementiser is still extremely marginal. However, since the end of the 19th century, Afrikaans began to catch up with WSAfE and by the end of the 20th century, Afrikaans overtook WSAfE as far as rate of complementiser omission is concerned (Kruger & Van Rooy, 2016, 2019). However, in the ‘non-native’ AfrE variety of SAfE, Kruger and Van Rooy (2016) find that the rate of complementiser omission is much lower, which they explain as a result of the additional processing constraint the accompanies language production in the second language. WSAfE is an important factor in the innovation and diffusion of complementiser omission in Afrikaans, but Kruger and Van Rooy (2016) argue for a reciprocal relationship in which Afrikaans also exerts further influence on WSAfE to accelerate the rate of ongoing change in WSAfE.

In an analysis of the historical data, Kruger and Van Rooy (2019) find that the relationship between WSAfE and British English is quite varied and differs considerably across registers. WSAfE letters don’t show much change, whereas fiction shows an extensive change from a strong preference for an overt complementiser at the beginning of the 19th century to the opposite preference for the omission of the complementiser by the end of the 20th century. In news reportage, there is a pattern of gradual increase in complementiser omission from the beginning of the 19th century to the end of the 20th century. In many ways, SAfE follows a clear trend that other L1 varieties also follow, and statistical modelling (Kruger & Van Rooy, 2019: 26–27) also points to negligible differences between WSAfE and British English. The Afrikaans historical data likewise show a clear trend of increase in the rate of omission, with less differentiation between registers than in WSAfE. Contact seems to work primarily in the direction from English to Afrikaans here. The results for WSAfE, taken from Kruger and Van Rooy (2019: 23–25), are presented in Figure 2, where the fluctuation and register differentiation are visible, but also the general downward trend over time, as omission increases and overt that decreases.
When we turn our attention to BSAfE, Partridge (2019) finds that complementiser omission is widely attested in the reportage register of BSAfE, more frequent than the corresponding data in WSAfE (which are the same data that were used by Kruger & Van Rooy, 2016), while it is much less frequent in the popular and instructional register than WSAfE, and equally infrequent in the academic register. Apart from reportage, then, BSAfE is similar to AfrE in allowing complementiser omission, but at lower omission rates than WSAfE, which Partridge (2019: 207) attributes to processing constraints as well, in agreement with the argument of Kruger and Van Rooy (2016). Overall, Partridge (2019: 207) argues against the view proposed by Mesthrie (2006: 119) and Makalela (2013: 103) that transfer of a compulsory complementiser in the substrate languages results in the ‘undeletion’ of the overt complementiser, as one manifestation of Mesthrie’s (2006) proposed anti-deletion, since she finds higher rates of omission in written language than what they report in spoken language from their data. It may also be the case that writers for publication are more proficient, and thus able to vary their choices more than the mesolect speakers investigated by Mesthrie (2006).

The finding in the reportage register is explained by Partridge (2019: 220) as due to influences in the publishing environment, where there is an overall trend in all SAfE varieties, and in Afrikaans, to omit the complementiser in order to condense newspaper reports. In addition, Partridge (2019: 207) documents a clear preference for the verb say in particular to condition complementiser omission far more than other complement-taking verbs, as illustrated by example (17), which corresponds to the finding that the Afrikaans equivalent sê also associates very strongly with omission (Van Rooy & Kruger, 2016).

(17a) Speaking to Sedibeng Ster Sports, Ncala says the only challenge they face is lack of sponsorship for township athletics clubs. (BSAfE reportage, contemporary)

(17b) Brown sê institusionele beleggers gebruik Satrix se Dividend Plus-fonds om hul beleggings te diversifiseer. (Afrikaans reportage, contemporary)

‘Brown says institutional investors use Satrix’s Dividend Plus fund to diversify their investments.’

While there are differences in the findings of Partridge (2019) on written language and the earlier findings by Mesthrie (2006), who used spoken data, complementiser omission in BSAfE has not been investigated in historical data. Figure 3 reports on the analysis of the historical BSAFE corpus.²

The data in Figure 3 show a fairly consistent trend towards higher rates of complementiser omission over time, and moreover less extensive differences between registers what was found by Partridge (2019) in a different corpus.³ Unlike the other two phenomena under investigation in this paper, there is similarity and convergence in the changes in omission rate for the that complementiser, even if this conclusion has to be limited to written language, since there is no historical corpus of spoken BSAFE. BSAFE therefore clearly does partake in some long-term changes affecting WSAFE.
The three grammatical features analysed in this paper do not show consistent similarity in their patterns of change. The modals show convergence in the changes taking place in Afrikaans and WSAfE, which in turn results in WSAfE diverging from other varieties of English, such as its parent variety, British English, as well as Australian and American English. By contrast, BSAfE follows neither the frequency changes nor participates in the semantic changes of WSAfE, but if anything, shows more similarity to British English. The progressive aspect shows initial change in BSAfE through transfer of the semantics of the persistitive in the substrate Bantu languages, which leads to extended use of the progressive with stative verbs. After the initial transfer, little further semantic change is observed in BSAfE. However, both BSAfE and WSAfE show a gradual increase in the frequency of use of the progressive construction since the 19th century. The frequency of complementiser omission increases over time in both WSAfE and BSAfE, while Afrikaans also shows evidence of complementiser omission under influence of English. The patterns of change are more similar here, except that the L1 variety is generally more receptive to complementiser omission, while the ‘non-native’ varieties BSAfE and especially AfrE lag behind (except in the reportage register of BSAfE), which Kruger and Van Rooy (2016) and Partridge (2019) attribute to a preference for the more explicit form to manage the processing strain from using a second language.

In general, it seems that long-term changes in frequencies of variable grammatical show surprising similarities across varieties, despite differences in the extent of contact. Where contact is more, it is possible that the changes in frequencies are matched more closely, as the case of complementiser omission and modal frequencies of Afrikaans and WSAfE show, with bidirectional influence in both languages, resulting in both of them showing divergence from their respective colonial parent varieties. However, BSAfE progressive and complementiser frequencies also follow a similar trajectory of long term frequency change. Such convergent changes are consistent with the view of Trudgill (2004) that convergence might take place without the prerequisite of favourable social circumstances, which is something that one would expect from Schneider’s (2007) notion of identity alignment. Closer contact seems to be a more important factor for functional extensions to be shared, though, which is more consistent with Schneider’s view, and in particular the proposal by Bekker (2012) and Van Rooy (2014b) that such closer contact plays out at sub-national level in individual contact settings, rather than at a larger level. Thus, the semantic changes in the modals of WSAfE and Afrikaans are shared in a situation where contact is more extensive, while the semantic innovation of the extended progressive is not shared by BSAfE and WSAfE speakers.

Kruger and Van Rooy (2017: 34–37) present evidence that the shibboleth features, such as the use of the progressive with having, are clearly avoided by WSAfE, and both WSAfE and AfrE editors remove such uses when they
encounter them in texts written for publication. However, they also find that more subtle extensions of the semantics of the progressive that are not clearly incompatible with the constructional semantics of WSAfE are not removed in the majority of cases by the editors; the editors just don’t use those forms as frequently in their own English, as can be seen from the way they rewrite certain passages more extensively when editing. Likewise, Partridge (2019: 208) finds that the same editors generally don’t change the choice of overt or covert complementizer very often, and only add an overt complementizer in a small number of cases, leading to an overall change of only 1 per cent in the omission rate, and not nearly reaching the level of statistical significance. If an innovative feature is clearly regarded as non-standard, it is unlikely to diffuse beyond the originating group, but the moment an innovative feature is below the level of public awareness, it has a chance to be accepted at least in passive use, if not yet in active use. The attitudinal dimension of the acceptance of change, emphasised by Kachru (1985) as precondition for nativisation to lead to long-term changes, is in evidence in this finding.

Convergent change in SAfE is conditional on close contact, even though such contact need not extend to identity realignment between the non-L1 speakers (IDG strand in terms of Schneider, 2007) and L1 speakers (STL strand). If the contact is extensive enough, then changes can diffuse in both directions, and also affect substrate languages, such as Afrikaans. However, if the contact is not very close, as has been the case between WSAfE and BSAfE until recently, the convergent changes seem to be limited to overall patterns of frequency distributions (‘probabilistic indigenisation,’ as proposed by Szmrecsanyi, Grafmüller, Heller, & Röthlisberger, 2016) without reaching deeper levels of structural or functional nativisation. If the context of contact changes, as is the case with South African since 1994, the possibility of convergence between WSAfE and BSAfE increases, as shown by Mesthrie (2010) and Mesthrie et al. (2015), and in future, such new contexts of contact may become visible in convergence of grammatical changes as well.

ACKNOWLEDGEMENTS

The funding of the National Research Foundation of South Africa (NRF) for the project ‘The development of South African English’ is acknowledged with gratitude, and likewise the contribution of the fellow researchers in the project: Ian Bekker, Caroline Piotrowska, and Ronel Wasserman. For discussion of various ideas in this paper, I also want to thank Susan Coetzee-Van Rooy, Johanita Kirsten and Maristi Partridge. Parts of the work reported here developed from collaborative research with Haidee Kotze, whose contribution I want to acknowledge as well. Earlier versions of some of the analyses in this paper were presented at the 2014 IAWE conference in New Delhi and the 2015 LSSA conference in Potchefstroom, and I would like to acknowledge the input from the audiences at these two conferences. All views and conclusions in this paper are my own and should not be attributed to NRF or anybody else.

NOTES

1 The expanded corpus used for this analysis consists of 726,339 words, but the coverage of the late 19th century is still limited compared to the other periods, with only 55,110 words. Register distributions within periods are also uneven, but all periods for which there are data contain at least 10,000 words per register in a given period. The periodisation is mainly per half century, except that the data from the 1990s and 2000s are grouped together to form the ‘21st century’ or better, the ‘present’ state of the language.

2 The data in Figure 3 are based on a manual analysis of the 20 verbs that were found to be the most frequent ones taking complement clauses in previous research on SAfE. The method adopted is closely aligned to the method in Kruger and Van Rooy (2019), ensuring comparability of the results in Figures 2 and 3, with the exception that a smaller number of verbs have been used. However, the 20 most frequent verbs accounted for 85 per cent of all the data reported for WSAfE, thus the overall picture should be sufficiently comparable.

3 Partridge (2019) used the WSAfE and BSAfE components of the parallel editing corpus developed by Haidee Kotze (Kruger) and her team. The WSAfE and AfrE components were used by Kruger and Van Rooy (2016) in their analysis. This corpus consists of original texts produced by primary authors before professional editing, and a parallel version that was prepared by a professional editor for publication, called the unedited and edited versions respectively. This is a specialised corpus and, during its compilation, finding unedited authentic texts written for publication and their edited counterparts was more important.
that achieving a balanced distribution of registers. The representativity of the editing corpus was therefore subordinated to the ecological validity. The historical corpus of BSAfE, by contrast, is somewhat more balanced in terms of diversity of texts in its composition, especially closer to the present, but does not contain all the registers of the editing corpus. Such limitations on the data are to be expected when working with commercially less viable varieties of English, as funding to develop larger resources is harder to come by. This situation has to be accepted for the time being, while it should remain an important research priority to expand corpus resources to ensure more representativity as time goes by.

REFERENCES
