Diachronic register change

A corpus-based study of Australian English, with comparisons across British and American English

Peter Collins,[1] Minna Korhonen,[2] Haidee Kotze,[3]
Adam Smith[2] & Xinyue Yao[4]
[1] University of NSW | [2] Macquarie University | [3] Utrecht University |
[4] Renmin University of China

A number of studies have found that grammatical differences across registers are more extensive than those across dialects. However, there is a paucity of research examining intervarietal register change, exploring how registers change differently over time in different regional varieties. The present study addresses this diachronic deficit, focusing on grammatical developments – from the early 20th to the early 21st century – in corpora representing three written registers and two speech-based registers in Australian, British and American English. We conducted a factor analysis on 68 lexicogrammatical features to identify six dimensions of register variation, and subsequently investigated the diachronic change of the five registers across these dimensions. We interpret our findings in terms of the differential effects of broad social changes on individual registers, in light of existing findings on trends of change in different registers and varieties.

Keywords: Australian English, British English, American English, multidimensional analysis, lexicogrammar, register change

1. Introduction

This study investigates diachronic change in five registers of three varieties of English, aiming to answer the following overarching research question: "How do registers change differently over time in different varieties?" A number of previous studies have observed that grammatical differences across registers (and even sub-registers) are more extensive than those across dialects (e.g. Biber, Johansson, Leech, Conrad, & Finegan 1999; Collins & Yao 2018; Kruger & Van Rooy 2018; Leech, Hundt, Mair, & Smith 2009; Hundt, Schneider, & Seoane 2016; Yao &

Collins 2019). However, there is a paucity of research addressing diachronic change in the grammatical "profiles" of particular registers in the 20th and 21st centuries. British and American diachronic register studies (typified by, for example, Biber & Finegan 1989; Atkinson 2001; Biber & Conrad 2009) are generally longer-term, covering the past three or four centuries (but see Mair & Hundt 1999; Westin 2002 for exceptions). Changes in Australian English (AusE) – an established Southern Hemisphere "Inner Circle" variety – have begun to attract scholarly interest in recent years. In fact the corpora used in the present study have also been used in several such studies.[1] Collins & Yao (2018) and Yao & Collins (2019) observe differences in the nature and extent of grammatical change in written AusE, British English (BrE) and American English (AmE) from 1931 to 2006. Kruger & Smith (2018) and Kruger, Van Rooy & Smith (2019) examine grammatical developments in the Australian Hansard from 1901 to 2015.

The present study complements previous multi-register corpus-based studies of the historical development of AusE as a regional dialect of English over the past century (Collins 2015; Collins & Yao 2018; Yao & Collins 2019). In this study we focus on the historical development of selected registers (as determined by the availability of suitable diachronic corpora of AusE, and of the two reference varieties that have influenced its development, BrE and AmE). We seek not only to redress this diachronic deficit, but also to make inroads into a hitherto under-explored area in register studies, namely intervarietal register change.

Accordingly, we aim to systematically describe diachronic change in three written registers and two speech-based registers, across three varieties of English – AusE, BrE, and AmE. We also offer some tentative explanations for the changes we observe, based on existing models of varieties of English (e.g. Schneider 2007). The speech-based material is drawn from corpora sourced from the Australian House of Representatives Hansard and the British House of Commons Hansard, covering the period 1935–2015. Unfortunately, AmE had to be excluded from the Hansard comparison, due to the US Senate congressional records not being directly comparable with the Hansard records. We distinguish two (sub)registers: monologues, which consist of uninterrupted parliamentary speeches; and dialogues, which include both debates, and material from Question Time (including only "questions without notice", i.e. questions not prepared and submitted in advance of the proceedings). Textual material for the three written registers – fiction, press, and learned writing – was derived from Brown-family

---

1. The first person to compile a diachronic corpus of AusE, and to produce studies based upon it, was Clemens Fritz (Fritz 2007a, 2007b). With two million words of text covering the period 1788–1900, Fritz's Corpus of Oz Early English (COOEE) comprises four registers: speech-based (15%), private written (35%), public written (40%), and government English (10%).

corpora, representing BrE, AmE and AusE over the period 1931–2006. For more information on the corpora on which the study was based see Section 4.

Change in these registers has been investigated in a number of studies. The Hansard has been found to have undergone contrasting trends of colloquialisation (the spread to other genres of grammatical features typical of casual speech such as first-/second-person pronouns and contractions) – partly through more permissive editorial practices – and anti-colloquialisation in the form of densification and decreased dialogic engagement (Kruger et al. 2019). Fiction has drifted towards a more oral, less literate, style since the 19th century (Biber & Finegan 1989; Biber & Gray 2016). Press has moved towards a more informal and less narrative style, but also a denser and more condensed style (Biber 1988; Mair & Hundt 1999; Biber & Finegan 2001). In stark contrast, learned writing has become more literate and compressed, and less oral (Mair & Hundt 1999; Leech et al. 2009). Our aim is to add to what is already known about change in these registers, and to relate this to varietal variation, invoking such explanatory factors as sociohistorical context, and evolutionary status.

The structure of the remainder of the paper is as follows. In Section 2 we provide a short historical sketch of AusE with a focus on the general sociohistorical background and prevalent language attitudes in 20th century Australia. Section 3 details the situational characteristics of the five registers included in the study. Section 4 describes the corpora used for the study. Section 5 presents the 68 grammatical features on which the study was based. Section 6 explains the methodology, specifically the two main steps in which we carried out our analyses, namely the exploratory factor analysis (EFA) and the regression analysis. Section 7 presents the six dimensions that result from the factor analysis and the distribution of dimension scores across registers, varieties and time. Section 8 is devoted to a summary and interpretation of our findings, and Section 9 to conclusions.

2.   Historical sketch of Australian English in the 20th century

The period of time covered in this study, from the early 20th to the early 21st century, saw AusE transition from a variety oriented to the norms of BrE to stabilised endonormativity. However, during the first half of the 20th century, trenchant criticisms of the Australian accent were commonly voiced (as for example in the letters to the editor of the *ABC Weekly*, which regularly railed at the putative slovenliness and ugliness of Australian speech: see Reeve 1989), and invidious lexical comparisons with British practices were also commonplace, with users of AusE under attack for their alleged overuse of slang, as in Desmond's (1911:19)

accusation that "The Australian interlards his conversation with large quantities of slang, which makes him frequently unintelligible to the visitor."

From the 1960s however, as argued by Moore (2008), Australians increasingly embraced their accent and vocabulary as standard:

> The final stage in the development of the Australian language was the discarding of the external English standard of language, as exemplified by Received Pronunciation and the class-based judgment of vocabulary, and the embracing of the Australian accent and vocabulary as standards in their own right […] the crucial period in this shift was 1966 to 1988.   (p.xii)

According to Moore (2008: 153–163) this development is linked to the "decline of Britishness" in Australia, the seeds of which were sown during WWII when Australians found themselves abandoned by Britain in the face of Japanese attacks following the fall of Singapore in 1942 (see Schneider 2007: 122–123). The major driver, however, is argued by Moore to have been Britain's moves to enter – and ultimately its entry into – the European Economic Community. Also identified as having a role to play in the shaping of Australia's linguistic identity, by Blair (1993), is the wave of nationalistic feeling from the early 1970s, generated largely by the election of the Australian Labor Party after twenty-one years of Conservative rule, and heralding Australia's coming of age and final break from its erstwhile British-oriented outlook. Some empirical support for this historical interpretation is also provided by Yao & Collins's (2019) finding that AusE experienced rapid grammatical developments in the period from 1961 to 1991.

Prefiguring Schneider's (2007: 122ff) identification of the 1980s as marking the achievement of phase 4 (endonormative stabilisation) and entry into phase 5 (differentiation), Blair declares that: "Australian English is currently a self-confident dialect, reasonably secure in itself, and prepared to set its own standards. It is confident enough to be open to outside cultural influences, as is the society in which it functions" (1993: 70).

One characteristic of AusE that has attracted a good deal of both popular and scholarly attention is the high level of colloquiality in its lexis, pronunciation and grammar (see for example Wilkes 1978; Seal 1999). In their recent diachronic investigation of grammatical colloquialisation in AusE, BrE and AmE, Collins & Yao (2018) note dramatic changes in AusE in the second half of the 20th century and suggest that their findings reflect the emergence of Australia's linguistic independence from its British "parent" and from possible American influence.

We shall use our diachronic corpus data to examine whether the developments in Australia's sociohistorical context and its linguistic identity described in this section (which are discussed in more detail in Collins & Yao 2018; Yao &

Collins 2019), provide us with a source of explanation for grammatical changes in AusE registers identified in the study.

3.   Situational characteristics of the five registers

The English registers examined in this paper are admittedly predetermined by the make-up of the corpora used. Fortuitously however, the majority are amongst the most intensely studied in the literature and similar to those used in Biber et al.'s (1999) influential and pioneering corpus-based *Longman Grammar of Spoken and Written English*. The reason we must refer to our register selection as "similar to" rather than "the same as" that in Biber et al. (1999) is that instead of conversation, our study includes the nearest available speech-based register material for which diachronic corpus data across the varieties in question are available, parliamentary Hansard. Using the records of parliamentary proceedings (verbatim in spirit, if not strictly in practice) allows us to provide representation of speech-based – or "spoken-in-origin" – material in the present study.

Table 1 below (modelled on Table 1.1 in Biber et al. 1999:16) differentiates our five registers in terms of their "situational characteristics". The registers are ordered from left to right according to how closely they represent conversation, with parliamentary dialogues the most conversation-like, and learned writing the least.

Hansard is the official verbatim record of parliamentary proceedings, including questions without notice, debates, and speeches. However, as already suggested, it does not consist of straightforward transcriptions of parliamentary speech, but is rather a highly edited, conventionalised, and traditionally conservative written representation of parliamentary speech (see Mollin 2007; Kruger & Smith 2018). Accordingly, situationally, Hansard has a multi-faceted register profile. On the one hand, it originates typically in spoken language which is interactive and produced extempore in a shared parliamentary situation (debates and questions) or as a premeditated – possibly scripted – mode of monologic speech (speeches). In other respects, however, Hansard exists as a set of written documents not directly interactive or involving specific addressees. Its communicative aim is less broadly personal and reflective of speakers' backgrounds than casual conversation, being focused more narrowly on the presentation of information and arguments pertinent to political issues and policies. The immediate audience for parliamentary proceedings is a set of elected politicians present in Parliament, but also, latterly, members of the public accessing them via live broadcast or the written Hansard records. These audience members can be assumed to have political backgrounds and/or interests, rather than being either specialists in the subject

Table 1. Situational characteristics of the five registers

	Parliamentary dialogues	Parliamentary monologues	Fiction	News	Learned
Mode[*]	Spoken (+"written")	Spoken (+"written")	Written (+"spoken")	Written	Written
Interactive	Yes	No	No (except dialogue)	No	No
Shared situation	Yes/no	Yes/no	No	No	No
Communic-ative aim	Argumentative	Informational	Reading for pleasure	Informational/ evaluation	Informational/ explanatory/ argumentation
Audience	Narrow/wide	Narrow/wide	Wide public	Wide public	Specialist
Dialect domain	National	National	Global/ national	Regional/ national	Global

* (+"written") = "also has features associated with typical writing"; (+"spoken") = "also has features associated with typical speech"

matter (as in the case of academic texts) or merely members of the public at large (as in the case of fiction and newspapers). The dialectal domain of the Australian and British Hansard is national, produced by speakers from across the country.

Fiction, like the Hansard, has a somewhat ambivalent character. While the mode of most works of fiction is written narrative, they also often contain representations of the conversational dialogue of fictional characters. The communicative aim of fiction is typically to provide a pleasurable reading experience by transporting the reader into an imaginary world. The audience is the public in general, diverse, and non-specialist. The dialect domain is typically global, that is "written for an international audience with relatively little influence from the national dialect of the author" (Biber et al. 1999: 16), although of course a national domain may be in evidence.

According to Biber et al. (1999: 9) newspapers "claim a relatively objective presentation of information – albeit, somewhat inevitably, moderated by an evaluative component". They are not designed for specific types of readers, their ostensible objective being to provide up-to-date information about newsworthy events for the general public. The dialect domain for newspapers is "regional/ national", insofar as the news reports that they provide address both events of local regional relevance and those of relevance to readers nation-wide (including international events).

Learned writing, like news, uses a written mode, is non-interactive, and lacks a shared immediate situation. Depending on the academic discipline, its communicative aim may be primarily to inform, to explain or to present an argument. The characteristic that arguably differentiates academic prose most sharply from the other registers is its narrow specialist audience and typically global dialect domain.

Given our interest in regional variation between registers, dialect domain is an important situational characteristic. It is this characteristic that would lead us to predict that news reportage would be the register most likely to yield insights into regional variation: news reportage, by its nature, has a regional and/or national orientation. By contrast academic writing, with its global specialist audience, is the least likely. In-between we have Hansard (whose debates originate in dialogic speech but in whose ultimate form transcribers/editors have a significant role to play), and fiction (whose authors may or may not wish to pursue national themes).

The situational characteristics outlined above may also be seen as drivers behind the distinction made by Mair & Hundt (1999) between "agile" genres (i.e. those that are open to innovation) and "uptight" genres (i.e. those that are prone to the retention of conservative features). In our material, Hansard debates and learned writing might be predicted to represent two extremes on a scale of agility/openness, with Hansard speeches, fiction and news reportage in-between. We would expect the register of Hansard debates – with its customary origins in interactive speech – to be the most agile of the five registers. It is well known that the spoken word tends to be more receptive to, and productive of, linguistic innovation than the written word. At the same time, however, we would expect Hansard debates to be less agile than prototypical (i.e. casual, conversational) speech, the traditional conservatism of Hansard historically providing a degree of resistance to innovation and change. That learned writing should prove to be the most uptight of our three written registers would not be surprising. The essence of (non-popular) academic writing is to represent technicalities and abstractions, a function for which the informality and innovativeness of everyday spoken language are largely irrelevant. Furthermore, the circumstances of its production/dissemination differentiate academic writing from fiction and press reportage: while the latter two registers are subject to competitive market forces – the pressure to sell as many texts as possible – academic writing is largely exempt from the pressure to win large audiences, its appeal being to a smaller audience of like-minded specialists.

4.   The data

Our written data are derived from a suite of parallel Brown family corpora with four target sampling years: 1931, 1961, 1991 and 2006. Our BrE data are derived from Before-LOB (BLOB), Lancaster-Oslo/Bergen (LOB), Freiburg-LOB (FLOB), and British English 06 (BE06), and our AmE data from Before-Brown (B-Brown), Brown, Freiburg-Brown (Frown), and American English 06 (AE06). AusE is represented by AusBrown, recently compiled with the same four target sampling years as the British and American Brown-family corpora, comprising texts published predominantly in Australia, by Australian authors (for details see Collins & Yao 2019; Yao & Collins 2019). Three registers are represented: fiction (from novels, short story collections and literature anthologies); academic writing (from academic journals and books representing humanities, natural sciences, politics/law/education, social/behavioural sciences, technology/engineering, medicine/veterinary science); and press reports (from thirteen newspapers with relatively long and continuous publication records, including *The Sydney Morning Herald, The Newcastle Sun*, and *The Canberra Times*). In its present form AusBrown contains around 720,000 words of running text, 240,000 per register (each of the three registers represented by thirty 2,000-word texts at each of the four time periods).

Our speech-based data are taken from an existing comparable Australian and British diachronic corpus of Hansard materials for the House of Representatives and House of Commons, sampled from the years 1901, 1935, 1965, 1995 and 2015 (see Kruger & Smith 2018; Kruger et al. 2019). The data used in this study were sampled from this corpus: thirty 2,000-word samples of debates and questions without notice (fifteen samples each), and thirty 2,000-word samples of speeches (each by a different speaker) were extracted for each period. Table 2 shows the makeup of our composite diachronic corpus.

Table 2. Corpus data used for the study

	1930s	1960s	1990s	2006–2015	Total words
AusE	AusBrown	AusBrown	AusBrown	AusBrown	720,000
	Aus Hansard	Aus Hansard	Aus Hansard	Aus Hansard	480,000
BrE	BLOB	LOB	FLOB	BE06	720,000
	Br Hansard	Br Hansard	Br Hansard	Br Hansard	480,000
AmE	B-Brown	Brown	Frown	AE06	720,000
Total words	780,000	780,000	780,000	780,000	3,120,000

5.   The grammatical features

In order to ensure comprehensiveness and comparability we included the same list of features used in Yao & Collins (2019), which in turn was compiled from the sets of features identified in previous studies as implicated in grammatical change (e.g. Leech et al. 2009; Aarts, Close, Leech, & Wallis 2013; Collins & Yao 2018); and/or in register variation (e.g. Biber 1988; Biber et al. 1999).

The list comprises 68 grammatical features representing five major grammatical categories:[2]

A. **The noun phrase**: definite articles, indefinite articles, demonstratives, possessive determiners, semi-determiners, quantity determiners, first-person pronouns, second-person pronouns, third-person pronouns, indefinite pronouns, non-personal pronoun *it*, singular nouns, plural nouns, nouns neutral for number, proper nouns, temporal-spatial nouns, adjective+noun sequences, noun+noun sequences, nominalisations, *of*-genitives, *'s*-genitives, *that*-relative clauses, WH-relative clauses, pied piping, present participial WHIZ-deletion relatives, past participial WHIZ-deletion relatives. (26 features)
B. **The verb phrase**: BE as main verb, DO as main verb, HAVE as main verb, present tense, past tense, progressives, perfects, passives, necessity modals, possibility modals, prediction modals, semi-modals, common activity verbs, common communication verbs, common existence verbs, common mental verbs. (16 features)
C. **Adjectival, adverbial and prepositional phrases**: predicative adjectives, common attributive adjectives as classifiers, common attributive adjectives as descriptors, inflectional comparatives and superlatives, phrasal comparatives and superlatives, time adverbs, place adverbs, stance adverbs, linking adverbs, degree adverbs as downtoners, degree adverbs as amplifiers, *-ly* adverbs, non-*ly* adverbs, prepositions. (14 features)
D. **Complementation**: infinitival verb complements, *ing*-participial verb complements, *that*-complement clauses, WH-clauses as object complements. (4 features)
E. **Other phrasal and clausal elements**: clausal coordination, clausal subordination, comparative clauses plus degree complements, phrasal coordination, *not* negation, *no* negation, existential *there*, emphatics. (8 features)

---

2. It is to be noted that there are some differences in the AusBrown frequencies between Collins & Yao (2018) and the present paper that are due to the incomplete state of compilation of AusBrown at the time that the research was being conducted for the earlier paper.

In order to facilitate the retrieval of these features from the corpora, the raw texts were all part-of-speech tagged with CLAWS, using the C7 tagset to maintain consistency with the existing Brown corpora (see http://ucrel.lancs.ac.uk/claws/). The eight BrE and AmE Brown-family corpora were accessed via CQPweb at Lancaster (see https://cqpweb.lancs.ac.uk/).

To calculate frequencies of the features, search routines were first written in regular expressions and input into PowerGREP, a grep software that extracts frequency information from textual data.[3] For the BrE and AmE texts of the Brown family corpora, the search routines were converted into CQP syntax and searches were performed on CQPweb. Raw frequencies retrieved therefrom were then normalised to tokens per 1,000 words (ptw).

In the next section we describe the quantitative aggregate analyses to which the feature frequencies were subjected in order to determine the nature and extent of their diachronic trajectories in the five registers and the three varieties under investigation.

6.   Methodology

6.1   Factor analysis: Identifying dimensions

The 68 annotated linguistic features were first subjected to exploratory factor analysis (EFA), using normalised frequencies ptw as input, in order to identify clusters of linguistic features that we interpret as reflecting an underlying shared functional motivation (as in Biber 1988). We performed the EFA using the package "Psych" (Revelle 2018) in R. To determine the number of factors to be extracted, we used the scree plot in Figure 1. Based on this, we extracted a six-factor solution: six factors have eigenvalues above 1, and the most obvious "break point" also occurs after factor 6.

We carried out the factor analysis using the promax rotation (as is typical for linguistic data, given that features are likely to be correlated; see Biber 1988: 84–85 for discussion), and the minimum residual/ordinary least squares factoring method, which yields results similar to the maximum likelihood method but without assuming multivariate normal distribution (see Revelle 2018: 120 for further explanation). To assess model fit, three values were considered. The root

---

3. For some features, 100% accuracy is not guaranteed by the search routines. In these cases, reliability was sought by comparing the results of the searches performed on a small sample with those derived from manual coding, with the results for any given feature being considered acceptable only when precision and recall were both over 95%.
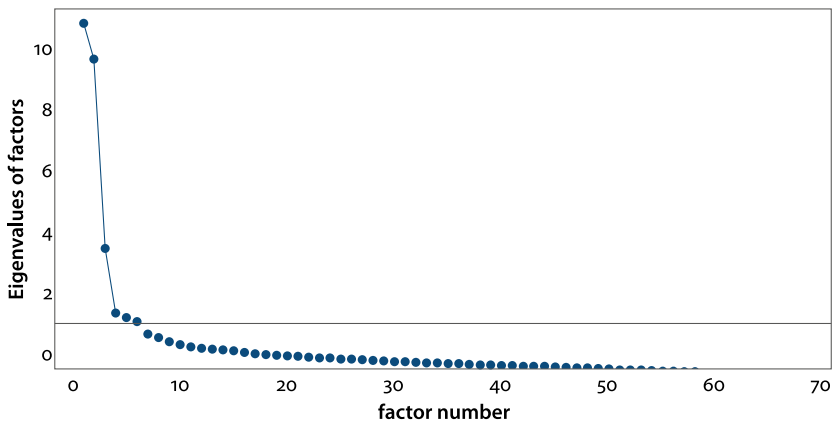
Figure 1. Scree plot

mean square of the residuals (RMSR), which is the square root of the difference between the residuals of the sample covariance matrix and the hypothesised model (reflecting the average residuals for the correlation matrix) is 0.03 (the recommended value is less than 0.08). The Tucker Lewis Index of factoring reliability is 0.746. A value of 0.95 indicates that the model improves the fit by 95% relative to the null model; this model thus improves the fit relative to the null model by approximately 75%. The RMSEA (Root Mean Square Error of Approximation) index is 0.065. Values closer to 0 represent a good fit, and values close to 0.06 are recommended (see Hu & Bentler 1999 for discussion of measures of fit). Based on these values we can conclude that there is a relatively good fit between the model and the observed data.

As in Biber (1988) we used an absolute factor loading of 0.35 and higher as the cut-off point for including a feature in the model. In the discussion of the six factors that follows in Section 7.1, we include items that load on more than one factor in order to characterise the communicative function underpinning the co-occurrence of features on each factor. However, in the subsequent calculation of factor (or dimension) scores for each text, we include each feature only once, on the factor where it has the highest loading.

Subsequently we calculated a dimension score for each text, by adding the standardised scores for each feature with a positive loading on the dimension and subtracting the standardised scores for each feature with a negative loading on the dimension. These dimension scores are used as the dependent variable in the regression analysis.

6.2   Regression analysis: Dimension scores across registers, varieties, and time

In order to determine whether particular registers undergo different patterns of change in the different varieties, we used generalised linear modelling (using the glm function in the "stats" package in R, with default parameters; R Core Team 2019), with the dimension score for each text as a dependent variable. For each of the five registers, we investigated the effect of two independent variables (or predictors) on this dimension score: Variety and Period. We include, in principle, an interaction term, since we are interested precisely in whether the varieties show differential change in each register, over time.

7.   Results

7.1   Factor analysis: Identifying dimensions

In interpreting the results of the factor analysis to identify the underlying dimensions of register variation, a few points need to be kept in mind. First, our model is derived from the input data (our corpus), which means that the particular kinds of registers included in the corpus inevitably shape the dimensions identified. In other words, the dimensions that emerge do so because they discriminate between the various registers included in the study, these all being relatively information-rich, and not prototypically informal and spoken.

Second, in interpreting the dimensions, we follow some of the general principles followed by Biber (1988). Accordingly we interpret the collection of features with positive and negative loadings on each dimension in terms of its contribution to an underlying function. It may be expected that some features demonstrate a less obvious "fit" with an underlying function.

7.1.1   *Dimension 1: Information presentation/elaboration and stance marking under production pressure*

As shown in Table 3, Dimension 1 consists mostly of features with positive loadings: There is only one feature with a negative loading (past participial WHIZ deletion relative clauses, a typically low-frequency feature that is usually associated with crafted writing).

The features with positive loadings can be broadly grouped into three sets. First, there are those associated with the presentation of static information, including BE as main verb, non-personal *it*, singular and plural common nouns, first- and third-person pronouns, definite articles, perfect aspect, existential *there*,

Table 3. Dimension 1 feature loadings[4]

Feature	Loading
BE as main verb	0.783
It non-personal pronoun	0.780
Clausal subordination	0.766
Singular common noun	0.720
Of-genitive	0.693
Demonstratives	0.659
That complement clauses	0.644
Degree adverbs downtoners	0.586
Necessity modals	0.542
WH relative clauses	0.524
Perfect aspect	0.520
(Definite articles)	(0.519)
Degree adverbs amplifiers	0.497
(HAVE as main verb)	(0.495)
Existential *there*	0.479
Quantifying determiners	0.458
that relative clauses	0.440
Prediction modals	0.438
(Third-person pronouns)	(0.434)
(Plural common nouns)	(0.412)
Stance adverbs	0.402
Predicative adjectives	0.396
(First-person pronouns)	(0.376)
Pied-piping relative clauses	0.368
(Total prepositions)	(0.366)
Past participial WHIZ deletion relatives	−0.537

and quantifying determiners. Second, there are features associated with the elaboration of information, but typically in syntactically less complex ways, suggestive

---

4. Bracketed features have a loading above the cut-off point, but have a higher loading on another dimension. See explanation in Section 6.1.

of information elaboration under conditions of cognitive constraint. These include various forms of clausal modification and elaboration: clausal subordination, *that* complement clauses, both WH and *that* relative clauses, and pied-piping relative clauses. *Of*-genitives and predicative adjectives may similarly be seen as more analytical and less syntactically compressed forms of information elaboration. Third, there are features that encode stance: demonstratives, downtoners and amplifiers, necessity and prediction modals, and stance adverbs.

Based on this, we interpret this dimension as reflecting information presentation and elaboration, combined with stance marking, taking place under production pressure (e.g. in spoken rather than written language).

Example (1) exemplifies Dimension 1, highlighting some of the features that characterise texts that have a positive score on this dimension. In this extract the specific presentation and elaboration of information is reflected by (amongst others) the use of BE as main verb and pronoun *it* (all in CAPS), while more analytical forms of elaboration in the form of *of*-genitives (underlined) and relative clauses is also evident (in *italics*). Stance marking is clear in the use of the obligation/necessity modal (in bold).

(1)   A charge is made against a man *that he has trespassed upon <u>the property of the Commonwealth Railways Commissioner</u>*. The case IS then very simple – the burden **should** rest upon the person *who makes the charge to prove IT*. IT has always BEEN <u>the endeavour of the Labour movement</u> to stand for that principle and resist encroachments upon IT *whenever they are made*.

(Australian Hansard: Dialogues, 1930s)

7.1.2   *Dimension 2: Personal-active versus impersonal-informational-compressed discourse*

As shown in Table 4, features with negative loadings on this dimension are mostly clearly associated with highly informational and largely static discourse: passives, noun-noun sequences, nominalisations, definite articles, classifying attributive adjectives, and prepositions. The NP-centred nature of these features is clear, and reflects a high degree of information compression.

The features with positive loadings on this dimension may be seen as expressing two complementary functions. On the one hand, there are features that are clearly verb-focused, and thus reflect a more active, verb-based discourse construction (DO and HAVE as main verbs, semi-modals, *not* negation, single common activity and communication verbs, non-*ly* adverbs, progressive aspect, *-ing* participial complement clauses). Some of these features overlap with the second set of features, which serve to express a personal rather than impersonal style of

discourse (*not* negation, first-, second-, and third-person pronouns, semi-modals, and emphatics).

Based on the above, we interpret this dimension as reflecting a contrast between discourse that is more personal and active (and more "verby") in discourse construction, and that which is more impersonal, informational, and static (and more "nouny") in discourse construction.

Table 4. Dimension 2 feature loadings

Feature	Loading
Not negation	0.784
Second person pronouns	0.731
DO as main verb	0.696
Semi-modals	0.586
First person pronouns	0.586
Indefinite pronouns	0.568
Phrasal coordination	0.549
HAVE as main verb	0.532
WH object complement clauses	0.513
Single common activity verbs	0.472
Third person pronouns	0.455
Non-*ly* adverbs	0.448
Progressive aspect	0.420
(Single common communication verbs)	(0.391)
Possessive determiners	0.388
(Emphatics)	(0.369)
-ing participial complement clauses	0.351
Passives	−0.392
Common attributive adjectives classifiers	−0.414
Noun-noun sequences	−0.519
Nominalisations	−0.580
Definite articles	−0.655
Total prepositions	−0.813

Examples (2) and (3) exemplify Dimension 2 with positive and negative loadings respectively. In (2), the personal, active character of the discourse is encoded

in the use of first- and second-person pronouns, HAVE as main verb, and *not* negation (in **bold**). The use of *-ing* participial complement clauses and progressives (in *italics*), and non-*ly* adverbs (in CAPS), is also evident.

(2) **I** thought of **you** YESTERDAY when **I** stood TOO CLOSE to the tracks and a tram, *keeling* FAST, clipped the tip of my shoulder.
Unlike **you, I** stepped back in time but **I** *was shaking*, because NOW **I** knew what **you** had ALREADY found out: the colossal weight of the thing, its dense rigidity, its utter lack of give.
Patrick was by nature **not** a guest but a host, the kind of person who **had** his own chair and ALWAYS sat in it.   (Australian fiction, 1990s)

In contrast, in (3) these features are nearly absent, and instead there is a high frequency of nominalisations and noun-noun sequences (in CAPS), prepositions (underlined), passives (in **bold**), and definite articles and attributive adjectives (in *italics*).

(3) In *the* 1966 DE BERG INTERVIEW, Wilkes noted that *Australian* literature differed from English 'in *the* sense that there are so many authors on whom no monograph exists, no biography **has been written**, no edition **has been prepared**, and *the* only way this **can be done** properly, I think, is by training people at *the* universities, RESEARCH STUDENTS, to do this kind of work'.
                    (Australian learned writing, 1990s)

7.1.3 *Dimension 3: Unmarked/non-perspectival versus perspectively marked information presentation*

Table 5 shows the feature loadings for Dimension 3. The features with positive loadings on Dimension 3 are, like those on Dimension 2, nearly all noun-based: adjective-noun sequences, singular and plural common nouns, and common nouns neutral for number. These are accompanied by present tense. The difference between the noun-based features on Dimension 3 and Dimension 2 is that those on the latter are associated with high informational density and compression, whereas those on the former are not, being instead more "general-purpose" nouns. In conjunction with present-tense verb marking, these features index, on the positive side, a kind of neutral, highly generalised, unmarked ("bland") discourse construction. On the negative side of this dimension are three features that introduce markedness in terms of polarity and tense, but also in terms of marking stance (in the form of the use of mental verbs, predominantly THINK). These features may also be interpreted as marking perspective, in respect of time, polarity, and viewpoint.

Against this background, we interpret Dimension 3 as setting apart registers in which information presentation is "bland", neutral and unmarked in respect of perspective, and those in which they are perspectively modified.

Table 5. Dimension 3 feature loadings

Feature	Loading
Adjective-noun sequences	0.582
(Singular common nouns)	(0.493)
Plural common nouns	0.485
Present tense	0.437
Common nouns neutral for number	0.390
(Inflected comparatives and superlatives)	(0.354)
No negation	−0.361
Single common mental verbs	−0.376
Past tense	−0.509

Example (4) illustrates a text that is typical of the "bland" all-purpose, non-perspectively marked informational construction at the positive end of this dimension, particularly evident in the use of common nouns (<u>underlined</u>) and adjective-noun sequences (**bold**), combined with present-tense verbs (*italics*).

(4) The <u>law</u> currently *includes* '**lock-out <u>rules</u>**' that *stop* **small <u>businesses</u>** that *elect* out of the **simplified depreciation <u>scheme</u>** from re-entering for five <u>years</u>. To ensure <u>fairness</u> and the **broadest <u>availability</u>** of this <u>measure</u>, this <u>schedule</u> *relaxes* those <u>rules</u>...   (Australian Hansard: monologues, 2000s)

7.1.4 *Dimension 4: Comparison and intensification (without stance marking)*

As shown in Table 6, Dimension 4 includes only features with positive loadings: phrasal and inflected comparatives and superlatives, emphatics and *-ly* adverbs. These features all reflect comparison and intensification, but without the strong stance marking that is evident particularly on Dimension 1.

(5) Significant differences were obtained between the groups on age, with specialists being <u>younger</u>. **Similarly** specialists had <u>less</u> experience and were <u>more</u> likely to be working within the government sector as opposed to private practice. The specialists also had **correspondingly** <u>more</u> confidence and interest in working with <u>older</u> adults.   (Australian learned writing, 2000s)

Table 6. Dimension 4 feature loadings

Feature	Loading
Emphatics	0.695
-ly adverbs	0.569
Phrasal comparatives and superlatives	0.563
(Stance adverbials)	(0.378)
Inflected comparatives and superlatives	0.368

The extract in Example (5) is from a text with a high positive score on Dimension 4, here resulting from the use of *-ly* adverbs (**bold**) and comparatives and superlatives (<u>underlined</u>).

7.1.5 *Dimension 5: Contextualised speech reporting*

The combination of positive features on Dimension 5 (see Table 7) – particularly proper nouns, temporal and spatial nouns, time adverbs, and single common communication verbs – serves primarily to reflect highly contextualised speech reporting, where words uttered at particular times and in particular places are reported.

Table 7. Dimension 5 feature loadings

Feature	Loading
Temporal and spatial nouns	0.610
Proper nouns	0.604
's-genitives	0.506
Time adverbs	0.453
Single common communication verbs	0.432

Example (6) exemplifies Dimension 5 in the high frequency of proper nouns (in **bold**), communication verbs (in italics) and *'s*-genitives (in CAPS).

(6) **Gippsland** regional fire co-ordinator **David Tainsh** *said* there was little likelihood of any significant fire spread over the next few days. "This is a comforting scenario for everyone who has been watching, waiting and worrying about the FIRES' recent behaviour," **Mr Tainsh** *said.*(Australian press writing, 2000s)

7.1.6   *Dimension 6: Descriptive specification of past events*

As there are so few features on Dimension 6, and they generally have fairly low loadings, a coherent interpretation is difficult. Three of the features, however, appear to be related to descriptive elaboration (descriptive attributive adjectives, place adverbs, and non-*ly* adverbs). The functional coherence of these features with clausal coordination and past tense, however, is not immediately clear. We tentatively describe this dimension as reflecting descriptive specifications of things and actions, and specifically activities recounted in the past tense.

Table 8. Dimension 6 feature loadings

Feature	Loading
Common attributive adjectives descriptive	0.493
Place adverbs	0.437
(Past tense)	(0.392)
Clausal coordination	0.384
Non-*ly* adverbs	0.375

Example (7) exemplifies Dimension 6 in the use of past tense (<u>underlined</u>) to recount descriptions of places and events using attributive adjectives (in *italics*), and place adverbs and non-*ly* adverbs (in **bold**).

(7)   I <u>looked</u> **forward** to *gentle* days on *uncrowded* beaches, and a *calm* sea, **still** *warm* enough to swim in. But it <u>was</u> like some *wild* thing <u>had</u> arrived – a *wild* wind stirring everything **up**, blasting **up** my memories so that images from the past <u>would</u> appear right in front of me.   (Australian fiction, 2000s)

7.2   Regression analysis: Registers across varieties and time

For each dimension we begin with a figure presenting a scatterplot of scores across all five registers included in the analysis, by variety and time period. Regression lines for changes in each variety over time are fitted. Full results of the regression analyses for each dimension are provided in the Appendices.

7.2.1   *Dimension 1: Information presentation/elaboration and stance marking under production pressure*

Comparing the scores for Dimension 1 for the five registers in our corpus (see Figure 2 and Appendix A), we find a clear distinction between the two speech-based parliamentary registers, which both score on the high positive end of this dimension, and the three written registers, which all score on the negative end of
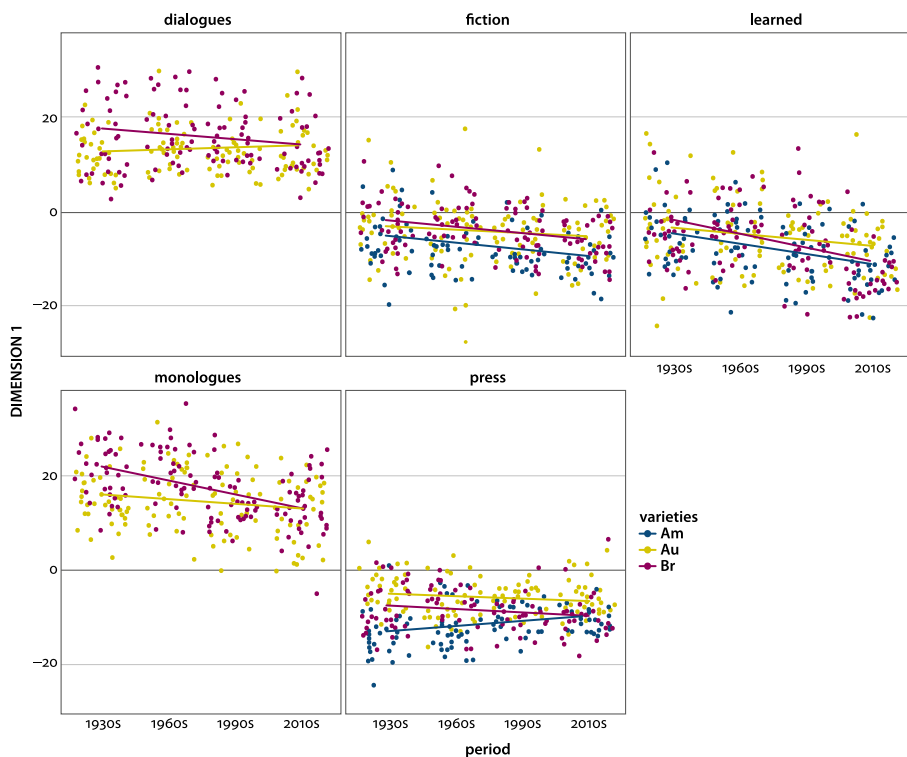
Figure 2. Scatterplot of scores for Dimension 1 across the five registers, with regression lines

the scale. Parliamentary debates and speeches are thus characterised by high frequencies of linguistic features that co-occur to express and elaborate information in a relatively static way, with stance marking present. There is also an element of production constraint evident, given that features with positive loadings are typically syntactically less complex. These characteristics are evident in Example (1), in Section 7.1.

In the case of parliamentary dialogues the two varieties are significantly different, but change over time is not significant and there is no compelling evidence of divergence between BrE and AusE over time. For parliamentary speeches the picture is somewhat more complex: the significant interactions of Variety and Period suggest that BrE and AusE speeches undergo differential change on this dimension with both showing a decline in the dimension score over time, but this is more pronounced for BrE.

In fiction there are no significant varietal differences, but in all three varieties there is a significant decline in dimension score over time, roughly of the same magnitude, with the change appearing to set in after the 1960s. Learned writing

shows a similar trend. There is a significant change over time, with the 2010s scores for all three varieties differing significantly from those of the 1930s. However, this change appears to set in later than for fiction: it is only from the 1990s that there is a significant effect for Period.

The press register shows distinct varietal developments over time. There are significant main effects for Variety and Period, as well as significant interactions between these predictors, which suggests that not only do all three varieties change over time, but they do so in significantly different ways. The 1960s emerges as decisive for AusE: the change sets in here for this variety, but only in the 1990s in BrE. AusE has a consistently less negative score over time, showing only a slight decline. BrE echoes this pattern, but with a more negative score. AmE takes a different trajectory, starting out from a strongly negative score, but then increases over time to converge with BrE by the 2010s.

7.2.2 *Dimension 2: Personal-active versus impersonal-informational-compressed discourse*

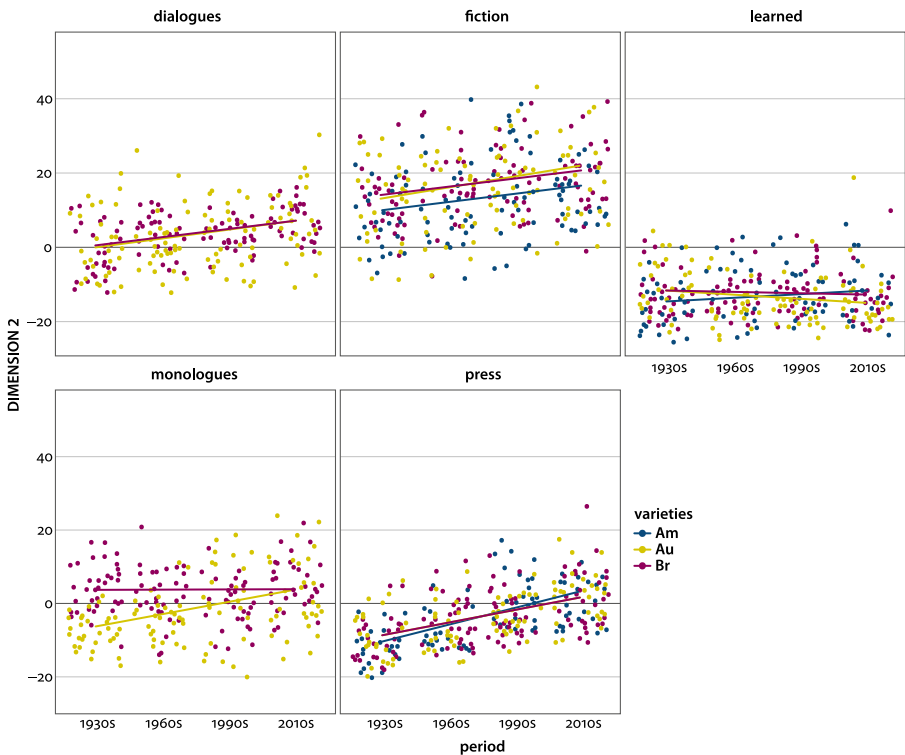**Figure 3.** Scatterplot of scores for Dimension 2 across the five registers, with regression lines

As is evident in Figure 3 and Appendix B, on this dimension, fiction is most distinct from the other registers, with consistently higher scores, although the high degree of dispersion also signals a great deal of individual variability. Fiction thus displays the greatest reliance on discourse presentation that is both personal and active (as also evident in Example (2) in Section 7.1). Parliamentary speeches and debates, as well as press writing have intermediate scores, while learned writing has the most strongly and consistently negative score, indicating that it is most typically characterised as an informational register that presents information in a highly compressed as well as impersonal fashion (as evident in Example (3) in Section 7.1). It is also a register that is well defined and uniformly realised.

Parliamentary dialogue has almost identical scores in BrE and AusE, and identical trajectories of change on this dimension, moving from a dimension score of around 0 to around 7. The regression analysis confirms that there is a significant effect for Period. In other words, parliamentary dialogues in both varieties become more personal and active, and less impersonal-informational-compressed in style, with a significant change from the baseline comparison with the 1930s emerging in the 2010s.

The picture for parliamentary speeches is very different. Significant effects for Period, Variety, and the interaction between these two predictors suggest that speeches in the two varieties change over time on this dimension, but do so in different ways. They start out at distinct points, with AusE parliamentary speeches demonstrating a negative score, and BrE speeches a positive score. Over time, the AusE parliamentary speeches score more positively on this dimension, so that by the 2010s speeches in the two varieties converge on a score just above 0. The change (and the divergence between BrE and AusE) appears to be already underway by the 1960s.

In fiction the differences between varieties are not significant (an effect of the large degree of individual variation). There is, however, a significant change over time for all three varieties. The change sets in after the 1960s, with significant effects for Period identified for the 1990s and the 2010s.

Learned writing demonstrates differential changes over time for the three varieties on this dimension. The regression analysis (with significant effects for Period and Variety, and interactions of these predictors) suggests that all three varieties undergo change (particularly evident after the 1960s). However AusE is significantly different in respect of how this change plays out (with significant main effects for Variety indicating a difference between AusE and AmE, but not AmE and BrE), and interaction effects that show AusE diverging from the other two varieties.

Press writing shows notable changes over time for the three varieties on this dimension, with AusE somewhat different from the other two varieties. There is

a consistent and significant effect for Period that cuts across the three varieties: all three demonstrate an increase in dimension score over time from negative to positive, a change which sets in even before the 1960s. In addition, there is a main effect for Variety, with AusE differing from AmE. There are also interaction effects indicating differential changes over time: primarily for BrE and AusE in the 1990s, when the three varieties appear to significantly diverge.

7.2.3 *Dimension 3: Unmarked/non-perspectival versus perspectively marked information presentation*
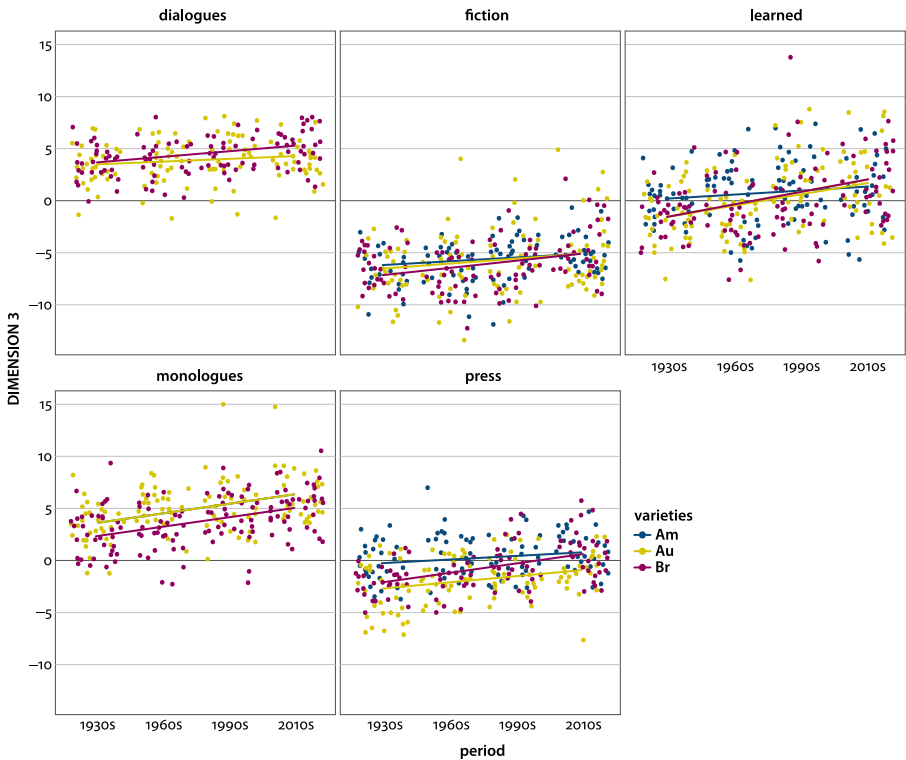
Figure 4. Scatterplot of scores for Dimension 3 across the five registers, with regression lines

The two speech-based registers (parliamentary debates and speeches) have the most clearly positive scores on this dimension. The negative scores for fiction writing place it at the other end, and learned writing and press writing are very close with scores around 0. This dimension therefore essentially serves to distinguish the general-purpose discourse construction of the two speech-based parliamentary registers (see Example (4) in Section 7.1) from the more marked

constructions of fiction in its use of past tense (for narrative reporting), negation, and mental verbs, all of which introduce perspectival variation.

The differences between periods and varieties that emerge for parliamentary dialogues and monologues are minimal, and do not demonstrate clearly interpretable trends. For fiction, too, there are no significant differences between the three varieties, and change over time appears to be marginal – it is only by 2010 that the dimension score for fiction is significantly different from the dimension score for 1930s fiction. If there is register change underway on this dimension towards a less negative score, it is slow and appears to be shared by all three varieties.

For learned writing the results of the regression analysis show that there is a significant main effect for Period – the 1990s data differing significantly from that for the 1930s – but no varietal differences. In contrast, press writing evidences significant varietal variation on this dimension: AmE press writing has the highest scores (consistently slightly positive), followed by BrE (which moves from negative to positive), and then AusE (which despite an increase maintains a negative score). For Period it is only the 2010 data that differs significantly from the 1930s data (indicating a possible slow change). Interactions between Period and Variety are limited.

7.2.4   *Dimension 4: Comparison and intensification (without stance marking)*

As shown in Figure 5 and Appendix D, the register with the most clearly positive scores on this dimension is learned writing, albeit with a notably large dispersion of scores. Also positive – if only marginally – are the scores for parliamentary speeches. We interpret this dimension as reflecting comparison and intensification, but without overt stance marking. The scores suggest that it is learned writing that makes the most extensive use of the resources for expressing these categories (as shown in Example (5)).

Parliamentary dialogues in BrE and AusE are very similar on this dimension, and also do not undergo significant changes. In contrast, the dimension scores for parliamentary speeches show distinctive changes across BrE and AusE: the BrE data show a mild decrease across the four periods, whereas the AusE data show a clear increase, ending in 2010 with a dimension score further to the positive end of the scale than BrE. The regression analysis shows a significant effect for Period (the 2010 data differing significantly from the 1930 data), but the interactions between Variety and Period suggest that this difference is largely accounted for by a change after the 1990s, when the AusE scores increase sharply but those for BrE remain largely constant.

The scores for fiction show no significant changes over time, and the differences between the varieties are slight and not significant. Learned writing shows
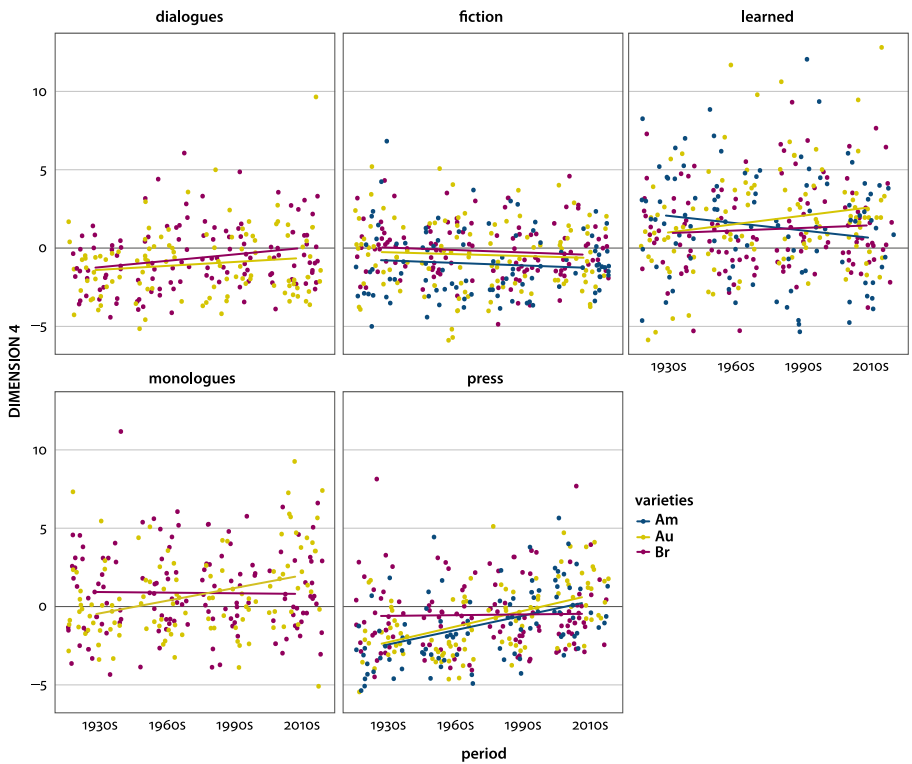
Figure 5. Scatterplot of scores for Dimension 4 across the five registers, with regression lines

not only more change than the other registers, but also distinct patterns of change for the three varieties. While BrE and AusE show a dimension score increase over time (more pronounced for AusE), AmE shows a slight decline. The regression analysis suggests that the changes observed are variety-specific, with no main effect for Period. AusE learned writing is the most different, the only variety with a significant change evident from the 1960s onwards.

Press writing also shows significant and variety-specific register change on this dimension. AusE and AmE press writing have almost identical trajectories of change – from a negative to a positive dimension score – across the entire period investigated, while BrE maintains a slightly negative score. Not surprisingly then it is BrE that differs from the other two varieties in showing no Period effect in the regression analysis, and in showing a significant interaction between Period and Variety.

7.2.5   *Dimension 5: Contextualised speech reporting*
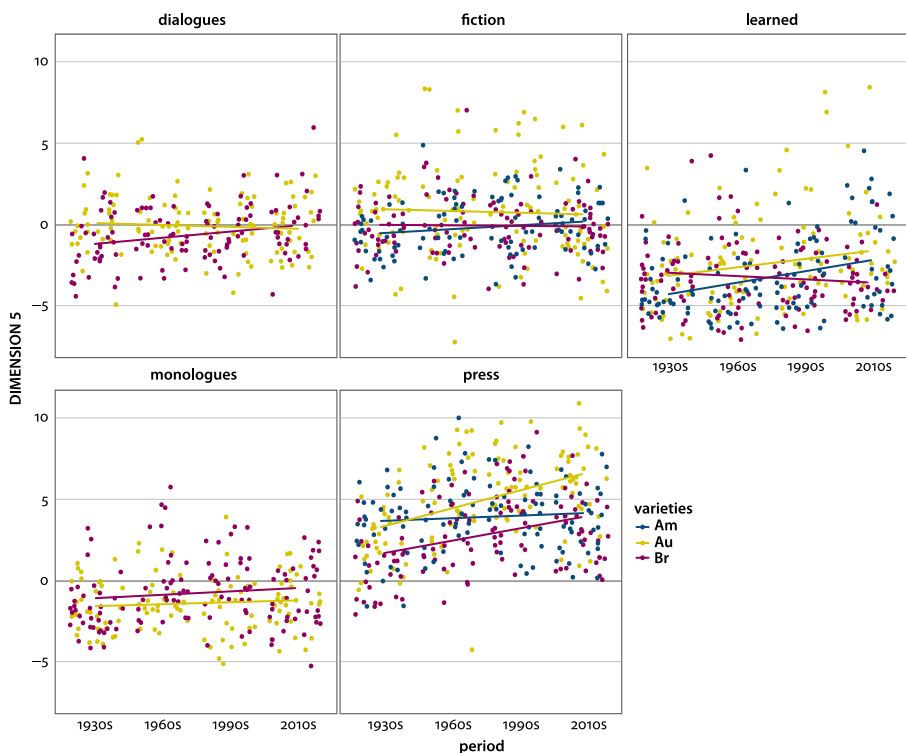
Figure 6. Scatterplot of scores for Dimension 5 across the five registers, with regression lines

As shown in Figure 6 and Appendix E, on this dimension, learned writing and press writing contrast sharply, the former with negative scores, and the latter with positive scores. We interpret these findings as reflecting highly contextualised speech reporting (providing indications of whose speech is being reported, and under what contexts that speech was uttered), a feature that clearly sets press writing apart from the other registers (see also Example (6) in Section 7.1).

With parliamentary dialogues there is a main effect for Variety (with AusE debates scoring higher), but parliamentary speeches do not demonstrate any significant varietal differences, or significant changes over time. In fictional writing, AusE consistently has a significantly higher score than BrE and AmE, but there is little change in dimension score in any of the three varieties.

Learned writing has consistently negative scores across all three varieties. The regression analysis does not find a significant effect for variety (not surprising given the large degree of dispersion of scores in this register). There is, however,

a significant effect for Period, with the 2010 data significantly different from the 1930s. There is also a significant interaction for Variety and Period, with BrE and AusE diverging significantly after the 1990s.

Press writing, the register demonstrating the most dramatic effects on this dimension, shows little change in AmE, whereas BrE and AusE both show an increase in dimension score. BrE press writing starts out in the 1930s with a relatively lower score than the other two varieties, but by 2010 it has increased to converge with AmE. AusE starts from approximately the same baseline as AmE, but enjoys a sharp increase by 2010. The regression analysis shows that changes over time on this dimension are strongly conditioned both by Variety (with BrE, but less so AusE, differing from AmE) and by Period (as already detailed). Interaction effects confirm that AmE is set apart from AusE and BrE on this dimension.

7.2.6   Dimension 6: Descriptive specification of pastevents
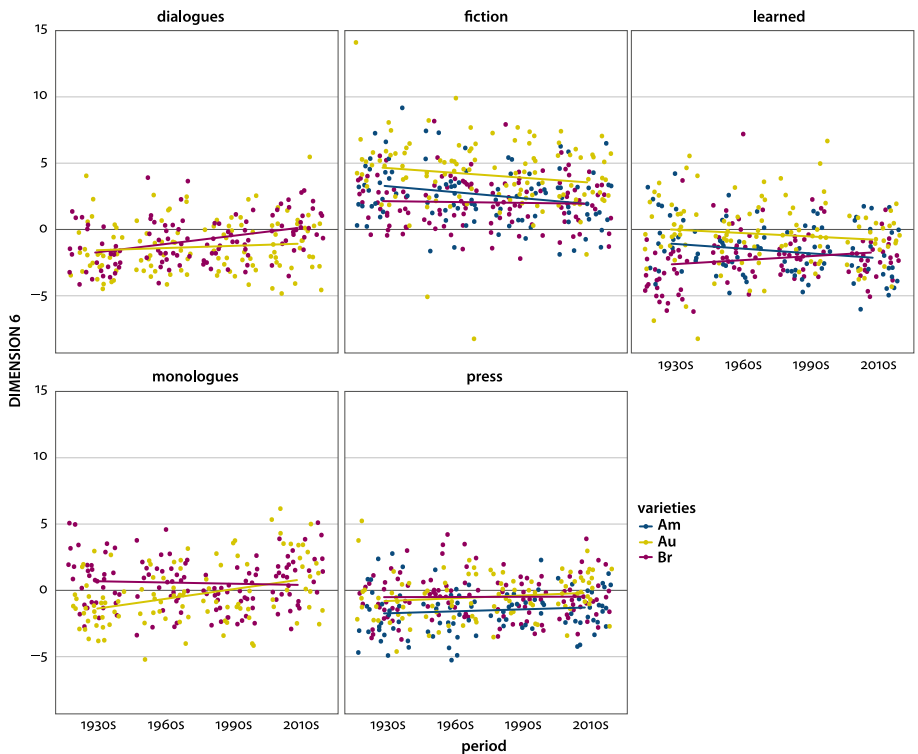
Figure 7. Scatterplot of scores for Dimension 6 across the five registers, with regression lines

On Dimension 6, the only register with a consistently positive score is fiction (see Figure 7 and Appendix F). We interpret this dimension as reflecting descriptive specification of past events, a function clearly associated with fiction, the only register in which description of nouns (people, places and things) regularly combines with the description of activities, recounted in the past tense (typical for narrative fiction in English; see Example (7) in Section 7.1).

The dimension scores for parliamentary debates are very similar for BrE and AusE, but BrE is the only variety to exhibit change – albeit small – over time. This difference is evident in the single significant interaction for Period and Variety in the regression analysis, which confirms that the 2010 British parliamentary debates differ significantly from the Australian debates, and from the 1930s data.

The parliamentary speeches display the opposite pattern, with only AusE undergoing change. The BrE speeches hold steady around 0, while the AusE speeches start with a negative score in the 1930s and then slowly increase over time. The regression analysis shows that there is a significant difference between AusE and BrE speeches on this dimension, and a significant main effect for Period. However, the latter is clearly mediated by Variety, with the BrE and AusE speeches diverging significantly after the 1960s.

In fiction, there is strong evidence of consistent register differences across the three varieties. AusE fiction has a significantly higher score on this dimension, followed by AmE and then BrE. There is some evidence of slow change over time, with a main effect for Period evident for the 2010 data. The lack of interaction effects suggests that all three varieties remain distinct over time, while undergoing the same slow process of change.

Learned writing, like fiction, shows strong evidence of varietal differences on this dimension, with AusE again consistently having the highest score, followed by AmE and BrE. The regression analysis shows that these varietal differences are significant. However, the interaction of Variety and Period additionally shows that BrE learned writing undergoes a significantly different trajectory of change from the other two varieties, with a slow increase from the 1960s which contrasts with a decline in both AusE and AmE from the same period.

In the case of press writing, there are no significant main effects for Variety, and only one significant interaction for Period and Variety, limiting the identification of any overall trends.

8. Summary and interpretation of findings

8.1 Synthesis of findings

In sum, our analysis suggests the following main trends for the five registers and three varieties in our study. **Parliamentary dialogues** show limited varietal differences, as well as limited change over time. It is only on Dimension 1 that BrE and AusE are significantly different, with BrE parliamentary debates scoring higher. Given the large number of features with positive loadings on this dimension, and the fact that these features index a complex interplay of underlying functional constraints, more detailed investigation of individual features would be required to determine the exact nature of this difference between parliamentary debates in the two varieties as far as this dimension is concerned. Beyond this, there is only limited evidence for possible (slow) change over time. Interaction effects (suggesting differential patterns of change over time in the two varieties) are almost completely absent. Parliamentary interactions, thus, are somewhat different in the two varieties, but do not change notably over time.

Parliamentary monologues (speeches) show much more significant evidence of change over time, as well as some evidence of varietal difference. The significant difference between BrE and AusE parliamentary interaction on Dimension 1 is also evident for speeches, but there are also strong effects for Period on Dimensions 2, 3 and 6. On Dimension 3 there are no interaction effects: BrE and AusE parliamentary speeches both increase in the dimension score, suggesting a move towards a more general-purpose, perspectively unmarked discourse style. Considering the features with negative loadings on this dimension (negation, mental verbs), this may also index a move away from a more interactive and more personal style, a finding also highlighted in Kruger et al. (2019). On Dimension 2 and Dimension 6, the strong effects of Period are modulated by Variety, meaning that the two varieties take different trajectories. In the case of both Dimension 2 (personal-active vs impersonal-informational-compressed discourse) and Dimension 6 (descriptive specification of past events), BrE speeches remain largely stable over time, whereas AusE parliamentary speeches start with a more impersonal-informational-compressed and less narrative-descriptive style, and move towards a more personal-active and narrative-descriptive style, converging with BrE by the 2010s. Parliamentary speeches, thus, show stronger varietal differences than parliamentary debates, and change occurs only in AusE speeches, which over time come to approximate BrE speeches in what may be seen as a more involved and less compressed style.

These findings extend those of Kruger et al.'s (2019) comparison of register change in the Australian and British Hansard. In their analysis, no distinction was

made between debates and speeches, and the current analysis suggests that the locus of change towards a more involved, personal and active style (which may be interpreted as a colloquialising tendency) has been in speeches, rather than in debates, with the former thus proving more "agile" in comparison to the latter (somewhat contrary to expectation).

Fiction shows limited evidence of register change over time, but some evidence of varietal differences (that remain largely consistent over time). The only substantial evidence of change over time is on Dimension 1 where there are significant effects for Period; however, as has already been suggested above, interpreting changes and differences on Dimension 1 is complicated by the array of features with positive loadings on this dimension. Varietal differences in fiction emerge on Dimensions 5 and 6. Australian fiction is set apart from British and American fiction in its use of highly contextualised speech reporting (Dimension 5), which occurs consistently more frequently in AusE fiction than in the other two varieties. These findings to some degree echo those of Kruger & Van Rooy (2018), who propose that speech reporting or retelling may be a stylistic feature particularly attuned to local contexts. In respect of the descriptive specification of past events (Dimension 6), all three varieties differ from each other: AusE fiction uses this stylistic resource the most; followed by AmE and then BrE. This suggests that there is a distinct stylistic imprint in respect of the use of adjectival and adverbial modification in narrative registers in the three varieties, which remains consistently distinguishable over time. Fiction, thus, shows limited areas of change over time, but does seem to have some distinct stylistic differences in varieties of English.

Learned writing, surprisingly given its general characterisation as a fairly "global" register, and one that is "uptight" and resistant to change (see Mair & Hundt 1999), shows some evidence of very individualised varietal differences that emerge over time, particularly on Dimension 2 (Personal-active vs impersonal-informational-compressed discourse), Dimension 4 (Comparison and intensification without stance marking), and Dimension 6 (Descriptive specification of past events). There is thus only mild support for our prediction that academic writing, with its global specialist audience, would display an absence of regional variation. As is the case in Hundt, Röthlisberger & Seoane (2019), subtle (but significant) differences between individual varieties are clearly evident, our study showing that these differences extend beyond individual features (such as voice alternation, as analysed in Hundt, Röthlisberger & Seoane 2019) to more general register profiles.

In terms of Dimension 2, the interaction effects for Period and Variety provide strong evidence that it is AusE learned writing that changes in different ways from the other two varieties. The implication of this is that AusE learned writing, unlike

the other two varieties, becomes *more* impersonal, informational and compressed over time. This finding is in line with that of Collins & Yao (2018), who find that in learned writing the mean "colloquiality score" (an index based on sets of colloquial features) shows a distinct drop in AusE learned writing (although they also identify a similar decline in BrE, not evident in our study). In terms of the use of resources for comparison and intensification, the regression analysis again identifies AusE as undergoing distinct developments from the other two varieties, in demonstrating a more pronounced increase in the use of such resources. Lastly, it appears that in the use of resources for descriptive specification (Dimension 6), there are differences in learned writing across the three varieties: all three varieties are distinct (with AusE most descriptive in style, followed by AmE and BrE). However, there is also evidence of change in BrE academic writing, which becomes more descriptive in style over time (as evidenced by the interaction effect for Variety and Period for BrE).

Lastly, as anticipated, **press writing** reflects complex varietal differences as well as patterns of change, across all dimensions with the exception of Dimension 6 (Descriptive specification of past events), which is largely constant across varieties and across time periods. Varietal differences (without much evidence of change) are evident on Dimension 3 (Unmarked/non-perspectival vs perspectivally marked information presentation) and Dimension 5 (Contextualised speech reporting). In terms of the former, AmE newswriting scores highest on Dimension 3, followed by BrE and then AusE. Interpreting these findings would require more detailed analysis of individual features. As for Dimension 5, the statistical analysis shows a significant difference between BrE and AmE newswriting (but not between AusE and AmE) in the use of contextualised speech reporting, and the interaction effects suggest different trajectories of change in the three varieties: AmE remains constant, with AusE and BrE both increasing in their use of this feature (with BrE continuing to use this feature less frequently). The remaining three dimensions all show significant effects for Period and Variety, with significant interactions too. As already suggested, interpreting findings for Dimension 1 is made difficult by the complexity of this dimension. In the case of Dimension 2 (Personal-active vs impersonal-informational-compressed information presentation), the main effect for Period indexes a move towards a more personal and active style of presentation, across varieties. In the case of Dimension 4 (Comparison and intensification without stance marking), the main effect for Period reflects an overall movement towards more use of such resources across the three varieties; however, BrE is distinct in not participating in this change to the same degree as the other two varieties. Overall, therefore, it appears that newswriting has become more personal and subjective in its presentation, but the changes in the press register are also highly idiosyncratic to local contexts, reflecting the

"agility" of the genre in response to local conditions. We thus find considerable divergences between the dimension scores and trajectories of AusE, BrE and AmE press (particularly on Dimensions 1, 3 and 5).

8.2   Observable trends

In this section, we present some further interpretative generalisations of overall trends. In our analyses, we observed five more-or-less recurrent tendencies, namely:

a.   There is a shift in AusE scores from the 1960s
b.   The overall AusE pattern is divergent from the other(s)
c.   The AusE pattern is divergent from the other(s) from the 1990s
d.   The overall BrE pattern is divergent from the other(s)
e.   The overall AmE pattern is divergent from the others

We tabulated their frequencies of occurrence in each register and on each dimension (see Table 9). Since varietal comparisons in the two speech-based registers are based only on two varieties, tendencies are more difficult to identify; in the discussion that follows, we thus mainly focus on the three written registers.

Table 9. Tendencies in the five registers and six dimensions

	D1	D2	D3	D4	D5	D6
Debates	a		c			c
Speeches	a		a, b	c	b	b
Fiction	a, e	a, e	a	a, e	a, b	a, b
Press	a, b	a	e	a, d	a, b, d, e	e
Learned	a, c	a, b	a, b	b, d, e	a, b, d, e	b

Tendency (a) occurs the most commonly. It is the most frequent tendency in fiction, occurring across all six dimensions, and occurs across the majority of dimensions also for press and learned writing. The register change observed in this study for AusE in the period from the 1960s to the 1990s are in line with the findings of Collins & Yao (2018) and Yao & Collins (2019), who found that this period was one critical for the evolution of written AusE. As observed in Section 2, by the 1960s British influence had started to decline in Australia, external British language standards were being discarded and a new-found self-confidence in AusE had begun to emerge (reflected in, for example, the distinctly Australian accents and vocabulary used by the characters in movies, plays and on television).

It is from the 1960s, then, that the post-colonial evolutionary phase described by Schneider (2007) as "endonormativisation" truly begins.

Pattern (c), a minor tendency, may be seen as related to this tendency, suggestive of a continued pattern of differentiation that AusE has undertaken in recent decades, and one that might give some insight into future developments. This pattern is largely restricted to the speech-based parliamentary registers, and largely absent from written registers, suggesting that the formation of distinct stylistic norms for written AusE may not have continued to evolve much beyond the 1960s. While these findings may reflect ongoing change in spoken AusE compared to BrE, they may just as well reflect a differential change in the nature of editorial practices in the two parliaments, as suggested by for example Kruger et al. (2019).

Tendencies (b), (d) and (e) should evidently be interpreted together. Tendency (b) (i.e. AusE is distinct) is particularly strong in learned writing and in speeches (where it occurs across the majority of dimensions), and to a much lesser degree in the other registers. Tendency (d) (i.e. BrE is distinct) occurs on Dimension 4 and Dimension 5 only, and only for press and learned writing. Tendency (e) (i.e. AmE is distinct) is more common across various dimensions for all three written registers. Why is tendency (b) stronger than (e), and (e) stronger than (d) – and why do these tendencies play out differently across different registers? Perhaps the answer lies in the evolution of AusE and AmE, as "Postcolonial Englishes" by the process of endonormativisation, as the two varieties progressively established their own norms and identities via divergence from the parent variety, with AmE completing this process around one century earlier than AusE.

9.   Conclusion

In this paper we have explored grammatical developments from the early 20th to the early 21st century in a suite of corpora representing three written registers and two speech-based registers in Australian, British and American English. A factor analysis of 68 lexicogrammatical features was used to identify six dimensions of register variation. While limited varietal differentiation was noted with **parliamentary dialogues** and **monologues**, it was only in monologues that we found significant change over time. However this finding was limited to AusE, with Australian speeches over time approximating British speeches with a more involved and less compressed style. **Fiction** showed limited evidence of change over time, but there was evidence of varietal differentiation which remained largely consistent over time, most notably in the stronger Australian use than in the other two varieties of the stylistic resources of highly contextualised speech reporting (Dimension 5) and descriptive specification of past events (Dimension 6).

Learned writing, contrary to expectation in view of its reputation as a fairly "global" register, evidenced regional variation over time. Finally it was **press writing** that exhibited the greatest degree of varietal differentiation and change, on all dimensions except Dimension 6. In our data press writing has generally become more personal and subjective in its presentation, but there are considerable differences between the dimension scores and trajectories between AusE, BrE and AmE, suggesting the receptivity or "agility" of this register in response to local contexts.

The findings of the study revealed a number of general diachronic tendencies. The strongest of these, attested across all six dimensions and all five registers – most notably fiction – was a shift in the AusE scores from the 1960s, one that we have suggested to be linked to the progressive rejection by Australians of their historic cultural and economic ties with Britain from this time and the emergent celebration of their culture and language. There was also a milder tendency for AusE to diverge from BrE from the 1990s in the two parliamentary registers, suggestive either of ongoing change in spoken AusE, or in editorial practices, or in both of these (see also Kruger et al. 2019). Another tendency was for AusE to diverge more strongly from the other two varieties than AmE, and for AmE to diverge more strongly from the other two varieties than BrE. This finding we have interpreted from an evolutionary perspective as reflecting the progressive formation by AusE and AmE of their own norms and identities via divergence from their colonial 'parent', BrE, with AmE completing this process almost a century before AusE (see Schneider 2007: 125ff; 291ff). While interpretations of the patterns we observe remain tentative, it is to be hoped that this foray into intervarietal register change will prompt others to embark on research with a wider set of registers and varieties of English.

Corpora

The Hansard monologues and dialogues were extracted from the larger Australian Diachronic Hansard Corpus and British Diachronic Hansard Corpus (Kruger et al. 2019). The corpora are available from Haidee Kotze and Adam Smith.

The written British and American 'Brown family' corpora used in the study (B-BLOB, LOB, FLOB, BE06, B-Brown, Brown, Frown, and AE06) are available online at https://cqpweb .lancs.ac.uk/. AusBrown (see Collins & Yao 2019) is available online at: https://sydneycorpuslab .com/.

References

Aarts, B., Close, J., Leech, G., & Wallis, S. (2013). *The verb phrase in English: Investigating recent language change with corpora*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139060998

Atkinson, D. (2001). Scientific discourse across history: A combined multidimensional/rhetorical analysis of the *Philosophical Transactions of the Royal Society of London*. In S. Conrad & D. Biber (Eds.), *Variation in English: Multidimensional studies* (pp. 45–65). London: Longman.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Biber, D., & Conrad, S. (2009). *Register, genre and style*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511814358

Biber, D., & Finegan, E. (1989). Drift and the evolution of English style: A history of three genres. *Language* 65(3), 487–517. https://doi.org/10.2307/415220

Biber, D., & Finegan, E. (2001). Diachronic relations among speech-based and written registers in English. In S. Conrad & D. Biber (Eds.), *Variation in English: Multidimensional studies* (pp. 66–83). London: Longman.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511920776

Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Blair, D. (1993). Australian English and Australian national identity. In G. Schulz (Ed.), *The languages of Australia* (pp. 62–70). Canberra: Highland Press.

Collins, P. (2015). Diachronic variation in the grammar of Australian English: Corpus-based explorations. In P. Collins (Ed.), *Grammatical change in English world-wide* (pp. 15–42). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.67.02col

Collins, P., & Yao, X. (2018). Colloquialisation and the evolution of Australian English: A cross-varietal and cross-generic study of Australian, British and American English from 1931 to 2006. *English World-Wide* 39(3), 253–277. https://doi.org/10.1075/eww.00014.col

Collins, P., & Yao, X. (2019). AusBrown: A new diachronic corpus of Australian English. *ICAME Journal* 43, 3–29. https://doi.org/10.2478/icame-2019-0001

Desmond, V. (1911). *The awful Australian*. Sydney: Shepherd: J. Andrews and Co.

Fritz, C. 2007a. From Plato to Aristotle – investigating early Australian English. *Australian Journal of Linguistics* 24: 57–97. https://doi.org/10.1080/0726860032000203218

Fritz, C. 2007b. *From English in Australia to Australian English: 1788–1900*. Frankfurt: Peter Lang. https://doi.org/10.3726/978-3-653-01830-1

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Hundt, M., Röthlisberger, M., & Seoane, E. (2019). Predicting voice alternation across academic Englishes. *Corpus Linguistics and Linguistic Theory*. Ahead of print https://doi.org/10.1515/cllt-2017-0050

Hundt, M., Schneider, G., & Seoane, E. (2016). The use of the *be*-passive in academic Englishes: Local vs global usage in an international language. *Corpora* 11(1), 29–61. https://doi.org/10.3366/cor.2016.0084

Kruger, H., & Smith, A. (2018). Colloquialisation versus densification in Australian English: A multidimensional analysis of the Australian Diachronic Hansard Corpus (ADHC). *Australian Journal of Linguistics* 38(3), 293–328.
https://doi.org/10.1080/07268602.2018.1470452

Kruger, H., & Van Rooy, B. (2018). Register variation in written contact varieties of English: A multidimensional analysis. *English World-Wide* 39(2), 214–242.
https://doi.org/10.1075/eww.00011.kru

Kruger, H., Van Rooy, B., & Smith, A. (2019). Register change in the British and Australian Hansard (1901–2015), *Journal of English Linguistics*. 47(3), 183–220.
https://doi.org/10.1177/0075424219857114

Leech, G., Hundt, M., Mair, C., & Smith, N. (2009). *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
https://doi.org/10.1017/CBO9780511642210

Mair, C., & Hundt, M. (1999). "Agile" and "uptight" genres: The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics* 4, 221–242.
https://doi.org/10.1075/ijcl.4.2.02hun

Mollin, S. (2007). The Hansard hazard: Gauging the accuracy of British parliamentary transcripts. *Corpora* 2(2), 187–210. https://doi.org/10.3366/cor.2007.2.2.187

Moore, B. (2008). *Speaking our language: The story of Australian English*. Sydney: Oxford University Press.

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Reeve, J. (1989). Community attitudes to Australian English. In P. Collins & D. Blair (Eds.) *Australian English: The language of a new society* (pp. 111–126). Brisbane: University of Queensland Press.

Revelle, W. (2018). *Psych: Procedures for personality and psychological Research*. Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.12.

Schneider, E. (2007). *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511618901

Seal, G. (1999). *The lingo: Listening to Australian English*. Sydney: UNSW Press.

Westin, I. (2002). *Language change in English newspaper editorials*. Amsterdam: Rodopi.
https://doi.org/10.1163/9789004334007

Wilkes, G. (1978). *A dictionary of Australian colloquialisms*. Sydney: Sydney University Press.

Yao, X., & Collins, P. (2019). Developments in Australian English grammar from 1931 to 2006: A comparative, aggregate approach to dialectal variation and change. *Journal of English Linguistics* 47(2), 120–149. https://doi.org/10.1177/0075424219837337

Appendix A. Regression analyses for Dimension 1, by register
Significance levels: ** < 0.001 * < 0.05

a. Parliamentary dialogues

	Estimate	Std. Error	t-value	Sig.
(Intercept)	12.05	1.06	11.35	**
period1960s	2.54	1.50	1.69	0.09
period1990s	−0.17	1.50	−0.11	0.91
period2010s	1.15	1.50	0.77	0.45
varBr	4.38	1.50	2.92	*
period1960s:varBr	−1.44	2.12	−0.68	0.50
period1990s:varBr	−0.37	2.12	−0.18	0.86
period2010s:varBr	−4.15	2.12	−1.95	0.05

b. Fiction

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−4.99	1.01	−4.92	**
period1960s	−1.88	1.43	−1.31	0.19
period1990s	−3.34	1.43	−2.33	*
period2010s	−4.47	1.43	−3.12	*
varAu	2.60	1.42	1.83	0.07
varBr	1.81	1.43	1.26	0.21
period1960s:varAu	−1.10	2.01	−0.55	0.59
period1990s:varAu	1.06	2.01	0.53	0.60
period2010s:varAu	1.51	2.01	0.75	0.45
period1960s:varBr	3.43	2.03	1.69	0.09
period1990s:varBr	1.62	2.03	0.80	0.43
period2010s:varBr	1.23	2.03	0.61	0.54

c. Learned

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−5.30	1.20	−4.43	**
period1960s	−0.60	1.69	−0.36	0.72
period1990s	−3.13	1.69	−1.85	0.07
period2010s	−6.43	1.69	−3.79	*
varAu	1.53	1.69	0.90	0.37
varBr	0.91	1.69	0.54	0.59
period1960s:varAu	0.65	2.40	0.27	0.79
period1990s:varAu	−0.42	2.41	−0.17	0.86
period2010s:varAu	3.03	2.40	1.26	0.21
period1960s:varBr	1.96	2.40	0.82	0.41
period1990s:varBr	1.83	2.40	0.76	0.45
period2010s:varBr	−2.20	2.40	−0.92	0.36

d. Parliamentary monologues

	Estimate	Std. Error	t-value	Sig.
(Intercept)	14.60	1.08	13.55	**
period1960s	2.32	1.56	1.48	0.14
period1990s	0.19	1.56	0.12	0.90
period2010s	−3.16	1.56	−2.02	*
varBr	5.77	1.50	3.84	**
period1960s:varBr	−2.00	2.18	−0.92	0.36
period1990s:varBr	−6.97	2.17	−3.20	*
period2010s:varBr	−4.20	2.16	−1.95	0.05

e. Press

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−13.69	0.76	−18.10	**
period1960s	0.99	1.06	0.93	0.35
period1990s	3.22	1.06	3.04	*
period2010s	2.95	1.06	2.78	*
varAu	9.49	1.06	8.95	**
varBr	5.23	1.07	4.89	**
period1960s:varAu	−4.46	1.49	−2.10	*
period1990s:varAu	−5.56	1.49	−3.72	**
period2010s:varAu	−5.01	1.49	−3.35	**
period1960s:varBr	−0.07	1.50	−0.05	0.96
period1990s:varBr	−4.44	1.50	−2.96	*
period2010s:varBr	−4.62	1.50	−3.08	*

Appendix B. Regression analyses for Dimension 2, by register
Significance levels: ** < 0.001 * < 0.05

a. Parliamentary dialogues

	Estimate	Std. Error	t-value	Sig.
(Intercept)	0.11	1.29	0.09	0.93
period1960s	1.16	1.83	0.63	0.53
period1990s	2.62	1.83	1.43	0.15
period2010s	7.36	1.83	4.03	**
varBr	−1.42	1.83	−0.78	0.44
period1960s:varBr	4.12	2.58	1.60	0.11
period1990s:varBr	1.58	2.58	0.61	0.54
period2010s:varBr	0.97	2.58	0.37	0.71

b. Fiction

	Estimate	Std. Error	t-value	Sig.
(Intercept)	8.76	1.87	4.67	**
period1960s	2.61	2.65	0.99	0.33
period1990s	8.01	2.65	3.02	*
period2010s	5.15	2.65	1.94	0.05
varAu	3.96	2.63	1.51	0.13
varBr	3.03	2.65	1.14	0.25
period1960s:varAu	−1.23	3.72	−0.33	0.74
period1990s:varAu	−0.79	3.72	−0.21	0.83
period2010s:varAu	2.59	3.72	0.70	0.49
period1960s:varBr	2.67	3.75	0.71	0.48
period1990s:varBr	−1.22	3.75	−0.33	0.74
period2010s:varBr	1.94	3.75	0.52	0.60

c. Learned

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−16.27	1.16	−14.09	**
period1960s	2.95	1.63	1.81	0.07
period1990s	3.76	1.63	2.30	*
period2010s	3.33	1.63	2.04	*
varAu	5.88	1.63	3.60	**
varBr	2.97	1.63	1.82	0.07
period1960s:varAu	−8.24	2.31	−3.57	**
period1990s:varAu	−9.35	2.32	−4.03	**
period2010s:varAu	−7.13	2.31	−3.09	*
period1960s:varBr	−2.42	2.31	−1.05	0.30
period1990s:varBr	−2.22	2.31	−0.96	0.34
period2010s:varBr	−4.18	2.31	−1.81	0.07

d. Parliamentary monologues

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−6.99	1.31	−5.33	**
period1960s	3.41	1.90	1.80	0.07
period1990s	7.40	1.90	3.90	**
period2010s	9.81	1.90	5.17	**
varBr	11.21	1.83	6.14	**
period1960s:varBr	−5.20	2.66	−1.96	0.05
period1990s:varBr	−11.67	2.65	−4.41	**
period2010s:varBr	−8.89	2.63	−3.39	**

e. Press

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−11.26	1.14	−9.86	**
period1960s	4.08	1.60	2.55	*
period1990s	12.06	1.60	7.53	**
period2010s	12.16	1.60	7.60	**
varAu	3.25	1.60	2.03	*
varBr	1.69	1.61	1.05	0.30
period1960s:varAu	−3.22	2.25	−1.43	0.15
period1990s:varAu	−7.29	2.25	−3.23	*
period2010s:varAu	−1.89	2.25	−0.84	0.40
period1960s:varBr	1.16	2.26	0.51	0.61
period1990s:varBr	−5.26	2.26	−2.32	*
period2010s:varBr	−0.68	2.26	−0.30	0.77

Appendix C. Regression analyses for Dimension 3, by register
Significance levels: ** < 0.001 * < 0.05

a. Parliamentary dialogues

	Estimate	Std. Error	t-value	Sig.
(Intercept)	3.41	0.33	10.28	**
period1960s	0.36	0.47	0.77	0.44
period1990s	1.16	0.47	2.47	*
period2010s	0.42	0.47	0.89	0.37
varBr	0.24	0.47	0.52	0.61
period1960s:varBr	−0.27	0.66	−0.41	0.68
period1990s:varBr	−0.19	0.66	−0.29	0.77
period2010s:varBr	1.25	0.66	1.88	0.06

b. Fiction

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−6.25	0.45	−13.93	**
period1960s	0.01	0.63	0.01	0.99
period1990s	1.01	0.63	1.59	0.11
period2010s	1.29	0.63	2.03	*
varAu	−0.29	0.63	−0.46	0.65
varBr	−0.03	0.63	−0.06	0.96
period1960s:varAu	0.30	0.89	0.34	0.74
period1990s:varAu	−0.12	0.89	−0.13	0.90
period2010s:varAu	0.39	0.89	0.44	0.66
period1960s:varBr	−0.97	0.90	−1.08	0.28
period1990s:varBr	−0.74	0.90	−0.82	0.41
period2010s:varBr	0.60	0.90	0.67	0.50

c. Learned

	Estimate	Std. Error	t-value	Sig.
(Intercept)	0.03	0.53	0.06	0.95
period1960s	0.34	0.75	0.46	0.65
period1990s	1.62	0.75	2.16	*
period2010s	0.90	0.75	1.19	0.23
varAu	−1.34	0.75	−1.78	0.08
varBr	−1.11	0.75	−1.48	0.14
period1960s:varAu	−0.24	1.06	−0.23	0.82
period1990s:varAu	1.84	1.06	1.73	0.09
period2010s:varAu	2.05	1.06	1.93	0.05
period1960s:varBr	−0.05	1.06	−0.05	0.10
period1990s:varBr	−0.20	1.06	−0.19	0.85
period2010s:varBr	2.22	1.06	2.09	0.04

d. Parliamentary monologues

	Estimate	Std. Error	t-value	Sig.
(Intercept)	3.40	0.39	8.76	**
period1960s	1.89	0.56	3.36	**
period1990s	2.05	0.56	3.63	**
period2010s	3.03	0.56	5.38	**
varBr	−0.50	0.54	−0.92	0.36
period1960s:varBr	−2.00	0.79	−2.54	*
period1990s:varBr	−0.63	0.78	−0.80	0.42
period2010s:varBr	−0.66	0.78	−0.85	0.39

e. Press

	Estimate	Std. Error	t-value	Sig
(Intercept)	−0.21	0.36	−0.58	0.56
period1960s	0.91	0.51	1.78	0.08
period1990s	0.35	0.51	0.68	0.50
period2010s	1.22	0.51	2.40	*
varAu	−2.68	0.51	−5.25	**
varBr	−1.54	0.51	−2.99	*
period1960s:varAu	0.42	0.72	0.59	0.55
period1990s:varAu	1.37	0.72	1.90	0.06
period2010s:varAu	0.91	0.72	1.27	0.21
period1960s:varBr	−0.61	0.72	−0.85	0.40
period1990s:varBr	1.29	0.72	1.79	0.08
period2010s:varBr	1.44	0.72	2.0	*

Appendix D. Regression analyses for Dimension 4, by register
Significance levels: ** < 0.001 * < 0.05

a. Parliamentary dialogues

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−1.55	0.38	−4.13	**
period1960s	0.61	0.53	1.14	0.26
period1990s	0.63	0.53	1.18	0.24
period2010s	0.80	0.53	1.51	0.13
varBr	0.11	0.53	0.22	0.83
period1960s:varBr	0.15	0.75	0.21	0.84
period1990s:varBr	0.54	0.75	0.71	0.47
period2010s:varBr	0.36	0.75	0.49	0.63

b. Fiction

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−0.54	0.37	−1.47	0.14
period1960s	−0.64	0.52	−1.23	0.22
period1990s	−0.60	0.52	−1.15	0.25
period2010s	−0.50	0.52	−0.97	0.34
varAu	0.58	0.52	1.13	0.26
varBr	0.74	0.52	1.42	0.16
period1960s:varAu	−0.17	0.73	−0.23	0.82
period1990s:varAu	−0.06	0.73	−0.08	0.93
period2010s:varAu	−0.04	0.73	−0.05	0.96
period1960s:varBr	0.36	0.74	0.49	0.63
period1990s:varBr	−0.51	0.74	−0.69	0.49
period2010s:varBr	0.22	0.74	0.29	0.77

c. Learned

	Estimate	Std. Error	t-value	Sig.
(Intercept)	1.92	0.55	3.51	**
period1960s	−0.03	0.77	−0.05	0.96
period1990s	−1.09	0.77	−1.41	0.16
period2010s	−1.14	0.77	−1.48	0.14
varAu	−1.63	0.77	−2.10	*
varBr	−0.76	0.77	−0.98	0.33
period1960s:varAu	1.92	1.10	1.76	0.08
period1990s:varAu	3.38	1.10	3.07	*
period2010s:varAu	2.88	1.10	2.63	*
period1960s:varBr	−0.88	1.10	−0.80	0.42
period1990s:varBr	2.03	1.10	1.86	0.06
period2010s:varBr	1.12	1.10	1.02	0.31

d. Parliamentary monologues

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−0.17	0.44	−0.38	0.70
period1960s	0.35	0.64	0.55	0.58
period1990s	0.34	0.64	0.54	0.59
period2010s	2.61	0.64	4.08	**
varBr	1.03	0.61	1.68	0.10
period1960s:varBr	−0.00	0.89	−0.00	0.10
period1990s:varBr	−0.37	0.89	−0.42	0.68
period2010s:varBr	−2.73	0.88	−3.09	*

e. Press

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−2.34	0.37	−6.33	**
period1960s	0.40	0.52	0.77	0.44
period1990s	1.78	0.52	3.44	**
period2010s	2.61	0.52	5.03	**
varAu	0.39	0.52	0.76	0.45
varBr	1.86	0.52	3.56	**
period1960s:varAu	−0.61	0.73	−0.84	0.40
period1990s:varAu	−0.41	0.73	−0.56	0.57
period2010s:varAu	0.25	0.73	0.35	0.73
period1960s:varBr	−0.80	0.73	−1.09	0.28
period1990s:varBr	−1.60	0.73	−2.19	*
period2010s:varBr	−2.61	0.73	−3.56	**

Appendix E.   Regression analyses for Dimension 5, by register
Significance levels: ** < 0.001 * < 0.05

a. Parliamentary dialogues

	Estimate	Std.Error	t-value	Sig.
(Intercept)	0.24	0.36	0.67	0.51
period1960s	0.41	0.51	0.81	0.42
period1990s	−0.48	0.51	−0.94	0.35
period2010s	−0.36	0.51	−0.70	0.48
varBr	−0.88	0.51	−1.73	0.08
period1960s:varBr	−0.41	0.72	−0.57	0.57
period1990s:varBr	0.64	0.72	0.88	0.38
period2010s:varBr	0.89	0.72	1.24	0.22

b. Fiction

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−1.25	0.41	−3.03	*
period1960s	0.50	0.58	0.86	0.39
period1990s	0.46	0.58	0.79	0.43
period2010s	0.75	0.58	1.28	0.20
varAu	1.19	0.58	2.06	*
varBr	−0.01	0.58	−0.02	0.98
period1960s:varAu	−0.69	0.82	−0.85	0.40
period1990s:varAu	0.57	0.82	0.69	0.49
period2010s:varAu	−1.31	0.82	−1.60	0.11
period1960s:varBr	0.83	0.83	1.01	0.31
period1990s:varBr	−0.23	0.83	−0.28	0.78
period2010s:varBr	−0.50	0.83	−0.60	0.55

c. Learned

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−3.47	0.49	−7.09	**
period1960s	−0.62	0.69	−0.90	0.37
period1990s	−0.21	0.69	−0.31	0.76
period2010s	1.20	0.69	1.73	0.08
varAu	0.16	0.69	0.23	0.82
varBr	0.49	0.69	0.71	0.48
period1960s:varAu	1.19	0.98	1.21	0.23
period1990s:varAu	2.02	0.98	2.05	*
period2010s:varAu	−0.34	0.98	−0.35	0.73
period1960s:varBr	0.63	0.98	0.64	0.52
period1990s:varBr	−0.56	0.98	−0.57	0.57
period2010s:varBr	−2.42	0.98	−2.47	*

d. Parliamentary monologues

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−1.06	0.35	−3.01	*
period1960s	0.69	0.51	1.36	0.17
period1990s	−0.33	0.51	−0.64	0.52
period2010s	−0.17	0.51	−0.33	0.74
varBr	−0.23	0.49	−0.46	0.65
period1960s:varBr	0.83	0.71	1.17	0.24
period1990s:varBr	1.44	0.71	2.03	*
period2010s:varBr	0.46	0.70	0.65	0.52

e. Press

	Estimate	Std.Error	t-value	Sig.
(Intercept)	3.74	0.45	8.23	**
period1960s	1.20	0.64	1.88	0.06
period1990s	0.86	0.64	1.34	0.18
period2010s	0.20	0.64	0.31	0.75
varAu	−0.84	0.64	−1.32	0.19
varBr	−2.50	0.64	−3.89	**
period1960s:varAu	1.64	0.90	1.83	0.07
period1990s:varAu	2.76	0.90	3.07	*
period2010s:varAu	3.58	0.90	3.99	**
period1960s:varBr	1.06	0.90	1.18	0.24
period1990s:varBr	2.94	0.90	3.26	*
period2010s:varBr	1.61	0.90	1.79	0.08

Appendix F. Regression analyses for Dimension 6, by register
Significance levels: ** < 0.001 * < 0.05

a. Parliamentary dialogues

	Estimate	Std. Error	t-value	Sig.
(Intercept)	−2.15	0.35	−6.08	**
period1960s	0.41	0.50	0.82	0.42
period1990s	0.57	0.50	1.14	0.25
period2010s	0.80	0.50	1.60	0.11
varBr	0.02	0.50	0.03	0.97
period1960s:varBr	0.80	0.71	1.14	0.26
period1990s:varBr	0.69	0.71	0.98	0.33
period2010s:varBr	2.01	0.71	2.84	*

b. Fiction

	Estimate	Std.Error	t-value	Sig.
(Intercept)	4.49	0.44	10.26	**
period1960s	−0.78	0.62	−1.26	0.21
period1990s	−0.60	0.62	−0.97	0.33
period2010s	−1.04	0.62	−1.68	0.09
varAu	1.86	0.61	3.02	*
varBr	−1.46	0.62	−2.36	*
period1960s:varAu	−0.40	0.87	−0.46	0.64
period1990s:varAu	−0.32	0.87	−0.37	0.71
period2010s:varAu	−0.13	0.87	−0.15	0.88
period1960s:varBr	1.41	0.88	1.61	0.11
period1990s:varBr	0.56	0.88	0.64	0.52
period2010s:varBr	1.21	0.88	1.38	0.17

c. Learned

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−1.65	0.43	−3.87	**
period1960s	−0.72	0.60	−1.20	0.23
period1990s	−1.05	0.60	−1.74	0.08
period2010s	−0.95	0.60	−1.57	0.12
varAu	1.35	0.60	2.24	*
varBr	−2.35	0.60	−3.88	**
period1960s:varAu	−0.30	0.85	−0.35	0.73
period1990s:varAu	0.50	0.86	0.58	0.56
period2010s:varAu	−0.55	0.85	−0.65	0.52
period1960s:varBr	2.62	0.85	3.06	*
period1990s:varBr	2.56	0.85	3.00	*
period2010s:varBr	1.73	0.85	2.03	*

d. Parliamentary monologues

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−1.64	0.36	−4.55	**
period1960s	0.75	0.52	1.43	0.15
period1990s	0.55	0.52	1.05	0.30
period2010s	2.98	0.52	5.70	**
varBr	2.25	0.50	4.48	**
period1960s:varBr	−1.19	0.73	−1.63	0.11
period1990s:varBr	−1.57	0.73	−2.15	*
period2010s:varBr	−2.36	0.72	−3.27	*

e. Press

	Estimate	Std.Error	t-value	Sig.
(Intercept)	−1.97	0.32	−6.13	**
period1960s	−0.28	0.45	−0.62	0.53
period1990s	0.23	0.45	0.52	0.61
period2010s	0.31	0.45	0.68	0.50
varAu	0.82	0.45	1.82	0.07
varBr	0.69	0.45	1.52	0.13
period1960s:varAu	0.15	0.63	0.24	0.81
period1990s:varAu	−0.20	0.63	−0.31	0.76
period2010s:varAu	1.12	0.63	1.77	0.08
period1960s:varBr	1.55	0.64	2.43	*
period1990s:varBr	0.12	0.64	0.18	0.86
period2010s:varBr	0.49	0.64	0.77	0.45

Address for correspondence

Peter Collins
School of Humanities and Languages
University of NSW
NSW 2052
Australia

p.collins@unsw.edu.au
https://orcid.org/0000-0002-3510-3398

Co-author information

Minna Korhonen
Department of Linguistics
Macquarie University

minna.korhonen@mq.edu.au

Haidee Kotze
Department of Languages, Literature and
Communication
Utrecht University

h.kotze@uu.nl

Adam Smith
Department of Linguistics
Macquarie University

adam.smith@mq.edu.au

Xinyue Yao
School of Foreign Languages
Renmin University of China

xinyue.yao@ruc.edu.cn

Publication history