The Australianization of English

Clemens Fritz (Regensburg)

This paper deals with the question of the history of Australian English in the late eighteenth and nineteenth centuries. It delineates the path from English in Australia to Australian English, looking at specific relevant historical periods. Internal developments and external influences are in focus and Australian English (AusE) is placed in the context of world Englishes today.

The paper is based on the doctoral thesis: *From English in Australia to Australian English 1788-1900* (Fritz 2005).

1. What is Australian English?

Australian English is a fully-fledged variety of English. It has been codified in dictionaries (*The Macquarie* 1981 and 1997, the *Australian National Dictionary* 1988), usage guides (Peters 1995), even grammars (Jones 2001). But what is Australian about Australian English?

First and above all, AusE is a variety of English. It shares most of the lexicon, phonology and morphosyntax with other varieties. Words, idioms, spelling and phonology are the areas where varieties of English differ most from each other; some differences are categorical, others lie in different frequencies of certain items. In the area of morphosyntax variation is less pronounced, but it is still there. What distinguishes AusE from other varieties, e.g. American English (AmE), English English (EngE) and Irish English (IrE), are different choices within the paradigm that, as a whole, make up English.

The differences between these varieties of English are thus the result of unique settings and subsequent historical developments. Contact with indigenous languages, a peculiar mix of immigrants and the environment the colonists found themselves in determined the face of colonial Englishes.

2. Situation at the foundation of the colony

In 1788 the first permanent English colony was founded on Australian soil. The colonists were no ordinary settlers but convicts and their overseers. Altogether there were 717 convicts, 180 of which were women, 191 marines and 19 officers.

At that time the varieties of EngE and AmE had a much greater common core than today. Many features that today are seen as either American or British did not yet distinguish them. This is very obvious in the area of spelling where an American could easily spell *colour* and his English counterpart *color*. The American Revolution and the range of Webster's books, his *American Spelling Book* (1800) and his *American English Dictionary* (1828), gradually furthered greater linguistic independence of the colonial variety.

This openness also meant that Australians had a greater amount of variables to choose from and they did. The fact that they sometimes chose variants that are considered AmE today has often led to the wrong conclusion that AusE has become 'Americanized'. Of course, AmE is and always has been a source for AusE, but these external influences have to be distinguished from the selection from inherent alternatives.

These 'inherent alternatives' were enhanced by the varied origins of the immigrants. Speakers from England, north and south, east and west, met speakers from Wales, Scotland, Ireland, America, South Africa, India and so on. This mix was unique and the result was unique, too. Furthermore, the settlers literally had to come to terms with the exceptional Australian continent with its strange fauna, flora and features of landscape. Language was an important tool to familiarize and tame nature. Finally, there are the languages of the Aborigines which also contributed to what later came to be known as Australian English.

3. Sources of early English in Australia

Linguists observe language. While this is comparatively easy for investigations of present-day use, historical linguistics is restricted by the availability and quality of sources. All these sources are written by people who were able to write and the survival of the sources is often determined by chance. These qualifications show the restrictions of historical linguistics. The sources that can be investigated were written by more or less educated persons. For most periods this excludes the greater part of the population; only in the nineteenth century did literacy become a skill of the majority. Even if the sources report spoken language, e.g. in parliamentary debates, the writing substantially changed the original form. There can be a cline from more literary to more spoken styles, but no text can claim to be an authentic depiction of spoken language.

For the historical study of Australian I have collected millions of words from thousands of sources. Altogether there are more than ten million words of English in Australia 1788-1900 which are complemented by contemporary text collections of British English (BrE; 950,000 words), IrE (47,000 words), English in New Zealand (29,000 words), English in South Africa (34,000) English in Canada (33,000) and three Mark Twain novels (305,000).

Corpus linguistics is the art of making best use of written data. A corpus is a principled collection of texts designed for a specific purpose; therefore a corpus is much more than a mere mass of texts. The difference can most easily be shown by using an example. If I can find 800 instances of the word *dingo* in ten million words of early English in Australia, then I can state that the word existed. I can also look at the individual instances and ascertain its range of meanings and uses. On the other hand, if I have a corpus of early English in Australia controlled for periods, registers and text types I can state that the term *dingo* became more frequent in the course of the nineteenth century, eventually replacing the phrase *native dog*. I can also note if it was used earliest in literary writings or in private letters, more by men or by women, etc. Although the designing of a corpus naturally leads to a much reduced number of words, the explanatory power of the results rises considerably.

For the thesis *From English in Australia to Australian English 1788-1900* I designed the COrpus of Oz Early English (COOEE) consisting of two million words.

The two main design principles were: (1) there should be an equal number of words in each of the historical periods defined for the study (500,000 words each) and (2) there should be an equal number of words from each register within each period. The four periods defined for COOEE were:

Period 1 (1788-1825): This period starts with the foundation of a penal settlement in Australia. Demographically, convicts, emancipists, i.e. convicts whose term of sentence had ended, and their children are dominant throughout. There is a fairly limited settlement area around the ports of the later colonies of NSW, Tasmania and Norfolk Island. Outside contact was limited, mainly due to trade monopolies.

Period 2 (1826-1850): A rapid pastoral expansion takes place. Inland Australia is surveyed and land illegally seized by large numbers of squatters. Assisted immigration schemes are started and bring large numbers of free settlers to Australia. Transportation ends in the eastern colonies and the convict element of society shrinks quickly.

Period 3 (1851-75): Political emancipation comes to the eastern colonies. The Australian gold rushes bring about massive immigration and migrations. Demographically, politically, economically and linguistically Australia develops in leaps and bounds.

Period 4 (1876-1900): In the last period Australia matures considerably. The native-born Australians clearly outnumber the non-natives and their dialects become dominant. Growing urbanization and changes in the economy mark the beginning of a new era. The bush and Australia's convict past are celebrated in many literary works, but the shift to urban life is already visible. The period ends with the political independence of Australia through the foundation of the Commonwealth of Australia in 1901.

Four registers were defined for COOEE: the Speech-based Register (SB), the Private Written Register (PrW), the Public Written Register (PcW) and the register of Government English (GE). Each register contains the same number of words in each period.

Speech-based	15%	=	75,000
Private Written	35%	=	175,000
Public Written	40%	=	200,000
Government English	10%	=	50,000

The registers are subdivided into several text types. These are plays, minutes and speeches in SB, private letters and diaries in PrW, memoirs, newspapers, narratives, official letters, reports and verse in PcW and imperial correspondence, legal texts and petitions and proclamations in GE.

Each text was carefully edited and codified according to the following criteria:

- information about the author and the addressees: name, year of birth, gender, country/region of origin, social status, year of arrival in Australia, gender, status and abode of addressee (if applicable)

- information about the text: year of writing, place of writing, register, text type, number of words, source of text.

The question remains to what extent COOEE can be seen as representative of early English in Australia. The easy answer to that question is: it is not representative. No corpus can claim to adequately represent the language of an entire country or period. But the design of a corpus can try to ensure that the results are relevant. Numbers also make up an important part of the argument. If a structure is very frequent in COOEE, then it certainly was not marginal in nineteenth century Australia as a whole. And if a structure cannot be found in COOEE, then it cannot have been a dominant feature overall.

4. The Australianization of Spelling

As already mentioned above, spelling is an area where AusE had to find its own choices with a number of variables, because these had not yet been codified sufficiently in either BrE or AmE. Consistent spelling was required in public texts and public text production started early. The first newspaper was the *Sydney Gazette* (1803); and before that numerous broadsides, pamphlets and announcements had already been printed in Australia. Of course, British books and newspapers easily outnumbered Australian and American ones in the beginning, but local production became dominant by the middle of the century. As regards external influences, both Johnson's (1755) and Webster's (1828) dictionaries were well-known by nineteenth century Australians.

The following table shows some of the variables in question. Not all of these variables are fully opposed standards. Some are standardized in Britain, but variable in the US, e.g. *ae/oe* digraph retention and L-doubling. Others are standardized in the US, but variable in Britain, like <ise>/<ize> and <dg>/<dge>.

Table 1: Examples of spelling variables distinguishing present-day AmE from BrE (adapted from Sigley 1999:7)

Variable	AmE variant	BrE variant
<ae> and <oe> reduction	*e.g. anemia/fetus*	*e.g. anaemia/foetus*
<dg>/<dge>	*e.g. judgment*	*e.g. judgement*
<i>/<e> with certain words	*inquire/inquiry/insure*	*enquire/enquiry/ensure*
<ense>/<ence> nouns	*e.g. defense*	*e.g. defence*
<er>/<re>	*e.g. center*	*e.g. centre*
<ise>/<ize> and <lyze>/<lyse>	*e.g. criticize, analyze*	*e.g. criticise, analyse*
<or>/<our> and /<oul>	*e.g. color/mold*	*e.g. colour/mould*
<s> and <l> doubling	*e.g. focussed/traveler*	*e.g. focussed/traveller*

The most conspicuous variables are *center/centre, criticize/criticise* and *color/colour*. Their history in nineteenth century Australia is sketched here.

Centre vs. Center

The above is only an example of a larger set of words. The words in question are: *calibre, centre, fibre, lustre, meagre, mitre, ochre, reconnoitre, sabre, sceptre, sepulchre, sombre, spectre* and *theatre*. The frequencies in COOEE, in the British reference texts (Ref-BrE) and in the Mark Twain novels (Ref-Twain) are given below.

Table 2: Frequencies of <re>/<er> in COOEE, Ref-BrE and Ref-Twain and Johnson vs. Webster

Lexeme	COOEE	Ref-BrE	Ref-Twain	Johnson 1836	Webster 1828
calibre/caliber	4/0	1/0	0/0	<er>	<er>
centre/center	98/3	28/10	2/13	<re>	<er>
fibre/fiber	7/0	23/1	0/0	<re>	<er>
lustre/luster	13/0	4/0	0/0	<re>	-
meagre/meager	8/0	2/0	1/0	<er>	<er>
mitre/miter	2/0	0/0	0/0	<re>	-
ochre/ocher	4/0	4/0	0/0	<re>	<er>
reconnoitre/reconnoiter	10/0	3/0	0/0	-	<er>
sabre/saber	10/0	2/0	0/0	<re>	<er>/<re>
sceptre/scepter	6/0	3/0	0/1	<re>	<er>
sepulchre/sepulcher	3/0	2/0	0/0	<re>	<er>
sombre/somber	28/0	6/0	2/1	-	<er>/<re>
spectre/specter	9/0	3/0	0/1	<re>	<er>
theatre/theater	51/0	7/0	4/4	<re>	<er>/<re>
Total <re>	253	88	9	10	3
Total <er>	3	11	20	2	12
% of <re>	98.8	88.9	31.0	83.3	20.0

It is remarkable that early Australians were much more advanced in regularizing the <re>/<er> variable than their contemporaries in Britain and the US. The difference is significant for all comparisons. A look at the 1836 edition of Johnson's and the 1828 edition of Webster's dictionaries shows that both were still divided in their opinion.

There are three instances of *center* in COOEE, a very small number indeed. Two come from the first period and were written by British immigrants, and the third from the last period, probably written by an Australian. The first two could be a remnant of British spelling traditions, the last a simple performance error.

Obviously what Australians did here was to eradicate certain options to achieve greater orthographic consistency. There was already a tendency in BrE and AmE towards <re> and <er> respectively, but Australians took the apparent direction of change in BrE and completed it. This is an instance of colonial innovation where the oversea's variety is more modern than the home model.

The next table compares the findings for COOEE with those from present-day Englishes. The data from Australian websites were collected via *Google*. LOB and FLOB are one million word corpora of 1961 and 1986 BrE, Brown and Frown are one million word corpora of 1961 and 1986 AmE. The LOB, FLOB, Brown and Frown data are taken from Sigley (1997).

Table 3: Frequencies of <re>/<er> in COOEE and present-day varieties

Variable	COOEE	Australian websites	LOB	FLOB	Brown	Frown
<re>	253	2,552,415	417	448	53	24
<er>	3	258,070	14	14	535	533
Total	256	2,810,485	431	462	588	557
% of <re>	98.8	90.8	96.8	97.0	9.0	4.3

A look at the data reveals that Australians have become much more relaxed about using <er>, which is probably also due to the comparative informal genre of internet websites. On the other hand, BrE and AmE have increasingly narrowed their options. Again, Australia is on its own here.

Criticise vs. Critiize

This section explores the variable <ize>/<ise> which comes up in words like advertise, apologise, Australianise, authorise, characterise, civilise, colonise, familiarise, fertilise, generalise, harmonise, legalise, mobilise, organise, realise, recognise, scrutinise, sympathise and utilise.

Table 4: Frequencies of <ise>/<ize> in COOEE over time and some reference texts

Variable	COOEE Period 1	COOEE Period 2	COOEE Period 3	COOEE Period 4	Ref-BrE	Ref-IrE	Ref-Twain
<ise>	62	129	140	250	136	2	14
<ize>	106	108	100	86	74	2	64
Total	168	237	240	336	210	4	78
% of <ise>	36.9	54.4	58.3	74.4	64.8	50.0	17.9

There is a consistent development towards increasing <ise> in COOEE. Whereas in the first period <ize> is chosen in ca. two-thirds of all possible cases, this is completely reversed later. Compared to the reference texts, COOEE becomes more consistent than either Ref-BrE or Ref-IrE but remains less consistent than Ref-Twain.

The question is who is responsible for this development in COOEE? A look at the origins of the writers answers this. British writers use <ise> only in 46.9 per cent of all possible cases, for the Irish the number is already 62.8 and for Australians it is 72.6! It seems that the British immigrants were a retarding element not only preserving the variability they knew from home but actually increasing it (Ref-BrE = 64.8 per cent; the British in COOEE = 46.9). The Irish, on the other hand, developed in the direction of using <ise> more and more, thereby following the general Australian trend (Ref-IrE = 50 per cent; the Irish in COOEE = 62.8). The native-born Australians set the new standard, again surpassing contemporary BrE preferences.

Table 5: Frequencies of <ise>/<ize> in COOEE, in present-day varieties (from Sigley 1999:9) and the web

Variable	COOEE	Australian websites	LOB	FLOB	Brown	Frown
<ise>	581	2,067,930	836	1070	12	13
<ize>	400	205,208	523	604	1713	1908
Total	981	2,273,138	1359	1674	1725	1921
% of <ise>	59.2	91.0	61.5	63.9	0.7	0.7

It is interesting to compare COOEE with present-day varieties. The data from Australian websites indicate that Australians have increasingly standardised <ise>; nineteenth century traditions have not only been preserved, they have been developed into a new Australian spelling standard. BrE seemingly remained stable;

the numbers for Ref-BrE (64.8), LOB (61.5) and FLOB (63.9) are very similar. British writers have accepted that words like *civilise* can have two acceptable spellings. As for <re>/<er>, AmE is least tolerant of variation. This has also influenced other sets of words which, historically speaking, should not have <ize> at all, for instance *advertize, apprize, comprize* and *merchandize*.

Colour vs. Color

This section is the most colourful since the choice of spelling is often lexicalized. A whole number of sociolinguistic criteria have to be looked at to explain certain orthographic choices. Examples of variable words are: *behaviour, colour, endeavour, favour, honour, labour, neighbour, rumour, splendour, valour* and *vigour*.

For Johnson (1836) and Webster (1828) there was no choice. Each already prescribed either <our> or <or> and they are still followed today. Obstinate Australia, however, refused to follow where Britain or America wanted to lead.

Table 6: Frequencies of <our>/<or> over time in COOEE and in reference texts

Variable	COOEE Period 1	COOEE Period 2	COOEE Period 3	COOEE Period 4	Ref-BrE	Ref-Twain
<our>	710	539	429	458	1208	3
<or>	152	131	35	125	25	174
Total	862	670	464	583	1233	177
% of <our>	82	80	92	79	98	2

Spelling in early Australia was much more variable than elsewhere. It is clearly much more influenced by the British than by the American model. Consequently there is a rise towards greater consistency in the direction of <our> between Periods 1-3. It comes as a great surprise that this development is not only stalled but even reversed in Period 4! Something must have happened that stabilized the use of <or> in nineteenth century Australia.

A cross-analysis of period and origin of author showed that the Irish and British writers in COOEE continuously moved towards exclusive use of <our>; in the fourth period the British used it in 95.0% of all possible cases, the Irish did so in 97.6%. Australians choose to differ in dropping the frequency of <our> significantly from 98.5% in Period 3 to only 72.3% in Period 4.

Use of <or> in COOEE is also higher with males and the upper class than with females and lower class writers.

A comparison of COOEE with present-day Englishes is shown in the table below.

Table 7: Frequencies of <our>/<or> in COOEE, in present-day varieties (from Sigley 1999:9) and the web

Variable	COOEE	Australian websites	LOB	FLOB	Brown	Frown
<our>	2136	2,155,850	1394	1123	31	33
<or>	443	552,994	10	10	1425	1331
Total	2579	2,708,844	1404	1133	1456	1364

<or>/<our>	83	80	99	99	2	2

Remarkably the variability in nineteenth century English in Australia shows up on today's Australian websites, too. AusE has retained this feature despite the fact that BrE and AmE both have almost done away with any variability here.

If we look at the individual words under investigation in COOEE great differences show up. Many do not show an <our>/<or> variation at all, but consistently spell <our>. A number of words show an orthographic consistency of more than 90 per cent. These are *colour* (98.9), *endeavour* (95.7), *harbour* (99.3), *humour* (98.0) and *neighbour* (94.7). For these words <or> is clearly a very disvaloured variable. This leaves a small set of words where orthographic variability was indeed a factor worthwhile a closer look, namely *favour* (73.8 per cent), *fervour* (77.8), *honour* (48.1), *labour* (80.2), *splendour* (85.0) and *vigour* (86.3).

Honour is the only word where <or> is actually more frequent than <our>. But even here <or> is declining considerably in the course of the eighteenth and nineteenth centuries. An exception is *honorable* which significantly rises in frequency in Period 4 and is then seven times more frequent than *honourable*.

Regional Variation and Codification

Compared to BrE and AmE, AusE is very uniform, a fact much discussed in the linguistic community. But there are regionalisms which are increasingly noted and investigated. Fritz (fc.) has investigated the spelling of the Hansard staff that took down the proceedings of the Australian federation debates of the 1890s.

Table 8: Spelling in the Hansard protocols of various federation debates

Variable	Adelaide 1897	Melbourne 1890	Melbourne 1898	Sydney 1891	Sydney 1897
<re>/<ere>	<re> (100%)	<re> (100%)	<re> (100%)	<re> (95.7%)	<re> (100%)
<our>/<or>	<or> (97.9%)	<our> (100%)	<our> (96.1%)	<our> (85.1%)	<our> (91.7%)
<ise>/<ize>	<ise> (96.8%)	<ise> (92.9%)	<ise> (99.3%)	<ise> (99.4%)	<ise> (99.4%)

The results show that the spelling guidelines used by the clerks of the three parliaments in Adelaide, Melbourne and Sydney differed greatly from each other. This also means that there were guidelines and that they were followed.

Regionalized spelling choices can also be found in Australia today despite endo-normative efforts by the Australian government's *Style Manual* (sixth edition in 2002) and usage guides like the already mentioned Peters (1995) and the *Macquarie Dictionary* (1981, 1997). For example the education departments in Victoria and South Australia in 1987 prescribe <ise>, whereas New South Wales and Queensland also allow <ize> (Peters 1995:406f).

5. Conclusion

English has an Australian variant. During the first century after settlement stable accents and sociolects developed. Phonological, lexical, orthographic and morphosyntactic choices increasingly depended on local norms. This was certainly not a conscious process. In fact an Australian variant of English was viewed with deep suspicion by language purists; often they commented on the language of the Australian youths as 'detestable'.

Leitner (2004) has delineated the process that made AusE 'acceptable' in the eyes of the upper and middle class. This took another century. Now AusE has become an epicentre of English of its own.

6. References

Ausinfo, 2002, *Style Manual for Authors, Editors and Printers*, Forrutude Valley: John Wiley & Sons.

Delbridge, Arthur, ed, 1981, *The Macquarie Dictionary*, Sydney: Griffin Press.

Fritz, Clemens, 2005, *From English in Australia to Australian English 1788-1900*, unpublished PhD thesis, Freie Universitat Berlin.

Fritz, Clemens, fc., "The conventions' spelling conventions: Regional variation in nineteenth century Australian spelling", Proceedings of ICAME 24, Verona 2004.

Johnson, Samuel, 1755 *A Dictionary of the English Language*, London: Westley & Davis (1836).

Jones, Alex J., 2001, *Australian English Grammar*, Uglebe, NSW: Wild & Woolley.

Leitner, Gerhard, 2004, *Australia's Many Voices I. Australian English - the National Language*, Berlin: Mouton de Gruyter.

Peters, Pam, 1995, *The Cambridge Australian English Style Guide*, Cambridge: Cambridge UP.

Ramson, William, ed, 1988, *The Australian National Dictionary. A Dictionary of Australianisms on Historical Principles*, Melbourne: Oxford UP.

Sigley, Robert, 1999, "Are we still under England's spell?", *Te Reo* 42: 3-19.

Webster, Noah 1800, *The American Spelling Book*, Wilmington, Bonsal & Niles.

Webster, Noah, 1828, *American English Dictionary*, Springfield (MA): G. and C. Merriam.