

FAVORING AMERICANISMS? -OR/-OUR SPELLINGS IN EARLY ENGLISH IN AUSTRALIA - A CORPUS BASED APPROACH

1. Building a Historical Corpus of Australian English

1.1 Data Collection

1.1.1 WHAT SOURCES CAN BE USED?

The early instances of English as spoken or written on Australian soil pose major theoretical problems for the study of early forms of AusE. For example, in how far can Watkin Tench's two accounts, from 1789 and 1793, be said to be Australian? After all, he had been to Australia only for a very short time when he wrote it. There can only be one answer. His books are instances of early English in Australia. They may contain features that contribute to the formation of AusE, but they are certainly not AusE.

Another problem encountered can be illustrated by the career of William Charles Wentworth. He was one of the first children born in the Antipodes (in 1790 while his mother, a former convict, was *en route* from Sydney to Norfolk Island) and a prolific writer. This should qualify him as a first class source on the beginnings of the English language in Australia. But when we learn that very soon after his birth he was brought to England and educated there, we have to rethink our evaluation. He returned in 1810, at the age of twenty. In 1817 he went back to England, again, to study law. During his stay there, in the year 1819, he published his account of Australia, *A Statistical, Historical, and Political Description of the Colony of New South Wales*. It was only in 1824 that he returned to Australia.

What can we now expect from this source? True, the author was born in Australia and had spent a few years there. But his entire education was conducted in England. Is this a problem? It would be a problem if we assumed that there had already been some kind of Australian English at that time and that since Wentworth did not spend all of his linguistically formative years in the Antipodes, his evidence must be dismissed.

But there was nothing like AusE at that time. There were only various dialects and sociolects of English spoken and written in the Antipodes. These were influencing each other and were themselves influenced by their environment. If we do not accept Wentworth as a source, we cannot accept a single source from that time. And if we want to find the origins of Australian English, we have to look for them in documents like that.

1.1.2 SELECTION CRITERIA

Material to be included had to meet with a regional and a temporal criterion (1788-1900).

The required place of writing was Australia, New Zealand or Norfolk Island. But other localities were allowed, if the writer was a native Australian or had lived in Australia for a considerable time. Sample size was not an essential criterion. Although full texts were preferred, e.g. with letters, articles and speeches, this was not always possible.

1.1.3 CORPUS SOURCES

The data for the corpus come from more than 100 different sources. They cannot be named here in full, some examples must suffice.

A number of letters come from the Mitchell Library in Sydney, New South Wales, which holds a vast amount of original documents relating to the history of Australia from its earliest times. These unedited letters were transcribed during visits to the library by the author.

Another source is published material in book form. Many historians have striven to evidence the course of Australian history by editing historical official and unofficial documents, letters, diaries, proclamations, newspaper reports, legal texts, etc. Examples are the exceptional works of Manning Clark (1975, 1977), David Fitzpatrick (1994), Ward & Robertson (1969) and O'Farrell (1984).

By far the most accessible kind of material consists of historical texts which are published on the internet. The most comprehensive and ambitious undertaking in the Australian context is the SETIS programme. It is housed at the University of Sydney Library and provides online access to a large number of full texts. Many of these are literary but there are also some historical texts. Examples of texts from the Setis website are the works of Marcus Clarke and Henry Lawson. Also, the complete Federation Debates (Melbourne 1890, Sydney 1891, etc.) are to be found there.

Altogether, **more than ten million words** of early English in Australia were collected. Above that **additional contemporary data** from Great Britain, Ireland, New Zealand, South Africa and Canada were found and edited. This constituted reference material and was used for comparisons.

1.2 Editing and Codification of the Sources

1.2.1 EDITING THE DATA

After computerization, each text received a heading which states its Source Identification Number (SIN) and provides data about the author and the source. In the corpus, the SIN is assigned chronologically. It starts with a number between 1 and 4 (for the period the document was written in) and then, after a hyphen, has a three digit number for further identification. The SIN is always given in pointed brackets when a quote from the corpus is presented.

1.2.2 INFORMATION ABOUT THE AUTHORS

The following data about the authors (if known) were collected:

- name
- year of birth
- gender
- country/region of origin
- social status
- year of arrival in Australia

1.2.3 INFORMATION ABOUT THE TEXTS AND ADDRESSEES

The following textual properties were ascertained (as far as possible)

- year of writing (or of publication)
- place of writing
- register of the text
- text type
- the number of words (counted by Microsoft Word 2000)
- the name of the source and the pages in the original text (if applicable)
- gender, status and abode of the addressee (if applicable)

2. From Data to Corpus – Building Principles

Two principles were adhered to when building the corpus. First, there was a temporal criterion that would enable valid diachronic comparisons, and second there was a register/text type criterion.

2.1 The Principle of Periodization

The corpus material was divided into four different periods, namely: 1788-1825, 1826-1850, 1851-75 and 1876-1900. In every period there were to be an equal number of words (ca. 500,000).

These periods roughly correspond to Mitchell's (1995:1) divisions of Australian history in the eighteenth and nineteenth centuries. They are also in line with the periodizations used by historians.

- 1) Convicts and Settlements in the Cumberland Plain (1788-1825)
- 2) Pastoral Expansion and free (assisted) immigration (1826-1850)
- 3) The Golden Decade and its consequences (1851-75)
- 4) The rise of white natives and urbanisation (1876-1900)

2.2 The Principle of Register

The second principle stated that in every period there should be a like number of words in the different registers:

Speech-Based (SB):	15%	=	75,000 per period
Private Written (PrW):	35%	=	175,000 per period
Public Written (PcW):	40%	=	200,000 per period
Government English (GE):	10%	=	50,000 per period

The Public Written (PcW) register dominates the corpus, since these writings were most widely distributed and certainly made up the lion's share of Australia's linguistic output. Next comes the Private Written (PrW) register. This represents the thousands of letters and diaries in which almost everybody confided his or her private joys and sorrows. The Speech-based (SB) register is comparatively small. This is certainly not representative of total 'production' of English in nineteenth century Australia, but is due to a lack of sources. By far the smallest register is Government English (GE). GE was a register used only by a very restricted number of people in clearly defined situations.

Since the sources used are of uneven length and their word counts are computed differently by different programs, the actual numbers diverge somewhat from the idealized table given above.

Altogether, the corpus, based on the principles of periodization and register, comprises ca. **2 million words** in **1357 texts**.

3. A Description of a Corpus of Oz Early English (COOEE)

The self-collected and self-edited corpus was compiled in the course of the years 1995-2001, although work on it was intermittent. It started from a body of mainly Irish-Australian letters which formed the basis of the master's thesis *Early Australian Letters – A Linguistic Analysis* (Fritz 1996). The sources are of very uneven length, ranging from diary excerpts to book chapters. Therefore the number of words in a category gives a much clearer account of the available material than the number of sources does. For this reason the word counts and not the sample counts will mostly be used in the description.

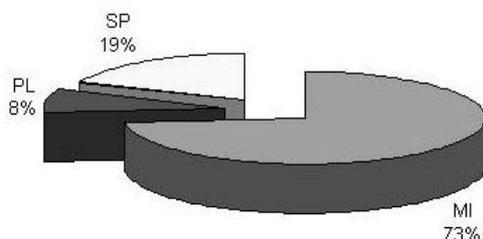
3.1 Register and Text Type

The individual registers are made up of several distinct text types. The following figures show the share of each text type in a register.

3.1.1 THE SPEECH-BASED REGISTER

The exact number of words in this register included in COOEE is 303,850 according to a *Microsoft Word 2000* count and 291,921 words according to a count by *WordList*.

Figure 1: Speeches (SP), Plays (PL) and Minutes (MI) in SB-Register

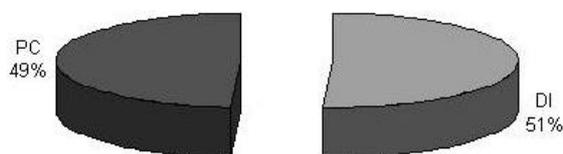


The high amount of MI seems at first astonishing, but is explained by the ready availability of the Federation Debates and the court minutes of the Superior Courts of New South Wales 1788-1899, published by the Division of Law, Macquarie University. So it was possible to keep the share of speeches and even more that of plays quite low. This was thought desirable since minutes of any kind should allow a better look at actual spoken language than the other two text types.

3.1.2 THE PRIVATE WRITTEN REGISTER

706,691 words for COOEE come from personal letters and diaries. Both have a very like share in the PrW register. So the intimate conversation with oneself is balanced against the need to stay in contact with loved ones.

Figure 2: Personal Communication (PC) and Diaries (DI) in PrW register



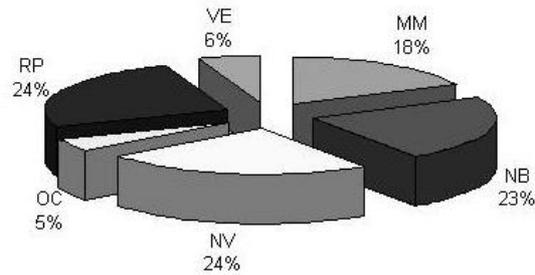
3.1.3 THE PUBLIC WRITTEN REGISTER

In this register we can find many different text types, which can also be very different from each other. The unifying bond is the intended publication, i.e. the address to persons unknown. Altogether, 793,593 words were included.

The distribution of the text types over time is likewise not even. For example, there is no narrative in the first period, indeed the first Australian novel appeared only in 1829. So a higher number of reports had to be included in the first period. When the number of reports is lowest, then memoirs

reach their highest point. Newspapers and broadsides, as well as verse and official correspondence do not differ much over time.

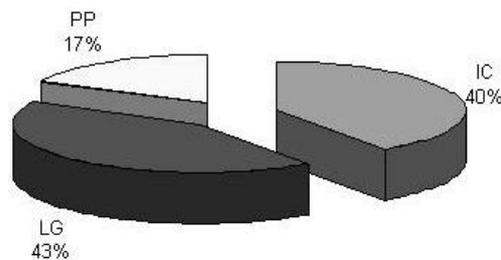
Figure 3: Memoirs (MM), Newspapers & Broadsides (NB), Narratives (NV), Official Correspondence (OC), Reports (RP) and Verse (VE) in PcW register



3.1.4 THE GOVERNMENT ENGLISH REGISTER

Legal English and Imperial Correspondence take the greatest share of the 231,526 words in the GE register. Petitions and Proclamations are trailing somewhat at 17%.

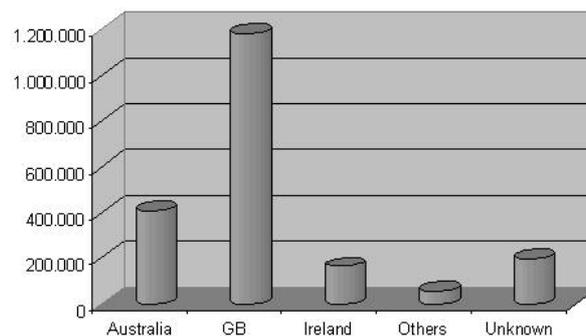
Figure 4: Imperial Correspondence (IC), Legal English (LG) and Petitions & Proclamations (PP) in GE register



3.2 Origins

The next figures show us where the authors of the sources came from.

Figure 5: Origins of Authors (all)



People hailing from Great Britain wrote most of the sources (1,160,619 words), but there is also a substantial amount from native Australians (400,670) and from Irish-born people (163,050).

Most of the writers whose origin is not known come from either the British Isles or were born in Australia. The label 'unkown' serves more as a precaution against an exaggeration of the number of either, but does not designate large numbers of people from outside the British Empire.

This mixture is, of course, not stable across the decades. Figure 6 shows that the amount of British born authors fell while those of the native borns rose considerably. But it was not only well after the gold rush period that native Australians contributed a greater percentage of writings to COOEE than those of British descent. This figure does not include *Others* and *Unknown* since it only serves to illustrate the rise of native writers.

Figure 6: Origins of Authors (Great Britain, Australia, Irish) divided into periods

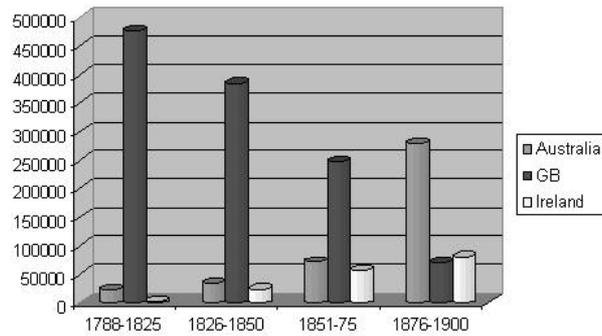
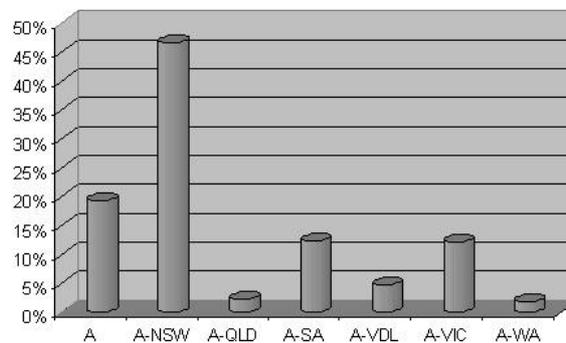


Figure 7 shows the origins of the native writers in relation to the number of words contributed to COOEE. Most were born in NSW, and many in Victoria and South Australia. There is also a very substantial number of people whose origin could not be exactly located in Australia (19%).

Figure 7: Origins of Authors (native Australians)

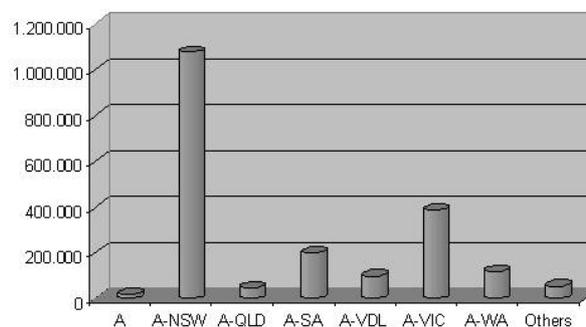


3.3 Place of Writing

All of the states of Australia are represented in the places of writing. Naturally, New South Wales takes the lead, followed by Victoria, South Australia, Western Australia and Van Diemen’s Land (today’s Tasmania).

For a text to be assigned to a state, today’s political borders were used, even if this state was historically not in existence at that time. Otherwise the regional distribution would have been skewed, e.g. if a text written at Port Phillip would be counted as coming from NSW.

Figure 8: Place of Writing

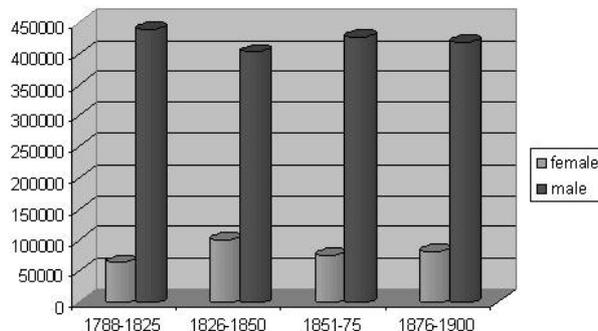


Texts written in Great Britain, at Sea or in other places outside Australia were included in the corpus if their author was a native Australian or had lived there for a considerable time.

3.4 Gender

Most of the writings comes from male authors, but there is also a substantial amount written by women. Considering the total size of the corpus we find that the women's total 16% equals 322,699 words.

Figure 9: Gender Distribution over time

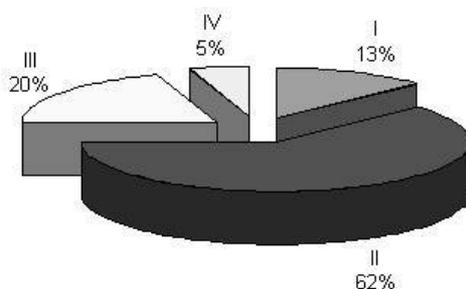


3.5 Status

Status is another important variable that can tell us a lot about an author. The governor of New South Wales is in frequent contact with other speakers of Standard BrE and in constant need to use it, so he is not likely to nativize his English early or quickly. Moreover, most of them returned to England after their service.

A convict's outlook on life, on the other hand, was different. He or she was not mobile, hoping to make a living from whatever possible source. This meant that a blending in was absolutely necessary, linguistically and socially.

Figure 10: Status Distribution (general)



The distribution of data from authors of different status owes a lot to the availability of the sources. Status I people were not frequent at all in early Australia, but their need to write was very high and historical interest in them has also added to the survival of their writings. People who were assigned status II were not the most frequent in total numbers in the population, of course, but all of them could read and write and had the leisure and the friends and relatives who encouraged writing. They also formed public opinion by speeches, articles, literary writings and other activities.

Not all of the people classified as status III were able to read and/or write. But being separated from their family basis at home and struggling to make a living, writing was a heart-felt necessity and comfort. Nevertheless they wrote infrequently, sometimes for a lack of means and sometimes simply out of shame.

Even more elusive are the letters and diaries from the lowest social classes. They have been preserved but seldomly and only meticulous historical work like that of Webby (1989), Fitzpatrick (1994) and O'Farrell (1984) has brought some of their writings to light.

4. EARLY AUSTRALIAN SPELLING –OR/-OUR

[T]here is no valid etymological reason for the preservation of the *u* in such words as *honor, labor*, etc. [...] The tendency of people in Australasia is to excise the *u*, and one of the Sydney morning papers habitually does this, while the other generally follows the older form. [...] [The American spelling is] the original and purer English — the English of Shakespeare, which has been preserved in the form in which the Pilgrim fathers took it away with them.

E.J. Forbes, Sydney manager of Merriam Webster company, pre 1900¹

4.1 Historical Insights into –or/-our

Today many people associate the different spellings of words like *favour/favor* with differences between BrE and AmE spelling practices.

Since *honor, labor, color, favor*, etc. are perceived as ‘Americanisms’ their appearance is often attributed to an increasing American influence. In this vein Peters (1986a:6f) writes:

Significant numbers of people do however accept different forms of some words as legitimate. **An increasingly familiar case** [highlight by C.F.] is that of words like *colour, honour*, [...] The practice of using *–or* is often spoken of as an [sic] “American” spelling, but the increasing frequency of its use here makes it a significant Australian variant, too.

In this publication Peters, like others, tacitly assumes that the spelling principles of AusE were once consistent with BrE ones and that American forms ‘intruded’ in the second half of the twentieth century. Peters also gives the mid-1985 circulation figures of newspapers following each spelling variant. The circulation figures for the *–or* newspapers are twice as large as those for the *–our* newspapers. The following data are taken from Peters (1986a:7).

Table 1: Newspapers using *–or/-our*

Newspapers using <i>–or</i>	Circulation figures	Newspapers using <i>–our</i>	Circulation figures
<i>Adelaide Advertiser</i>	215,956	<i>The Australian</i>	119,010
<i>Adelaide News</i>	176,044	<i>Australian Financial Review</i>	60,000
<i>The Age</i>	247,000	<i>Canberra Times</i>	45,253
<i>Brisbane Courier-Mail</i>	227,943	<i>Sydney Morning Herald</i>	258,700
<i>Daily Telegraph</i>	299,797	<i>West Australian</i>	236,031
<i>Herald</i>	337,003	Total	728,994
Total	1,503,743		

Leitner (1984) also comments on the spelling variabilities in Australian newspapers and hints at the possibility that American owned newspapers are introducing American spellings. Regional differences are alluded to, a suggestion supported by Peters (1995:546f).

This suggested ‘Americanization’ of AusE becomes less convincing when historical sources are taken into consideration. Leitner (2002:91) rightly observes that Australia’s first governor, Arthur Phillip, quite naturally uses a mix of spellings: *honor, harbour, favorable, labour, labor, encrease, expenses, expence* in his official letters. Peters (1995) does no longer mention AmE as a possible source or a reinforcing factor for *–or* spellings.

It is a well-known fact that the spelling of some words was variable in the eighteenth and nineteenth centuries, exactly the period when AusE was beginning to raise its head. Although the establishment of a standard orthography had been a growing concern of enlightened theorists, a complete and undisputed system was not achieved until late into the nineteenth century. The latter development was especially due to the habits of a rising middle-class who desperately wanted to ‘upgrade’ their language (first in spelling and grammar, later also in pronunciation).

Much ink was spilled in the eighteenth century on the question of *–or/-our*. Etymology, the science of true meaning, was used as an arbiter. Latin-derived words should have *–or* and French-derived words should have *–our*, producing a mixed spelling system. However, not always did scholars

¹ The quotes are from the unpublished pamphlet: *The So Called "American Spelling." Its Consistency Examined*. The *Brisbane Courier Mail* (22/12/99) printed parts of it in the article “Yankee slang rocks into ‘Strine’ via Internet”. Since the original article was not available, the comment on it by Annette Potts in *Bikwil* had to be used. This can be found under the following address: <http://www.bikwil.zip.com.au/Vintage19/Webster's-Dictionary.html>

agree on a word's history. Even three Old English words were erroneously given *-our* spellings, namely *harbour*, *behaviour* and *neighbour* (Peters 1986b:20).

Noah Webster's blue-backed *American Speller* was first published in 1785, but some of the spellings in its first edition were later modified to achieve greater consistency. 80 million copies in Webster's lifetime and many personal tours² later, AmE was on a course of accepting a consistent, simplified orthography.

In Britain, on the other hand, the trend towards *-or* spellings was arrested by successive, only minimally altered, reprints of Dr Johnson's dictionary (Peters 1986b:21; 1995:547).

During the early formative years of AusE there simply was no American or British standard and *-or/-our* was certainly not considered to be such a distinction. The *Melbourne Age* decided as late as 1854 that *-or* spellings are 'better' and that they therefore should be used in all articles. As we can see from Table 1, this policy has not change after 150 years!

4.2 Orthographic Standards in NZE, BrE, AmE and COOEE

Sigley (1999) has investigated spelling practices in NZE, AmE and BrE using the WWC (1986 NZE), Brown (1961 AmE), Frown (1991 AmE), LOB (1961 BrE) and FLOB (1991 BrE) corpora. He established three groups for the spelling variants (1999:8):

- (a) standardised in BrE but variable in AmE
- (b) fully opposed BrE and AmE standards
- (c) standardised in AmE but variable in BrE

The *-or/-our* and *-ol/-oul-* differences are in group b. Table 2 uses his findings (1999:9) which are complemented by the ones in COOEE:

Table 2: Numbers of *-ol/-oul-* and *-or/-our* variants

Variables	Brown	Frown	LOB	FLOB	WWC	COOEE
-OL/-OUL-	90/3	23/1	0/34	0/45	0/47	0/50
-OR/-OUR	1425/31	1331/33	10/1394	10/1123	10/1542	728/3234

The numbers clearly show that there is a categorical difference in the spelling of words like *mould* between AmE and all other varieties looked at. Although there are a few American examples of an *-oul-* spelling, the reverse is never found. The findings from COOEE tell us that this orthographic feature was already so well-established around 1800 that in two million words of nineteenth century Australian texts not a single instance of *-ol-* was used!

There are also words which lost *-oul-* spellings. One example of this is *contro(u)l*. Table 3 shows that *controul* was a negligible variant in nineteenth century Australia and that its use petered out in the course of time. The reference material for BrE contains three more examples of this variant. Joseph Banks uses it in his journal in 1769, British government committees in 1810 and 1812.

Table 3: Frequencies of *control/controul* over time

Variable	Period 1	Period 2	Period 3	Period 4
control	10	16	36	109
controul	5	2	1	0

Seven of the eight instances of *controul* were written by status II persons, the last by a status I person. Five times it comes up in legal documents and petitions and proclamations from the GE register, twice in court proceedings in passages where reported speech is used and once in verse. The origins of the author and the place of writing do not add further information. Clearly, *controul* was an obsolete variant of *control*, used very infrequently by people of higher status in very formal contexts.

The second important conclusion to be drawn from Table 2 is that the spelling *-or/-our* is variable in COOEE. This variability is much higher than in any other present day variety with *-or* spellings

² The story is told, by an old printer recalling his apprenticeship, of the day 'a little pale-faced man came into the office and handed me a printed slip, saying, "My lad, when you use these words, please oblige me by spelling them as here: *theater*, *center*, etc.'" It was Noah Webster traveling about the printing offices and persuading people to follow his 'improved' conventions. (quoted from McCrum, MacNeil & Cran 1992:258)

making up some 18.4% of the total. The investigation of this variability seems very much worthwhile.

Before proceeding to this investigation, one other word has to be looked at, namely *governor/governour* (1228/2). The latter variant comes up twice in a single text, a personal letter by the naval surgeon George Worgan to his brother in 1788 <1-014>. This orthographic idiosyncrasy seems like a relic from a dim and distant past, if we compare it to the 1228 instances of *governor* from later dates. Above that, Worgan himself uses *governor* 15 times.

4.3 –OR-/OUR Variation in COOEE

4.3.1 VARIABILITY OVER TIME

Total numbers can be deceptive. In order to establish exactly where and why there is orthographic variability in COOEE, individual words have to be looked at. The frequencies are given for COOEE as a whole, then for each period 1788-1825, 1826-1850, 1851-1875 and 1876-1900 and finally the frequencies in the AusE post1900 material (1901-23). Since the latter is not a balanced collection of texts, the results must not be taken with a pinch of salt. Nevertheless, they can be indicative in some cases.

Table 4: Frequencies of *-or/-our* spellings in individual words over time

Variables	COOEE	Period 1	Period 2	Period 3	Period 4	AusE post1900
ardor/ardour	0/8	0/2	0/1	0/3	0/2	0/1
armor/armour	0/9	0/1	0/2	0/4	0/2	0/6
*behavior/*behaviour	0/73	0/45	0/13	0/5	0/5	0/1
candor/candour	0/11	0/5	0/4	0/0	0/2	0/1
clamor/clamour	0/8	0/1	0/1	0/5	0/1	0/0
color/*colour*	5/268	1/100	3/58	1/50	0/60	0/44
*demeanor/*demeanour	28/20	14/6	3/10	11/3	0/1	0/4
endeavor/endeavour	10/359	1/113	0/120	0/76	9/50	0/13
favor*/favour*	126/389	51/114	36/117	11/81	28/77	2/59
fervor/fervour	2/7	2/0	0/3	0/2	0/2	0/0
flavor/flavour	0/20	0/5	0/5	0/6	0/4	0/3
harbor/harbour	2/305	0/175	1/82	0/31	1/17	0/3
honor/*honour*	328/344	126/81	94/84	35/127	73/52	0/44
humor*/humour*	1/55	0/9	0/18	1/12	0/16	0/7
labor*/labour*	200/783	25/188	59/202	28/148	88/245	2/31
neighbor*/neighbour*	12/405	1/71	5/135	3/135	3/64	0/24
odor/odour	0/15	0/1	0/6	0/4	0/4	0/2
parlor/parlour*	0/20	0/3	0/6	0/4	0/7	0/6
rigor/rigour	0/10	0/6	0/3	0/0	0/1	0/0
rumor/rumour	0/24	0/8	0/7	0/4	0/5	0/2
savior/saviour	0/21	0/7	0/14	0/0	0/0	0/6
savor/savour	0/9	0/6	0/0	0/2	0/1	0/1
splendor/splendour	3/18	2/4	0/7	0/5	1/2	1/4
succor/succour	0/7	0/1	0/2	0/4	0/0	0/1
valor/valour	0/5	0/0	0/3	0/0	0/2	0/0
vapor/vapour	0/20	0/5	0/4	0/4	0/7	0/0
vigor/vigour	2/20	0/5	0/4	0/8	2/3	0/8
Total <i>-our</i> :	3233	962	912	724	635	274
Total <i>-or</i> :	719	223	201	90	205	5

A X²-analysis of the overall frequency distributions of *-or/-our* spellings over time, as shown in Table 4, reveals a significant drop in *-or* spellings in the third period and a significant rise of these in the fourth period! This runs contrary to the expectation that *-or* spellings should become less

frequent continuously. If *labo(u)r* and *hono(u)rable*, for reasons explained later, were removed from Table 4, a continuous decline of *-or* would, however, be discernible.

It is clear that the spelling patterns did not stay stable and that there was no one-way development. It is notable that the data from AusE post1900 show even less evidence for *-or* spellings.

There are great differences between the number of occurrences in individual words. Many do not show an *-or/-our* variation at all, but are consistently spelled *-our*. These are: *ardour*, *armour*, *behaviour*³, *candor*, *clamor*, *flavour*, *odour*, *parlour*, *rigour*, *rumor*, *saviour*, *savour*, *succour*, *valour*⁴ and *vapour*.⁵

A number of words show an orthographic consistency of more than 90%. These are *colour* 98.2%, *endeavour* 97.3%, *harbour* 99.3%, *humour* 98.2%, *neighbour* 97.1% and *vigour* 90.9%. Obviously, the *-or* spellings for these words are idiosyncrasies and not stable features of a developing AusE standard.

Three of the five instances of *color* come from women's writings. Mary Vidal provides two instances of it. Interestingly, she uses the spelling *colour* only twice in a text of almost 10,000 words <2-308>. That means that for her *color* and *colour* were equally distributed orthographic variants.

Endeavor can be found ten times in COOEE. Once in the first period in a NSW broadside. But nine times it comes from the fourth period from a speech and from parliamentary debates and is used by natives from the state of Victoria (William Spence once, George Turner eight times). Turner's use of *endeavor*, is, however, subject to the process of minute taking by an unknown clerk. Yet, it can be said that Turner, or the clerk responsible for this session, was the only one to use the *-or* spelling here. In this text, <4-421>, there is also *favor*, *honor*, *honorable* and *neighbors*. According to Peters (1995:547) the Victorian Education Department endorsed *-or* spellings in 1910, in the 1930s and the 1970s.

The only two occurrences of *harbor* come from two very different sources. One is in a government proclamation by James Stirling in Western Australia who had already been in Australia for nineteen years <2-313> and the other in a private letter by the NSW-born Arthur Streeton <4-237>.

Apart from *humorist*, which was not taken into consideration, *humor* leads a solitary life in COOEE. It is used by John Cross who had just touched Australian soil when he writes his letter home <3-278>.

Considering raw numbers, the twelve instance of *neighbor* are a lot. Four of them come from Penelope Selby, <2-325>, <2-353>, <2-362>. Only one is from a native Australian, seven are from relatively recently arrived British and three are from recent Irish arrivals.

Vigor is only found in two texts written by native Australians.

This leaves a rather small number of words where orthographic variability was indeed a factor: **demeanour* 41.6%, *favour* 75.5%, *fervour* 77.7%, *honour* 50.9%, *labour* 79.7% and *splendour* 85.7%.

Fervour is a comparatively infrequent word evidencing a straight line of development. *Fervor* comes up only in period 1, whereas *fervour* is to be found only in periods 2, 3 and 4.

Looking at *splendour* we get a very similar picture. *Splendour* is only a minor variant coming up mostly in period 1. But we can also find it once in Ref-AusE post1900!

³ Peters (1995:546) mentions that in ACE *behavio(u)r* shows the least likelihood for *-or* of any of the words looked at (10:99).

⁴ *Ad valorem* comes up 9 times, *valor* never.

⁵ Note that there are 15 instances of *evaporate*. According to the OED *evapourate* is last recorded in the early seventeenth century. The same holds true for *invigorate* (9), despite the fact that *invigour* is recorded as late as 1899 in the OED!

**Demeanour* is the only word in COOEE where *-or* is found more frequently than *-our*. It has to be noted that the *-or* variable only occurs in *misdemeanor*. If *misdemeanour* is taken as the headword, the frequency of *-our* spellings decreases to only 28.2%!

Demeanour only occurs nine times and in a variety of text types: DI, PC, MM, NB, RP and IC. *Misdemeano(u)rs*, on the other hand, are much more restricted for semantic reasons. 19 times they are in legal texts, seven times in minutes of court proceedings, six times in imperial correspondence, only seven times in other text types, but always dealing with criminal matters. The decreasing frequencies of *misdemeano(u)rs* over time shows a lessening concern with crime in the colonies.

The favoured spelling *misdemeanor* only occurs in official writings, whereas *misdemeanour* is also found elsewhere. This suggests that in this word *-or* was thought appropriate for formal writings and that *-our* is a popularized variant of it, bowing to ‘public’ pressure. There is only a single legal document where *misdemeanour* comes up <2-367>. In the case of this word, *-or* is likely to have been supported by the fact that legal texts evidence a very conservative character and thus Latinate spellings.

Although there is a decline in spelling *favour** as *favor**, this is not statistically significant over all four periods. There is a clear drop in the third period, but in the last period under investigation, *favor** rises again, though not to the levels it had in the first period. Change is least pronounced in *favor* (period 1 = 21 instances, period 4 = 17 instances) and *favorite* (period 1 = 10 instances, period 4 = 7 instances) and most evident in *favorable/ly* (period 1 = 17 instances, period 4 = 4 instances). After 1900 *-or* spellings are almost completely gone.

*Honour** and *Honor** are to be found with almost like frequencies in COOEE. The first three periods show a consistent decline of *honor** from 60.9% to 21.6%, the big surprise is its rise to 58.4% after 1876. Again, a look at specific words yields interesting insights.

Table 5: Frequencies of Honor/Honour and Honorabl*/Honourabl*

Variables	COOEE	Period 1	Period 2	Period 3	Period 4
honor	216	118	71	12	15
honorab*	110	7	22	23	58
honour	222	62	52	67	41
honourabl*	109	16	31	54	8

There is semantic distinction between *honor* and *honour* in COOEE in that the first is commonly used to address heads of court, the Governor and the Chief Justice whereas the latter is mainly reserved for figurative use.

the Troops in garrison fired 3 Volies of Small arms in honour to her Majastyes Birth day <1-034>

When the importance of the system of justice and of the Britain-sent Governor declines, so does the number of occurrences of *honor*. *Honour*, on the other hand, is in use in almost like numbers in all periods. Here the *-our* spelling wins out.

Hono(u)rable goes into the opposite direction. This is most noticeable in the fourth period. *Honorabl** significantly rises and *honourable* declines in complementary fashion, although 25 years earlier the situation was almost reversed. This is evidence of a swift and powerful linguistic change the exact causes of which are unclear.

There could be a functional reason behind it. Most instances of *honorable* come from the Federation Debates of the 1890s where members address each other as *Honorable member*, only once *honourable member* can be found. This is certainly an effect of spelling conventions established by Parliamentary clerks.

The last word with variable spelling is *labour/labor*. Whereas the first spelling stays relatively stable, the latter actually increases in frequency over time. The significant rise after 1876 is mainly due to the establishment of the Australian Labor Party, which is mentioned in one way or another 78 times out of 88 instances of *labor**.

If we discount Australian Labor associations, perception of the development changes. Although *labor* continues to be a spelling used, it becomes a considerably less important variant in the late nineteenth century. The two post1900 instances of *labor* fit in well with this pattern.

Nevertheless, the fact that *labor* was the chosen spelling for *Labor Party*, *Labor platform*, *Capital and Labor*, *labor bodies*, *labor candidates*, *labor members*, *labor interest*, *labor organisations* and other collocates testifies to its strength.

Today AusE seems to favour *labor* over *labour* even more than in COOEE. Peters (1995:546) reports that here *-or* numbers are highest even after the exclusion of all references to the Australian Labor Party (129:95)!

4.3.2 THE INFLUENCE OF ORIGIN OF AUTHORS

One of the hypothesis to be tested is that the frequency distributions of *-our/-or* spellings is related to the origin of the writer. The results are shown in Table 6.

Table 6: Frequencies of *-our/-or* in Relation to Origin

Country of Origin	# of <i>-our</i>	# of <i>-or</i>	%
Great Britain	1,955	422	82.2%
Ireland	213	21	91%
Australia	637	191	76.9%

If the relations are compared for statistical significance, the result is that the origin of the author is indeed a significant factor determining the choice of the word ending.

All are significant at a level of confidence of 0.001, i.e. writers from Australia, Great Britain and Ireland differ from each other and there is only a 1/1000 chance that this statement is wrong.

Usage is most variable with Native Australians and most consistent with Irish men and women. British writers are between these two extremes, but in the end, they follow the Irish (!) and today almost exclusively use *-our*, whereas AusE has kept some of its variability.

As regards the development from early English in Australia to Australian English, this finding clearly shows that the native population was the decisive factor in shaping a new variety of English.

4.3.3 THE INFLUENCE OF STATUS OF AUTHORS

The last hypothesis to be tested is that the status and education of the writers is related to the choice of a certain word ending. Table 7 states the frequency distributions.

Table 7: Frequencies of *-our/-or* in Relation to Status

Country of Origin	# of <i>-our</i>	# of <i>-or</i>	%
Status I	421	133	76%
Status II	2106	542	79.5%
Status III	559	45	92.5%
Status IV	112	14	88.9%

There is a clear line of demarcation between Status I and II and Status III and IV. There is no significant difference in usage between Status I and II writers. Neither is there one between Status III and IV authors.

All other differences are statistically significant well below the 5% level that is generally agreed in the social sciences to prove the correctness of a hypothesis. This means that *-or* is significantly more frequent in the writings of the educated than in the writings of the less educated.

Table 8 looks at the frequencies of *labour/labor* spellings. A X^2 analysis reveals that Status III authors are significantly less likely to use *labor* than expected.

Table 8: Frequencies of *labour/labor* in Relation to Status

Variable	labour	labor	Total
Status I	172	33	205
Status II	474	148	622
Status III	122	12	134
Status IV	15	7	22
Total	783	200	983

The last individual word looked at is *favour/favor*. If X^2 is applied it can be shown that Status I authors are significantly more likely to use *favor*, whereas Status III authors are significantly less likely to use this variant.

Table 9: Frequencies of *favour/favor* in Relation to Status

Variable	favour	favor	Total
Status I	33	37	70
Status II	272	77	349
Status III	72	9	81
Status IV	12	3	15
Total	389	126	515

In conclusion, it can be stated that the differences in the frequencies of Latin *-or* and French *-our* spellings are certainly class related. The more educated a writer is, the more likely he/she is to use the Latin form. The French variant was a demotic upstart, which, in Britain and in Australia, succeeded in becoming the majority word ending.

4.4 Comparison of COOEE with Reference Data

In a last step, the findings from COOEE are compared with those from reference material, collected alongside COOEE, containing 950,000 words of BrE (1768-1886), 29,000 words of writings from New Zealand (1840-60), 60,000 words of IrE (1791-1923) and 33,000 words of English in Canada (1801-47). As already stated above, the reference material is not a systematic collection of texts and thus does not form a corpus. All results derived from it must therefore be interpreted cautiously. Table 10 lists the individual and the total frequencies in COOEE and the reference data.

Table 10: Frequencies of *-or/-our* spellings in COOEE and Reference Material

Variables	COOEE	Ref-BrE	Ref-NZE	Ref-IrE	Ref-CanE
ardor/ardour	0/8	0/1	0/0	0/0	0/0
armor/armour	0/9	0/5	0/0	0/0	0/0
*behavior/*behaviour	0/73	9/40	0/0	0/0	2/1
candor/candour	0/11	1/3	0/0	0/0	0/0
clamor/clamour	0/8	0/8	0/0	0/0	0/0
color/*colour*	5/268	0/282	0/7	0/2	1/2
*demeanor/*demeanour	28/20	1/4	0/0	0/0	0/0
endeavor/endeavour	10/359	0/184	0/1	0/3	0/5
favor*/favour*	126/389	9/275	0/12	1/9	1/4
fervor/fervour	2/7	0/3	0/1	0/0	0/0
flavor/flavour	0/20	0/42	0/0	0/0	0/0
harbor/harbour	2/305	1/245	0/21	0/1	0/0
honor/*honour*	328/344	7/156	0/0	2/2	0/6
humor*/humour*	1/55	5/49	0/0	0/1	0/1
labor*/labour*	200/783	11/172	0/11	0/2	2/3
neighbor*/neighbour*	12/405	0/193	0/8	0/19	4/9
odor/odour	0/15	0/1	0/0	0/1	0/0
parlor/parlour*	0/20	0/19	0/0	0/3	0/3
rigor/rigour	0/10	0/5	0/0	0/0	0/0
rumor/rumour	0/24	0/6	0/1	0/0	0/0
savior/saviour	0/21	0/9	0/1	0/0	0/0
savor/savour	0/9	0/7	0/0	0/0	0/1
splendor/splendour	3/18	1/13	0/0	0/0	0/0
succor/succour	0/7	0/1	0/0	0/0	0/0
valor/valour	0/5	0/1	0/0	0/0	0/0
vapor/vapour	0/20	0/5	0/0	0/0	0/0
vigor/vigour	2/20	0/16	0/0	0/0	0/0
Total <i>-our</i> :	3233	1745	63	43	35
Total <i>-or</i> :	719	45	0	3	10

Due to the differences in size, especially with the Irish, New Zealand and Canadian material, the number of instances differ a lot. What can be compared then are a few individual words, namely *behaviour*, *favour*, *honour*, *labor* and *neighbour*, and the total numbers.

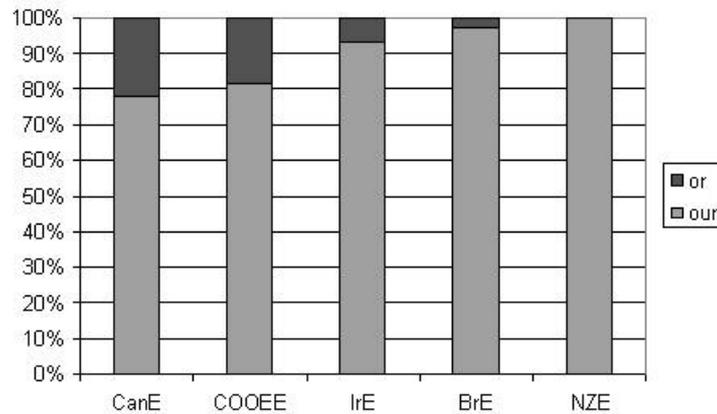
Interestingly, *behaviour* is spelled with *-or* in some instances in the British and the Canadian writings, something not found in COOEE. This can be interpreted as COOEE being more advanced than contemporary BrE which today has completely lost this variant. The Canadian examples, on the other hand, already show a majority of *-or* spellings.

COOEE and Canada show great variability in the spelling of *favour*, whereas everywhere else in the world, the *-our* variant clearly predominates.

The Australian language data for *honour* are very similar to those from Ireland and very different from contemporary British usage. This highlights the early development of differences from the parent variety.

The same is true for *labour*. Early AusE finds a partner in the texts written in Canada, not in those from Ireland. The final individual difference is in *neighbour*, which in Canada has a much higher percentage of *-or* spellings than anywhere else.

Figure 11: Comparison of Total Frequencies



When comparing total frequencies of *-or/-our*, significant differences between the varieties are observable. The language found in COOEE does not differ significantly from the Canadian writings, but it does so from BrE, IrE and the New Zealand material. The last finding suggests that language use in Australia and New Zealand was divided from the start.

BrE usage is related to the one found in New Zealand and Ireland, but there is a significant difference to spelling practices in Canada.

The texts written in New Zealand are significantly unrelated to the Irish and Canadian ones.

And finally, there is a significant contrast between IrE and Canadian texts.

5. CONCLUSION

There is a great need for the corpus-based study of early Australian English. COOEE marks a significant step in this direction.

The investigation of the spelling variable *-or/-our* has shown it to be a well-established feature in nineteenth century Australia. The assumption that *-or* spellings are only due to recent American influence has been disproved. A comparison with contemporaneous varieties has shown early AusE to follow its own standards.

Usage of one or the other variable was found to be determined by semantics, by the time of writing, by the origin of the writers and by their status.

All in all, *-or* was on the decline in the nineteenth century, but it was always there. It seems possible that after 1945 a rising influence of AmE has reinvigorated this spelling tradition.

6. LIST OF WORKS CITED

- Clark, Manning, 1975, *Select Documents in Australian History 1850-1900*, Sydney: Angus & Robertson.
- , 1977, *Select Documents in Australian History 1788-1850*, Sydney: Angus & Robertson.
- Fitzpatrick, David, 1994, *Oceans of Consolation*, Cork: Cork UP.
- Fritz, Clemens, 1996, *Early Australian Letters: A Linguistic Analysis*, unpublished MA thesis.
- Leitner, Gerhard, 1984, "Australian English or English in Australia: Linguistic identity and dependence in Australian broadcast media", *English World Wide* 5, 1: 55-85.
- Leitner, Gerhard, 2002, "Varietätenkontakt: Der amerikanische Einfluss auf das australische Englisch", In: Bader, R. B. Braun and Adi Wimmer, eds, *Vergangenheit und Zukunft in Australien*, Tübingen: Stauffenberg Verlag, 83-99.
- McCrum, Robert, MacNeil, Robert and William Cran, 1992, *The Story of English*, London: Faber and Faber.
- Mitchell, Alexander G., 1995, *The Story of Australian English*, Sydney: Dictionary Research Centre.
- O'Farrell, Patrick, 1984, *Letters from Irish Australia*, Sydney: New South Wales UP.
- Peters, Pam, 1986a, "Spelling principles", In: Peters, Pam, ed., *Style in Australia: Current Practices in Spelling, Punctuation, Hyphenation, Capitalisation, etc.*, 5-18.
- Peters, Pam, 1986b, "Problem word endings", In: Peters, Pam, ed., *Style in Australia: Current Practices in Spelling, Punctuation, Hyphenation, Capitalisation, etc.*, 19-29.
- Peters, Pam, 1995, *The Cambridge Australian English Style Guide*, Cambridge: Cambridge UP.
- Sigley, Robert, 1999, "Are we still under England's spell?", *Te Reo* 42, 3-19.
- Webby, Elizabeth, 1989, *Colonial Voices: Letters, Diaries, Journalism and other Accounts of Nineteenth-Century Australia*, St. Lucia: University of Queensland Press.