From Plato to Aristotle -
Investigating Early Australian English

1. Early Australian English – the State of the Art

The possible origins and developments of early Australian English (AusE) are still under debate. Many different questions have been asked. They can be summed up in the following way:

    1. How did AusE come to be a distinctive variety of English?
    2. What is the linguistic input of AusE?
    3. How did it develop lexically?
    4. How is it possible that AusE is so remarkably uniform across a whole continent?
    5. Where, when and how did the three sociolects (Broad, General and Cultivated) arise?

Many scholars have tackled these questions and have come up with a number of answers. However, no consensus has been achieved yet. This is because mostly 'Reason' has formed the basis of the answers. There are as yet too few empirical studies of early AusE.

Although 'Reason' is always a good companion for the scholar, it cannot stand alone. Empirical facts, especially the investigation of actual instances of early AusE, **must** be an integral part for good answers to the questions posed above. Thus Plato, the founder of Rational Investigations, must be followed by Aristotle, the founder of Empiricism.

What answers have been given so far?

1. How did AusE come to be a distinctive variety of English?
Basically there are three different positions here. The majority view is that AusE is the result of a mixing of dialects with Collins (1975) and, to some extent, Horvath (1985) claiming that the mixing had already taken place already in the greater London area while Bernard (1969, 1981) and Trudgill (1986, the only one to take not only pronunciation but also lexis and grammar into consideration) contend that most of the mixing must have taken place on Australian soil.

Turner (1960), Gunn (1972), Hammarström (1980) and Cochrane (1989) believe in the direct 'transplantation' of a London dialect to Australia, which has not changed much. As evidence they use investigations of the phonological systems of **present-day Broad** AusE and **present-day** or **earlier 'Cockney'**. This seems not above criticism methodologically.

Mitchell (1995) shows his superior historical insight when he suggests a compromise by pointing to the fact that due to the Agrarian and the beginning Industrial Revolution the regional dialects of England had already started to break down in eighteenth and nineteenth century England. So the dialects had already been mixed and levelled, especially in the great urban centres, prior to the European settlement of Australia. This process was then continued in the colonies.

2. What linguistic input do we have?
AusE is undoubtedly very closely related to the English spoken in south-eastern England. But where did the first white Australians actually hail from?

The most reliable figures about the origins of the early Australians come from the counties of trial of the convicts (cf. Robson 1955:155). The early statistical material for the population in Australia is, however, very thin. The first colonial counts, up to the 1850s, did little more than note whether someone was born in Australia or not (cf. Price 1987:3).

Taking Robson's calculations we find that 17 per cent of the male convicts were tried in London. Lancashire accounted for 7, Dublin for 5, Yorkshire for 4 and Warwickshire for 3 per cent. All other counties had 2 per cent or less. Overall, 71 per cent of all male convicts had been tried in England, 22 per cent in Ireland, 5 per cent in Scotland and the rest overseas. A comparison of place of trial and place of birth shows that at the country level the varations between these is small (cf. Jupp 1989:24).

Considering the above facts, it seems surprising that so much emphasis has always been laid on the London heritage of AusE. Demographically the picture is much more complicated.

a) What did the Irish contribute?
According to Ramson (1966), Bernard (1969) and many others, the Irish did not contribute much to the formation of AusE, despite forming some 25 per cent of the white antipodean population in the nineteenth century. Of late, several scholars have criticised this opinion. Troy (1992:460) contends that no Irish influence has been found, because no one has looked for it thoroughly. Horvath (1985:39) and Trudgill (1986:139f) have provided lists of features that could have originated in Irish English (IrE) or even Gaelic, which was still spoken by a number of Irish in Australia.
O'Farrell (1989, 1996) and Fitzpatrick (1994) have shown the close family patterns which the Irish maintained in Australia. Fritz (2000a,b and no date) has shown that this could lead to a preservation of IrE elements. Taylor (1992, 1998, 2001) and Lonergan (fc.) discuss lasting impacts on AusE.
A closer look at actual language data, an empirical approach, could thus prove previous assumptions not to be correct.

b) What is the input of the convicts?
The proportion of convicts and their contribution to early Australian English has always been very much exaggerated. The numbers Ward (1958) gives for the convicts and those of convict origin can be shown to be misleading (Mitchell 1995:44) and were taken over too uncritically by scholars like Horvath (1985), Gunn (1992) and others.
Although some convict terms and even some instances of 'cant' have entered AusE (cf. Langker 1980, 1981), their homogeneity as a speech group, and thus their impact, has been more surmised than proved (cf. Mitchell 1995:7). It is very questionable if there was ever something like a homogeneous group of convicts and/or ex-convicts who formed a coherent linguistic community that itself was distinguished from the language of those who had arrived free or were born free. The convicts came from very diverse backgrounds, they did not live under one roof, they could even be rather isolated from each other. Moreover, the divisions between convicts, freed and free were not as sharp as many assume and certainly much less so than in 'Merry ol' England'.
The protestations of some free immigrants that the convicts and emancipist **should be** separated much more from the rest of society were not effective. The rise of many ex-convicts to respectable social status is testified in numerous cases, e.g. the Reibey family as described by Irvine (1992).
Thomas Fellon's letter to his wife in 1835 praises the good living-standard and opportunities for convicts and ex-convicts. This would not have been possible if he, or all convicts, had shown a use of language easily recognisable as being the speech of a criminal and thus detested. Indeed, his language is non-standard, but not in the way described by author of 'cant' dictionaries.

> Der mary, I never work one day but fourteen days for myselfe since I been in this cuntry because it is not allowed by Government but if i wonst got my liberty I cud [...] ten shillings per day Der mary let me know in youre next letter is my fathere live or know or did my sister go to meracar or know Der mary this is fine cuntry is there is in the wourld for ateing and drinking Der mary if you wore in this cuntry you cud be worth pound per week but by owne labour [...]. <2-129>[1]

Empirical investigation of frequencies of convict terminology is asked for. Over and above that, actual language use of convicts and ex-convicts should be compared with that of other social groups before drawing far-reaching conclusion about their contribution to the formation of AusE.

3. How did it develop lexically?
The lexical development of AusE is well-documented in the works of William Ramson (e.g. 1966, 1988). Ramson (1966) uses data culled from personal readings of thousands and thousands of pages of early AusE. Due to the method employed, the number of words is somewhat restricted and frequencies or links between language user and item used are not taken into account. Sometimes,

---

[1] For an explanation of this number see below. Details about the source are given in the Appendix.

Ramson's findings resemble more a literary discussion than a linguistic analysis. A modern style corpus investigation could thus make the findings much more thorough and insightful.

His 1988 *Australian National Dictionary: A Dictionary of Australianisms on Historical Principles* (AND) is a very valuable work since it not only lists the words but also quotes extensively from the sources used. Thus it is possible to go back to the sources and to ascertain the meanings and functions of the words under investigation. However, even this laudable book provides us with little or no information about frequencies and usage by particular groups.

4. How is it possible that AusE is so remarkably uniform across the continent?

It has always been thought remarkable that AusE is so uniform all over the continent. Even the three acknowledged sociolects of AusE seem to be the same everywhere. There are only a few regionalized lexical items which make a distinction between speakers of AusE from different states possible. This has been called into question recently.

Basically, there are three theories that try to explain this uniformity. Those who believe in a direct transplantation of London English to the antipodes see no problem at all.

The second, and strongest group, sees AusE originating in the early phase of the colony in a kind of 'Sydney mixing-bowl'. This amalgam would then spread and level out all other influences through the astonishing mobility of Australians and the fact that 'new chums' always try to blend in linguistically with the 'old hands'.

Görlach (1991:150) questions the 'Sydney mixing-bowl' approach in the following way:

> 1. As the spread of settlements shows, the early speech communities in the east were separated by hundreds of miles from those in the west, and what internal movements there were can hardly have sufficed to spread, say, Sydney norms throughout the continent; and
>
> 2. convicts and their descendants formed the majority in NSW until at least 1840 [this again repeats Ward's misleading numbers], when transportation to this state was ended. Apart from NSW, only Tasmania (from 1804 to 1852) and Western Australia received convicts, but WA did so only from 1850-68. This means that their speech cannot possibly have had any large impact on the entire WA speech community - as indeed the contemporary Irish immigrants failed to have.

It is true that the early communities were very far apart. But Hobart and Norfolk, the earliest settlements outside the Sydney area were settled from Sydney and contact between all of them was frequent.

The second point Görlach takes up can be easily refuted, too. It stems from his assumption about an alleged uniformity of convict speech and its lasting impact on AusE. He even seems to believe that the convicts of Western Australia (WA) spoke exactly the same, and to his mind very distinct, English as the ones who had reached Australia's shores 62 years earlier.

Undeniably the settlement of WA was very isolated and could well have developed its own dialect. But whatever dialect there may have been was surely swamped by the gold rushes of the 1890s.

A rather isolated position is taken up by Bernard (1969). He also thinks of a mixing-bowl, but in his opinion, this operated in all the major sea-ports (Port Jackson, Hobart, Port Phillip, etc.) and since we have the same linguistic input everywhere, the outcome of the mixing was the same and later also homogenized by the above mentioned high mobility.

Although this question seems by and large settled, there are some points which warrant closer investigation, in fact investigation which can only be done empirically. It should prove very rewarding to study the actual linguistic accommodation processes taking place in Sydney and elsewhere. Another point worthy of interest is the study of early language use in South Australia (SA) and WA, the two colonies where local dialects seem likeliest. In order to do this we have again to take recourse to empirical data, nothing else will suffice.

5. Where, when and how did the three sociolects (Broad, General and Cultivated) arise?

The existence of three sociolects of AusE has been convincingly shown by the pioneering work of Mitchell and Delbridge (1965). All of these, Broad, General and Cultivated AusE, are very similar and movement from one variety to the other seems easy (Bernard 1969:70). The question of the origin of the sociolects is, however, a very contentious issue.

Cochrane (1989) and Horvath (1985) believe that the sociolects, at least Broad and Cultivated were there from the beginning. Görlach (1991) also talks of two sociolects, although he remains cautious about the supposed sharp divisions between the two speech communities. According to Horvath, General developed later when social class barriers broke down in Australia and the speakers of the two original sociolects mixed.

This assumption is based on the view of early Australia as a sharply divided society, where brutally treated criminals could never be the equals of free immigrants. This clearly contradicts historical fact.

Borrie (1994:34f), for example, presents figures, collected in 1821 by an Emancipist Committee protesting against alleged attempts to restrict their rights, that show that they had twice the number of sheep and colonial vessels, three times more land under cultivation and four times more town houses than the free immigrants.

Of course there were different sociolects (as well as dialects) in early Australia. But it is also very clear that these sociolects do not correspond to today's sociolects.

Influential historians like Robson (1955), Shaw (1966) and Clark (1975, 1977) have depicted convicts as monstrous, lacking education, moral standards or the ability to love their children. However, a modern generation of scholars has shown a very different picture. Indeed, what Shaw, Robson and Clark have to say about the convicts seems to reveal more about them than about the transportees. Nicholas & Shergold (1988:5f) rightly state:

> For example, even though [Manning] Clark found that the transported criminals had surprisingly high levels of literacy, he argued that the criminal class was characterised by mental imbecility, low cunning and ignorance. The fact that the percentage of […] skilled urban trades people, was higher than the percentage of labourers and agricultural labourers combined, is ignored.
> In his 37-page analysis of 'Who are the Convicts?', A.G.L. Shaw barely mentions their occupational backgrounds. And the most thorough and careful quantitative study by Lloyd Robson displays a near total disregard for the statistical evidence on occupations. […]
> Much of the analysis of the convict system in Australia rests on two assumptions by historians; that the organisation of forced convict labour differed significantly from free labour; and that convictism was inefficient. Both assumptions have received unanimous assent; neither assumption has been explicitly tested.

Convicts were not treated too badly, divisions were not as rigid as post-convict era literary writings (such as the ones by Marcus Clarke in the 1870s and Price Warung in the 1890s and others) suggest (cf. Kociumbas 1992:257; Nicholas & Shergold 1988:11).

Australian society, including the convicts, was better educated than the overall population of the British Isles (cf. Cleverly, 1971:134; Jupp, 1989:555; Nicholas & Shergold, 1988:9).

As shown above, class distinctions were by far cry not as rigid in Australia than in Great Britain. So the sometimes encountered criticism of the rise of convicts and emancipists reflects more the fears of middle-class people who are not sure of their standing than the reality of early Australian society. This breaking down of class barriers is confirmed in the sources. John Maxwell, an Irish immigrant, writes in 1884:

> I saw M. Hawthorn [his social superior] today. He was telling me he had got a situation. [...]. He is very sociable here and stops and shakes hands with either Hugh or I when he meets us but Australia and the crossing of the line makes a great change on people's sociability. <4-076>

Another argument against Horvath's hypothesis runs as follows. How should it be possible that there were two or three sociolects from the beginning which all evolved in the same direction, becoming ever more similar, with no common model to aim at? Did Governor Macquarie (with his Scottish burr) or Governor Bourke (with his Anglo-Irish accent) speak Cultivated AusE or a proto-form of that? Certainly not. People of this class either left Australia, and thus had no lasting linguistic impact, or, if they stayed and made Australia their home, assimilating their language slowly towards an arising accepted standard. The examples of Macquarie and Bourke also show that even England's upper class was not speaking with a unified accent, well into the 19th century not even using a unified spelling or grammar, until the establishment of certain codes at public schools furthered the rise of RP later.

Bernard (1969, 1981; supported by Gunn 1972) claims that Broad AusE was there from very early on and social pressures to adapt the language led to the emergence of General and from that to Cultivated. Mitchell (1995:61) and others are extremely sceptical of such a course of development, since the social forces must have been extraordinary and yet only pertaining to a limited number of people. Moreover, phonologically nothing inherent Broad suggests a logical development into General.

Convicts, freed, free, soldiers, pastoralists and others, they all came from similar geographical and dialectal backgrounds. They all contributed to the emergence of Broad Australian and provided the language pattern that was then indigenised by the first generations of the native borns. Mitchell (1995:62) therefore claims convincingly that external influences must have brought about the existence of General alongside an already established Broad. He suggests that this influence can be found in the second wave of British immigration beginning in 1830 and greatly expanded in the gold rush period. The newcomers were simply too many to be assimilated completely. Cultivated, according to Mitchell (1995:63), developed later out of group choices.

The most convincing theory about the development of the sociolects of AusE has been put forth by Mitchell (1995). It is most in line with historical facts and reasonable interpretation of these.
Still, it is only a theory and only an empirical investigation could deepen our understanding of what happened exactly in the formation processes of AusE.
What has to be looked at is the actual existence of sharply divided social groups using linguistically definable sociolects.

2. Early Australian English – A New Approach

Most of the answers that have been found to the questions stated at the beginning are based on rational thinking. In many cases the quality of the answers has reached a level that leaves little or no room for dispute. But it would be wrong to stop here. Plato cannot stand alone. He must be followed and complemented by Aristotle's empirical approach, just as Aristotle historically came after Plato.

How can we investigate early AusE?

There is only one way to investigate early AusE. We have to look at historical instances of it in written sources. This greatly limits the range of possible studies. We cannot study its phonology in greater depth, because the data to be gathered from the sources are too few and far between to be able to serve as a statistically significant basis for far-reaching claims.

Of course, we can never be certain how the written records relate to actual language use in early Australia. Moreover, the existence of historical records is very much dependent on chance and hence it is not possible to create corpora with the strictness in design which is required for corpora of present-day English.

The data may be bad, and the art difficult, but there is no choice. If we want to investigate earlier forms of English, we have to look at written historical evidence. But there are, of course, many aspects to consider in order to find the best data and to make the best use of them.

Horvath (1985:26) claims that there is sadly no possibility to investigate convincing instances of early AusE. She says:

> Even if labourers wrote letters back home, it is unlikely that these were thought to be important enough either by the sender or the receiver to be saved for posterity.

In that she is clearly wrong and greatly underestimates the emotional sophistication and levels of literacy of 'labourers'. Only because they did not belong to the upper class does not mean that they did not value news from their relatives and did not think it worthwhile to preserve them. Even a short look into any major archive in Australia, for instance the Mitchell Library, or in the British Isles, e.g. the Public Record Office of Northern Ireland, tells a very different story.

During a conference on Celtic Englishes in Potsdam in 1998 Görlach raised doubts about the linguistic value of historical letters. He claimed that there were only the writings of country parsons who pin down a few dialectal spellings and misrepresentations of pronunciations. Otherwise useful data were not to be got. When read one of the letters used in the investigation for Fritz (2000b) he however readily admitted defeat.

The data are there. What we have to do then is to construct a corpus of early AusE. It must be built on such principles as to allow us to do rewarding investigations and to claim reasonable representativity for the results of these.

What sources can be used?

The choice of material is fundamental for any corpus building. Consider the following example:

> Stockdale, John, 1789, The Voyage of Governor Phillip to Botany Bay with Contributions by other Officers of the First Fleet and Observations on Affairs of the Time by Lord Auckland, edited by J.J. Auchmuty. Sydney: Angus and Robertson, 1970.

Although the above title seems an excellent starting point, the book itself reveals that its content is only a retelling of the original sources by the English editor. The material is not authentic and therefore cannot 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 much famed account of Australia, namely 'A Statistical, Historical, and Political Description of the Colony of New South Wales'. It was only in the year 1824 that he came back 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.

How were the Sources Selected?
Material to be included in a corpus of early English in Australia had to meet with a regional and a temporal criterion. The latter means that only texts written between 1788 and 1900 were used.

The required place of writing was Australia, New Zealand or Norfolk Island. But in a few cases, other localities were allowed. For example, if a person who was a native Australian or who had lived in Australia for a considerable time, wrote a shipboard diary, this was found to be acceptable. The same applies to people who travelled to other countries or published their memoirs in Great Britain.

Sample size was not an essential criterion. Although full texts were preferred, e.g. with letters, articles and speeches, this was not always possible.

Of course, as with every corpus, there is a sampling problem involved. What should the corpus be like in order to represent English in early Australia comprehensively?

Something that can help to achieve representativeness to some extent is the definition of genres.

So the definition of registers[2] and text types were of central importance for the creation of the corpus. It also contributes to the ultimate aim of an empricial and quantitative analysis of the evolution of Australian English. Without it only an investigation of the temporal pattern would have been possible, with it, we have a tool to look for different speeds of development in different registers. The definition of these was much influenced by the works of Biber (1988), Biber, Conrad & Reppen (1998) and Biber & Finegan (1992, 1997).

What Sources were Used?
The data for the corpus come more than 100 different sources. They cannot be named here in full, so some examples must suffice.

A number of private and official letters come from the Mitchell Library in Sydney, New South Wales, which holds a breath-taking amount of original documents relating to the history of Australia from its earliest times.

Another category is published material in book form. Many historians have striven to evidence the course of Australian history by editing official and private documents, letters, diaries,

---

[2] The term 'genre' designates a formally distinguishable variety of text, but it is also used to identify categories of artistic composition. 'Register', on the other hand, is a variety of language defined according to its use in social situations. The latter was preferred, because it emphasizes the differences between the categories chosen better, than the term genre.

proclamations, newspaper reports, legal texts, etc. Examples are the exceptional works of Manning Clark (1957, 1975, 1977), David Fitzpatrick (1994), Ward and Robertson (1969) and O'Farrell (1984).

A third category, and by far cry the most accessible kind of material, consists of historical texts published on the internet. Without doubt the most comprehensive and ambitious undertaking in that area is the *Setis* (The Scholarly Electronic Text and Image Service) programme.[3] Many of these are literary but there are also some historical texts. Examples of texts from the website are the works of Thomas Alexander Brown (Rolf Boldrewood), Marcus Clarke and Henry Lawson. Also, the complete Federation Debates (Melbourne 1890, Sydney 1891, Adelaide 1897, Sydney 1897 and Melbourne 1898) are to be found there.

Editing the Data

All sources had to be converted into computer text-files. For some this was easy (the internet documents), for some extremely hard (the original letters from the Mitchell Library).

Each computerized text received a heading which states its *Source Identification Number* (SIN) and provides data about the author and the source. The SINs are roughly assigned chronologically. They start with a number between 1 and 4 (for the period the document was written in) and then, after a hyphen, are given a three digit number for further identification.

Whenever a quote from the corpus is used, the SIN is given in pointed brackets, e.g. <1-093>. Due to considerations of space, it was not possible to print the full table containing all information about all the texts of the corpus. But all the information about the sources that do come up in this paper can be found in the Appendix.

The following data about the authors were collected (as far as possible):
  o name
  o year of birth
  o gender
  o country/region of origin
  o social status
  o year of arrival in Australia

The following textual properties were also noted down (as far as possible):
  o year of writing (or of publication)
  o place of writing
  o register of the text
  o text type
  o the number of words (as counted by MS Word 2000™)
  o where the source was found
  o the page numbers in the original text (if applicable)
  o gender of the addressee (if applicable)
  o status of the addressee (if applicable)
  o abode of the addressee (if applicable)

---

[3] Here is the hyperlink: http://setis.library.usyd.edu.au/ .

3. From Data to Corpus – Building Principles

Two principles were adhered to when building the corpus. First, there was a temporal criterion to enable diachronic comparisons, and second there was a register/text type criterion. The latter made sure that defined types of text were distributed evenly in each period.

The Principle of Periodization

The corpus material was divided into four 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 18th and 19th centuries. They are also in line with periodizations commonly 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 the native white population and urbanization (1876-1900)

The periods given do not only mark important changes in the history of white Australia. They also signify transitions in the linguistic history of early AusE.

Between 1788-1825, when Australia was primarily a penal settlement, proto-Broad AusE was established. Convicts and free immigrants came roughly from the same geographical and dialectal backgrounds and a recognisable colonial dialect was spoken by the native born early on. Yet their number was still very small.

After 1825 the number of free, often assisted immigrants grew considerably and soon outnumbered the convicts and emancipists. By 1841 the latter formed only 39% of the population (cf. Ward, 1958:16). The number of natives, and thus the number of speakers of a colonial dialect increased, especially since the sex-ratio, that had been very imbalanced in the era of the penal colony, improved rapidly (Borrie 1994:65).

The Golden Decade had a significant impact on the linguistic history of Australia. Between 1850 and 1860 the total population of Australia had risen from 405,356 to 1,145,585 (Greenwood, 1955:448 and Borrie, 1994:67)! Nevertheless, Broad AusE survived these mass migrations, since its speakers, though fewer in number, were by far the more stable element of colonial society (cf. Mitchell, 1995:23, 28, 35). Thus, although the Broad speakers did not manage to assimilate the newcomers completely, their language was modelled close on Broad AusE and can be thought of as the start of General AusE.

Historically and linguistically the years between 1875 and 1900 form the next important period. In 1861 37.2% of the population had been born in Australia. One decade later, this had risen to 53.5% and was climbing steadily (Price, 1988:8f). Natural increase 'leapt to over 70%' (Mitchell, 1995:28) already between 1861 and 1870.

This 'nativization' of the Australian population went hand in hand with the final establishment of AusE. It does not come as a surprise that it was in the 1890s that the Bulletin school voiced its opinion on the vigour of local writings and that Australia was finding a sense of nationhood, which finally manifested itself in 1900-01 in the creation of the Commonwealth of Australia.

The Principle of Register

The second principle stated that in every period (1788-1825, 1826-1850, 1851-1875 and 1876-1900) there should a like number of words in the different registers:

Table 1: Number of Words in each Register

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

The Public Written (PcW) register dominates the corpus, and rightly does so, since these writings were most widely distributed and certainly made up the lion's share of Australia's linguistic scene.

Next comes the Private Written (PrW) register, which also is quite large. This represents the thousands of letters and diaries in which almost everybody confided private joys and sorrows. The Speech-based (SB) register is comparatively small. This is certainly not representative of total 'production' of English in 19th century Australia, but is due to a lack of sources. By far the smallest register is Government English (GE). For this we have to bear in mind that GE was used only by a very restricted number of people in clearly defined situations.

Proportional sampling was, of course, not possible and, according to Biber, Conrad and Reppen (1998:247f), is not always a good solution anyway.

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

SPEECH–BASED TEXTS

This register come as close as possible in a historical corpus to the spoken English in early Australia. It differs greatly from the others, even if the writing down of them changed the language to a greater or lesser extent (Halliday 1985:41f).

Minutes (MI; for example testimony in court, parliamentary proceedings, etc.), *Speeches* (SP; addresses and sermons) and *Plays* (PL) are the text types to be found in this register.

Most of the *Minutes* come from parliamentary debates, but there are also some from courtrooms. Although we cannot observe spoken language in the same way as with a tape recorder, this is certainly the closest approximation that we can get.

Speeches can also be drawn from debates, but also from newspapers where they are reported. Sometimes speeches were even published in separate books. This text type has some characteristics of spoken language, but is planned and largely a monologue, which makes it similar to written language, too.

There is much debate about the value of the language as used in *Plays*. Biber and Finegan (1992:701), however, having thoroughly studied dialogue as represented in fiction and in plays conclude that "authors capture the linguistic characteristics of conversations fairly accurately with respect to Dimension B (Elaborated Reference) and Dimension C (Abstract Style)."

PRIVATE WRITTEN TEXTS

The language used in this register is often very informal, intimate and relatively unmonitored. Of all written registers, this is the one closest to spoken language. There are only two text types in this category, namely *Private Correspondence* (PC) and *Diaries* (DI; diaries, journals).

Astonishingly, there are still scholars who deny the value (Görlach: personal communication) or the bare existence (Horvath 1985:26) of linguistically interesting *Private Correspondence*, despite ample prove to the contrary in linguistic (e.g. Fritz 1996 and 1998, Arnaud 1998, Palander-Collin 1999) and historical studies (e.g. Fitzpatrick 1994, O'Farrell 1984).

Personal letters written home make excellent linguistic sources, because they show variable usage between native and non-native elements. Australian terms are often glossed to make them understandable for the family at home.

So we have, for instance, John Maxwell writing home in 1883:

"Hugh Sheils is working to Alexander Perry at ringing trees. I suppose you would not know what ringing trees is if I did not tell you. It is to take about 6 broad of bark and the last spring's growth off right round its circumference." <4-065>

Personal letters are not an art form to most writers in nineteenth century Australia. They are a vital means of maintaining contact with home. That not too many were proficient in this can be seen in the following quotes (see also Fitzpatrick 1994:473):

Michael Normile jr. in 1860: "Actualy my Dear Father I fancy I am speaking to you verbaly while I am writing this scroll to you but my grife I am not." <3-213>

There is a heart-warming comment by the unknown scribe who lent his, clearly uneducated, hand to Richard Dillingham in 1838:

"there Is one thing that I the writer this Letter have to State to I daresay your Great Sattisfaction that Is your Son Richard I can say Is Verry [illegible] Steady for though Liquor Is verry Cheap And he have the means

and the opportunity of Geting it he Never gets more than do him Good Rum is 2s the Pint and that is the Cheif that is drank Wine is verry Cheap It is 9d the Pint and the Cape Wine is 7d the Pint the distance from England to Vandemans Land is 18000 Miles and It is day here when it is Night in England" <2-138>

Diaries are sometimes the most intimate correspondents of people and the language used in them does not have to be adapted to another person or to an outward norm. They were often used as confidantes and comforters, something which many were in need of in early Australia.

Diaries can show the most intimate of conversations, namely with oneself, but can be intended for later publication or documentation, this was often the case with explorers' diaries.

PUBLIC WRITTEN TEXTS

These are written for the general public or to some person who was not a friend. This makes the language more formal and monitored, although sometimes informal passages can be included (e.g. dialogue in literary texts). The text types that belong to that register are the following: *Memoirs* (MM), *Newspapers and Broadsides* (NB), *Narratives* (NV; novels and short stories), *Official correspondence* (OC; letters to office bearers, letters to the editor, pastorals and business letters), *Reports* (RP; histories, accounts, statements) and *Verse* (VE).

The text type *Memoirs* was very popular in Australia. In the early years they told prospective emigrants what they could expect to find in the Antipodes, whereas later examples were mainly written for the local market, remembering the glorious days of early settlement and exploration.

Newspapers were established early in the colony (Goodwin 1986:11f), the first being the *Sydney Gazette* (1803-1842) followed by *The Derwent Star* and *Van Diemen's Land Intelligencer* (1810).

How important the printing-press was in Australia can be seen in the following extract from David Mann in 1811:

> The art of printing had been gradually improving from the period of its establishment, by the judicious care of Governor Hunter, and its advantages became daily more and more obvious. On the 5th of March, "The Sydney Gazette" was instituted by authority, for the more ready communication of events through the various settlements of the colony The utility and interest of such an establishment were speedily and universally acknowledged; and its commencement was soon succeeded by the publication of an almanack, and other works calculated to suit the general taste and increase the general stock of amusement. <1-126>

Broadsides, on the other hand, not inhibited by the heavy burden of the stamp tax, came even earlier (cf. Ingleton 1988), dating from virtually the first years of settlement.

The term *Narrative* covers short stories and novels. This text type is somewhat problematic, since it often contains monologues and dialogues to some extent. Therefore, extracts can be closer to the *Speech-Based* register (cf. Biber & Finegan (1992) and Haan (1996:23)), although the texts as a whole certainly belong to the *Public Written* register. Investigation of this kind of source has to keep the dual character of such texts in mind. Taylor (1997) and Ramson (1972:38) are very critical of literature as revealing the 'true' linguistic picture, but this looks at AusE only from a very special angle. If we accept that every text written in Australia is part of the English produced in Australia and thus a possible contributor to early AusE, then we must count every instance of it. It may be the case that the author was trying to potray a certain dialect and missed it. Nevertheless, his English is evidence of potential usage in Australia at that particular time.

Official Correspondence are letters written for business or other official purposes, sometimes requesting a favour from an office bearer. As such they can transcend the status boundaries that normally kept people from different classes apart, even if such contact was infrequent and largely one-way.

The text type of *Reports* is similar to that of *Memoirs*. But, whereas *Memoirs* talk of the experiences of one's own life, *Reports* are writings that describe a situation without reference to the life of the author. One of the most famous texts is certainly the influential, if sometimes inaccurate, *Bigge Report* <1-198> dating from 1822, where the conditions in the colony of NSW are described. With the help of this report, the Colonial Office wanted to find out, if transportation and the penal system was still an effective deterrent to crime. This uncertainty is most telling and contradicts the mostly literary presentations of convict life as a series of hardships (cf. Baker, A., 1984:2-3).

The last text type in the *Public Written* register is *Verse*. This is a highly specialized writing and

hardly typical of everyday language. It is included only, because it also forms part of the body of early English in Australia. Moreover, its artifical structure only states more openly what other text types try to hide, namely that written language mostly involves a deliberate choice of grammar and words.

GOVERNMENT ENGLISH

The language used in such texts is formulaic, explicit and very close to the one written all over the Empire in this context. The writers of such texts had seen service not only in Australia, but also in India, South Africa and elsewhere.

> But he will be absent for eighteen months or two years during which time if I am appointed Chief Justice pro tempore something more permanent may offer for me but at all counts the addition of five hundred pounds per annum to my income for that period will make up for my losses at the Cape and enable me to pay all I owe, and by that time my ten years will nearly have expired and I shall be anxious to revisit England. (W.H. Burton in 1834, <2-092>)

Therefore the language to be found in these is most resilient to change and shows the least level of nativization throughout the period under investigation. *Imperial Correspondence* (IC; letters written among office bearers), *Legal English* (LG; laws, verdicts, grants, contracts, regulations) and *Petitions and Proclamations* (PP; petitions, resolutions, addresses to government bodies, proclamations, resolutions, official recommendations) are the relevant text types.

The third type of letter under investigation is *Imperial Correspondence*. These are letters written by officials (civil servants, the military) to other officials. Legal opinions by lawyers or judges that were addressed to officials were also included in this category.

Correspondence of that kind is very explicit and influenced by the linguistic experiences of their writers who may well have served in many different parts of the Empire.

Legal English has been described as having a 'formal' or even 'frozen' style (Hiltunen 1990:65). Mellinkoff (1963) and Hiltunen (1990) have described the language of the law extensively and come to the conclusion that it is full of archaisms, uses uncommon lexis and structures, and is not at all prone to change. Functionally this is explainable (Halliday 1985, Halliday and Hasan 1985) due to the need for legal language to be explicit and valid for long periods of time. Biber & Finegan (1997: 273), however, after a lengthy study, show that legal English did indeed change to some extent from more oral to a more literate style.

The last text type defined is *Petitions and Proclamations*. In fact, there is an inherent discrepancy since a *Petition* is addressed *to* an official, whereas a *Proclamation* is issued *by* an official. Despite this, it was considered valid to group both kinds of texts under one heading, since the language used was very similar, being conditioned more by established formulae of expression than by other considerations. Also included are resolutions passed by non-government bodies, e.g. patriotic associations.

4. A Description of a CORPUS OF OZ EARLY ENGLISH (COOEE)

COOEE has taken about seven years to complete. Work on it started in 1995 and finished, as much as any work on a corpus can finish, in 2002. Admittedly, efforts in building COOEE were intermittent due to my full-time teaching obligations at a German secondary school. But singly-handed working slowed progress a lot, too.

COOEE 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).

This early corpus comprised 143,565 words in 359 letters, diaries and various excerpts. Today, the total number of words collected is **more than ten million** (ca. 70 times larger than the original corpus). This means that the overall database is much larger than that used for Brown, LOB, ACE, ARCHER or CONCE. Its size is even twice that used for the acclaimed *Longman Grammar of Spoken and Written English* (Biber, Johansson, Leech, Conrad & Finegan 1999).

However, a corpus is much more than an unprincipled collection of texts. In order to gain argumentative strength its make-up had to follow the principles outlined above. This, of course, reduced the number of words actually part of COOEE greatly.

Altogether, **COOEE**, based on the principles of periodization and register, **comprises ca. 2 million words in 1356 texts**. The remaining eight million words are not lost but make up several reference corpora (see below).

The sources are of very uneven length (mean length: 1,481 words), 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 of COOEE.
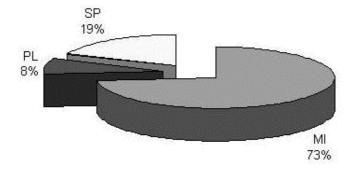
Register and Text Type

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

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*, a program from the *WordSmith Tools* program suite. The latter is used to provide all of the following data about the individual parts of COOEE, i.e. types, tokens, type/token ratio, mean word length, mean sentence length and keywords.

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 insight into actual spoken language than the other two text types.

Table 2: Ten most positive Key Words[4] in the SB-Register

1. *plaintiff*	2. *witness*	3. *prisoner*	4. *jury*	5. *defendant*
6. *question*	7. *hon*	8. *attorney*	9. *would*	10. *case*

Table 2 clearly shows the court character of the texts sampled in this part of COOEE. The prominent place of *would* which characterizes the hedging strategies of much of the dialogue recorded is of interest.

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.

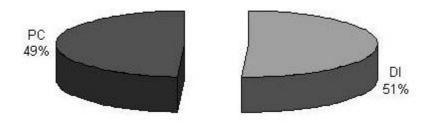Figure 2: Personal Communication (PC) and Diaries (DI) in PrW register

Table 3: Ten most positive Key Words in the PrW-Register

1. *miles*	2. *water*	3. *dear*	4. *north*	5. *west*
6. *camels*	7. *east*	8. *camp*	9. *good*	10. *got*

Travelling, mostly *north, west* and *east*, but not south, is prominent here. Distances are counted in *miles*, *water* is a prerequisite of life and *camels* (!) are often used to get to the next *camp*. There letters were written that started with *Dear* ….
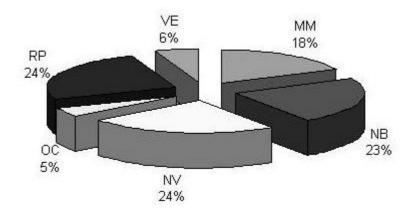
THE PUBLIC WRITTEN REGISTER

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, 793,457 words were included.

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

---

[4] The key words were calculated by the *KeyWords* program from *WordSmith*. In order to find the key words of a part of a corpus, or of a reference corpus, its wordlist was compared with the wordlist from all of COOEE. To compute the "key-ness" of an item *KeyWords* computes its frequency in the first wordlist, the number of running words in the first wordlist, its frequency in the reference corpus, the number of running words in the reference corpus and cross-tabulates these.

Positive key words are those that appear more often than expected, negative key words are those that appear less often than expected. The key words displayed here have a maximum p value of 0,0000000001.

High frequency words like *the, of, and, to, a, an,* etc. were excluded from the wordlists before proceeding to the keyword search. Proper names were excluded after the search.

The distribution of the text types over time is likewise not even. For example, there is no narrative in the first period. So a higher number of reports had to be included for that 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.

Table 4: Ten most positive Key Words in the PcW-Register

1. *dad*	2. *convicts*	3. *eyes*	4. *while*	5. *station*
6. *captain*	7. *among*	8. *sheep*	9. *head*	10. *himself*

Most surprising is the top position of *dad* in texts that are not intimate. The mystery is solved when looking at the actual instances. They come mainly from narratives which 'pretend' to record intimate conversation. *Convicts*, *sheep*, *heads* of cattle and *stations* are the main topics.

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 at 17%.

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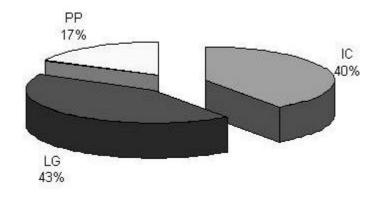

Table 5: Ten most positive Key Words in the GE-Register

1. *shall*	2. *court*	3. *act*	4. *council*	5. *colony*
6. *majesty*	7. *supreme*	8. *said*	9. *may*	10. *judge*

Shall and *may* indicate prescribed uses of modals, whereas *supreme court, acts, councils*, *colonies*, *majesty* and *judges* are the main issues mentioned.

A COMPARISON OF REGISTERS

The different registers are now compared with each other regarding the number of different types, the standardized type/token ratio[5], mean word and sentence length.

Table 6: Types, Tokens and Mean Lengths

	types	type/token ratio	mean length of words	mean length of sentences
SB	15151	37,89	4,44	29,55
PrW	24999	41,01	4,17	25,66
PcW	30237	44,15	4,47	29,89
GE	9130	33,72	4,71	51,28
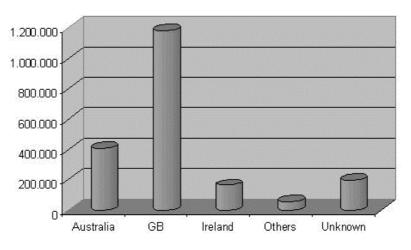
The number of different word types seems significant, but the differences in sizes of the various parts of COOEE invalidate such a reasoning. A standardized type/token ratio is the only viable option for a comparison. Thus we find that GE has the least number of different words available. Obviously GE texts are carefully worded, especially legal texts, in order to achieve unambiguous communication over great temporal and spatial distances.[6] The fact that the texts in SB also contain comparatively few different words hints at their authenticity. Spoken language generally uses a smaller set of words that written language, so this finding from COOEE shows the texts in SB to be more like spoken than written language. Unsurprisingly, the highest variation is found in the PcW register. This reflects not only the literary ambitions of the authors, but also their careful choosing of the right words.

The GE register also stands out because of the greatest mean length of words and sentences. This is certainly due to greater numbers of words of Latin, Greek or French origin and to the syntactical structures of official and legal texts as they were explained above. The PcW register does not differ too much from the other two registers in this respect. This can be explained by the large portions of narratives and memoirs that aim at imitating spoken and/or informal language. In these two categories they succeeded, only the type/token ratio betrays them somewhat. The comparatively low mean sentence length in the PrW register can be attributed mainly to the diaries that often do not contain complete sentences at all, but evidence sometimes rather idiosyncratic stops.

Origins of the Authors
The next figures show us where the authors of the sources for COOEE came from.

Figure 5: Origins of Authors (all)

---

[5] The standardized type/token ratio is computed every 1000 words by *Wordlist*. In other words, the ratio is calculated for the first 1,000 running words, then calculated afresh for the next 1,000, and so on to the end of a text or corpus. A running average is computed. This method ensures that type/token ratios are comparable for copora of different sizes.
[6] Cf. Halliday and Hasan (1985) for such a need and the structure of legal texts as explained above and in Mellinkoff (1963).

People hailing from Great Britain wrote most of the sources (1,160,619 words), but there is also a substantial amount of native white Australians (400,534) and 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.

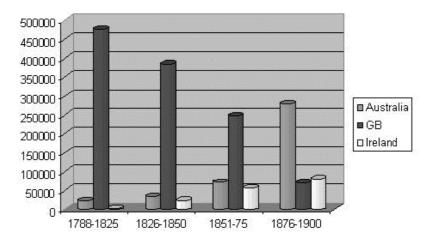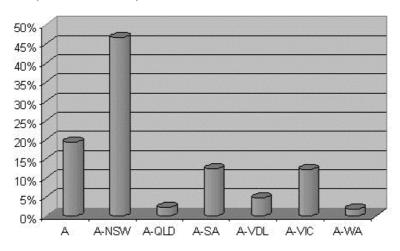Figure 6: Origins of Authors (Great Britain, Australia, Irish) divided into periods

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%).

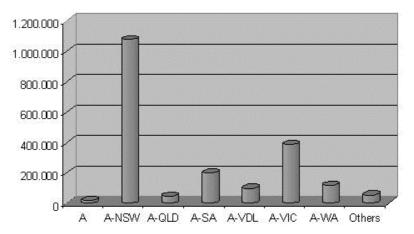Figure 7: Origins of Authors (native Australians)

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 (Tasmania since 1859).

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.[7] Otherwise the regional distribution would have been skewed by historical names, e.g. if a text written at Port Phillip would be counted as coming from NSW.
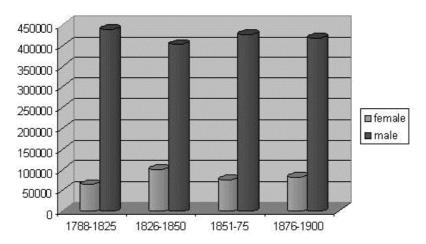
Figure 8: Place of Writing

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.

Gender of the Authors

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. This does not differ much in the four periods looked at, only between 1788-1825 there are considerably fewer writings by women than in other periods (12.6%).
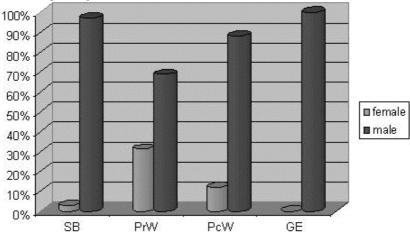
Figure 9: Gender Distribution over time

The number of different word types for women is 18,306 which leads to a standardized type/token ratio of 42.64. The mean length of words is 4.17 characters, that of sentences 28.07 words. The number of types for men is 38,749 (among 1,666,578 tokens) which leads to a standardized type/token ratio of 41.0. The mean length of words is 4.42 characters, that of sentences 29.48 words.

---

[7] For example Victoria (1851) and Queensland (1859) were separated from New South Wales as new states only well into the nineteenth century. Today's Northern Territory was a part of South Australia until 1901 and the Australian Capital Territory was formed in 1926.

In this respect, texts written by women do not differ much from those written by men. However, we must bear in mind that authorship of sources is very unevenly distributed in the different registers as Figure 10 shows. The above mentioned statistics, however, are not significantly changed by leaving out all the texts from the SB and the GE registers as research on COOEE has shown.

Figure 10: Authorship of Texts per Register

Women did not write any GE and next to none (3%) SB texts. Their overall contribution of 16% stems from the PcW (12%) and mostly the PrW (31%) register.

If we compare the texts written by women with those written by men, the following keywords can be found:

Table 7: Ten most positive Key Words in Sources Written by Women

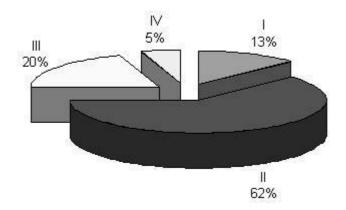| 1. *dear*	2. *Mrs*	3. *husband*	4. *see*	5. *do*
6. *Miss*	7. *tell*	8. *quite*	9. *Papa*	10. *aunt*

These key words provide us with very clear and detailed insight into the writings by women in COOEE. They write to relatives, *husbands, fathers, aunts,* keeping up family communications and *telling* what they *see* and *do.*

Status of the Authors

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

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

Figure 11: Status Distribution (general)

The distribution of the 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. This is illustrated by the following excerpt where William Dysart writes from Victoria to his native Ireland in 1877:

> I do not know how to begin to write after such a long silence, and the longer one is without writing the more difficult it is to begin, but though so long silent, there was none of you ever forgotten for a single day. The reason I did not write from the first was I had bad luck from the very first and as I was always in the hope of something turning up and then I would have written with some heart. <4-007>
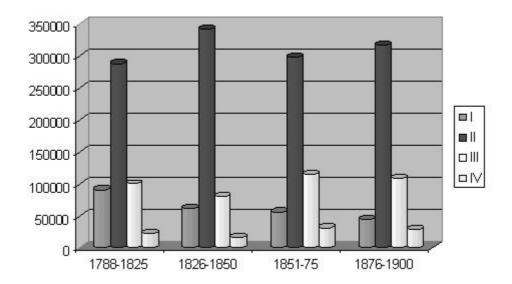
Even more elusive are the letters and diaries from the lowest social classes. They have been preserved but seldomly and only meticulous historical work like that of Webby (1989), Fitzpatrick (1994) and O'Farrell (1984) has brought some of their writings to light.

Bridget Burke writes from Queensland to Ireland in 1889:

> My Dearest father & mother
> I for once in 12 mounths sit to have a few words of conversation with you by a message which I must say is my hardest Job to get through. I got Your letter about 10 days ago which I often wanderd ware You dead or what became of You. As for my Brother & sisters I Quite forgive them as they have got children of their owne to [?look two] & me a child or a lost lam far away from home & nation. To think my father & mother at the end of a long year could cast a thout on me & wright me a letter. Wasent it a chearing present to a true hart from the Dearest frind. <4-182>

Figure 12: Status Distribution over time

As can be seen in Figure 12, the contributions of writers of different status did not change much over time. However, a slight decrease in status I people and a corresponding increase of status III and IV people can be discerned.

5. Methodology – Using the Corpus

The Policy

The aim pursued is the investigation of the origins of Australian English. When looking for the origins of AusE, we must look at Australia and the kinds of English spoken and written there from 1788 onwards. In this context, it is of no particular interest that *to fossick* comes from a British dialect, possibly from Cornwall. The only question of relevance is: When and how does this word enter Australia and how does it fare there? Who uses it and how often can we find it?

Likewise if *ain't* is found to be a feature of English in early Australia, then criticism of the kind "How can he say that *ain't* is a feature of early AusE since it is clearly a feature of so many English dialects?" is unacceptable.

The aim is not to highlight only the things that are peculiar or even unique to Australian English. This 'freak show' approach does not give due credit to the subject.

So what constitutes proper areas of study? Lexis, morphology and syntax that is <u>important</u>, in terms of frequency or sociolinguistic usage, in historical periods of AusE and/or today.

The Issues

Early AusE and the use of English in early Australia was very much marked by a unique lexis. These lexical items, their occurrences and frequencies can be studied in relation to social factors:

- status, convict or free, gender, origin, year of arrival, age, length of stay in Australia
- differences between newcomers, established immigrants and natives
- for letters: gender of writer and addressee, status of writer and addressee, origin of writer and abode of the addressee

It seems also worthwhile to take textual properties of the sources into consideration. For example,

- year of writing, place of writing
- register of the text, text type

But modal usage, spelling traditions and others can be investigated in the same vein. For lists of possible areas of study see Newbrook (1989, 2001).

Intra- and Inter-Corpus Studies

When studying the language of a corpus it is always important to bear in mind the following question: What is the point of comparison?

We can restrict ourselves to intra-corpus investigations and try to ascertain developments that may or may not be relevant for today's Australian English. Dead-ends can raise equally important questions as ongoing pathways.

Another possibility is to relate the results of corpus research to findings on present-day AusE, BrE, IrE or AmE or to relate them to contemporary BrE, IrE or AmE.

Findings from COOEE can and should be compared with contemporaneous evidence from other corpora or at least from historical descriptions of other varieties of English. Sadly, such descriptions, not to mention proper corpora, of nineteenth century varieties are not plentiful. Possible are ARCHER, CONCE and Bliss (1979) for early Irish English.

Note that in collecting material for COOEE a great number of additional material was collected that can serve as a point of comparison.

These are:
- about 8 million words of contemporary AusE (1789-1900)
- about 250,000 words of post-1900 AusE (1901-23)
- about 950,000 words of contemporary BrE (1768-1886)
- about 47,000 words of contemporaneous IrE (1791-1905)
- about 29,000 words of contemporary NZE (1840-1860)
- about 34,000 words of contemporary SAE (1797-1862)
- about 33,000 words of contemporary CanE (1801-47)

6. A Progress Report on Lexis

Research using COOEE is currently carried out and will be published in a doctoral thesis under the title *From English in Australia to Australian English – the first 100 years*. Lexical fields, spelling variables, modal usage and morphology are among the areas to be covered.

The chapter on Aborigines and AusE lexis can already be presented in this paper.

General Remarks

When E.E. Morris (1896) compiled his *Austral English: A Dictionary of Australasian Words, Phrases and Usages, with Those Aboriginal and Maori Words Which Have Become Incorporated in the Language and the Commoner Scientific Words That Have Had Their Origin in Australia,* it was one of the first attempts to codify certain words that seemed sufficiently different in their meaning to warrant separate treatment in a book.

Although Baker (1966:10) is highly critical of this publication it is still a very interesting indicator of a need felt very much by Australians at that time, a fact acknowledged by Turner (1994:13) and to be found in COOEE. Philip Muskett, for instance, writes in 1893:

> It has, however, been reserved for Australia, strange even from the first, to prove an exception to this universal law. Yes, strange even from the first! For did not the earliest arrivals find that the seasons came at the wrong time of the year; that Christmas-tide came with sunshine, and that the middle of the year was its coolest part? Were there not found in it curious animals, partly quadruped, partly bird, and partly reptile? Were there not discovered, also, other animals who carried their young in a pouch? Moreover, did Dot these first settlers see that the trees shed their bark, and not their leaves; and that the stones were on the outside, not the inside, of the cherries? <4-291>

Australia is a continent with a unique fauna and flora. Everything the first convicts and settlers saw and encountered demanded a new name making an expansion in the vocabulary absolutely necessary. For the same reason, many words they brought with them, for instance *meadow, copse, thicket, brook and village,* fell into disuse.

Newcomers to Australia would therefore find themselves in a natural and also in a social environment that differed very much from their previous everyday experience. The need to talk about new concepts and to designate referents that had no existence in the home country was so urgent that an Australian lexis developed quickly.

This functional requirement was reinforced by the social conspicuity the use of some items involves. For instance the use of the word *paddock* to designate a small field where horses are kept would not only cause an almost certain misunderstanding but probably would also draw rather embarassing guffaws from an *old hand*.

Thus, John Maxwell's remark on his use of an Australian term in a letter to his brother is partly an apology and partly an explanation:

> "I have to *graft* just as hard as him - there is get out of it (please excuse slang for you will pick it up in spite of yourselves)." <4-160>

The Languages of Aborigines and their Contribution to early AusE

The words to be studied were decided upon by not-computer-aided readings of the sources and by inclusion of the word lists found in publications like Ramson (1966), Baker (1966) and others.

The latter procedure provided the chance of testing and quantifying the usages claimed by various authors as being current in early Australia. In this way, a non-finding of a word can be as significant as the finding of one. Therefore all the words looked for are mentioned in the text, even if there were no appearances of them in COOEE.

James Cook and his crew were the first to come into closer contact with Aborigines in eastern Australia. A vocabulary list was compiled that was intended to be useful for future meetings. They had, however, been unaware of the linguistic diversity of Aboriginal languages and the list proved to be more of an obstacle than a help.

One of the most fascinating lexical histories is that of the word *kangaroo*. The 'First Fleeters' used it in application to a number of different marsupials, a use Cook had recorded in his list. The Aboriginal people of the Port Jackson area had no knowledge of this word and obviously thought it

was an English word. This misunderstanding is testified to in different writings. We have, for instance, Newton Fowell, who writes home in 1788:

> I forgot to mention in the proper place that when the man was taken he denied ever seeing the Cowes even to the last & we have not seen them since, nor do we ever expect to hear of them again. this is a very great loss to the Colony as we must go a long way before we can replace them. either to the Cape of Good Hope or the Philipine Islands"; I am of opinion the Natives have killed them as once before they threw a spear at one of them calling them Kangooroo at the same time. <1-011>

Watkin Tench, whose two accounts of the early settlement contain most valuable insights, relates:

> Whatever animal is shewn them, a dog excepted, they call kangaroo: a strong presumption that the wild animals of the country are very few.
>
> Soon after our arrival at Port Jackson, I was walking out near a place where I observed a party of Indians, busily employed in looking at some sheep in an inclosure, and repeatedly crying out, Kangaroo, kangaroo! As this seemed to afford them pleasure, I was willing to increase it by pointing out the horses and cows, which were at no great distance. But unluckily, at the moment, some female convicts, employed near the place, made their appearance, and all my endeavours to divert their attention from the ladies became fruitless. <1-015>

As late as 1847, John Dunmore Lang was able to reconstruct the linguistic history of early Aboriginal-white interactions as his following comments show:

> A considerable number of words of the Sydney Aboriginal dialect known to the convicts or other white persons, in the earlier period of the penal settlement in Moreton Bay, were naturally enough made use of in attempting to hold communications with the black natives. These words, which were quite as unintelligible to the natives as the corresponding words in the vernacular of the white man would have been, were learned by the natives, and are now commonly used by them in conversing with Europeans as English words. Thus, *corrobory*, the Sydney word for a general assembly of natives is now commonly used in that sense at Moreton Bay; but the original word there is *yanerville*. (quoted from Ramson 1966:110)

WEAPONS AND IMPLEMENTS

A number of words of Aboriginal origin can be found when looking at weapons and tools used by the natives. Considering the size of COOEE the total number of words in this category is rather small.

The most famous, the *boomerang*, is found only 22 times in the corpus. The term *corvak* for the same thing is used once only by Alex Crawford in Western Australia in 1882 <4-030>, the Nyungar (WA) word *kylie* does not come up at all.

Different kinds of clubs feature more prominently. The *waddy* is used 28 times, both as noun and verb, the *wom(m)era*, a Dharuk word, twice. Interestingly, by 1875 this word from Dharuk, an Aboriginal language in NSW, had travelled to Western Australia to appear in the diary of John Forrest, a native of this young Australian state. *Nulla nulla*, another Dharuk term, appears in verse by Charles Harpur <2-292> and in the collection of Aboriginal tales by Langloh-Parker <4-381>. *Leangle/langeela* is there twice in Victorian writings, once in the poetry of George McCrae <3-250>.

> Blood to be spilt in future fight.
> The long leangle's nascent form
> Forespoke the distant battle-storm;

Spears do not appear to have been assigned Aboriginal names often by whites. But the implement that helped to propel spears, was obviously strange enough to warrant such treatment. This was the *woomera* (four instances).[8] Its first appearance in COOEE dates from 1829. Like the *womera*, this Dharuk word travelled from its native NSW. John Batman, one of the pioneers of white settlement in Victoria uses this word there in 1835 <2-131>, although it was most probably not originally used by the Aboriginal population in this state. By 1851 it had already reached Western Australia, where we find it again in 1882. Thus only one out of four instances is actually from the original place. Then this word journeyed with the settlers all over Australia.

---

[8] This spelling is from the AND, in COOEE only the spelling with a single *'o'* can be found, thus merging it with *womera*, the already mentioned club.

The Wemba word for a shield, *malka* (also *mulga* or *malga*), is used six times. Five times it is used in poems by McCrae (!) <3-250>, only once (in the spelling *mulga*) in a different register, namely in the memoirs of William Thomas <3-119>.
The Dharuk word is *hieleman,* which is there once in the spelling *ëlamán*.

There are hardly any Aboriginal implements mentioned in COOEE. The *coolamin*, a hollowed out knot of wood for carrying water, comes up twice in the diary of the Calvert Expedition <4-330>, but nowhere else. Neither *dilly*, a small bag, nor *mogo*, a stone hatchet, nor a *wirri*, a shallow wooden scoop, are there. The only thing that does come up are *yam-sticks* (five times), used for digging up the edible roots of plants of the *Dioscorea* species. Although the word *yam* itself is not Aboriginal, all the references to yam-sticks are applied to Aboriginal women using this tool.

> One old and hideous hag, in particular, dabbed her yam-stick into the ground dramatically, <3-078>

> The lubras fished up crawfish from the shallow muddy water-holes with their toes and yam-sticks, and exchanged them for the dainties of civilized life.. <3-078>

> She [an Aboriginal] was not over fastidious about her wardrobe, which consisted of a yam stick. <4-369>

> with which to chop out the bees' nests and opossums, and with yam sticks to dig up yams. <4-381>

> The wives, taking their goolays and yam sticks, went out as he told them. <4-381>

DWELLINGS
The names of Aboriginal dwellings are numerous and come from many different languages situated in different parts of Australia. Most of these have travelled to some extent and entered into figurative use for temporary shelters or ramshackle or dilapidated buildings.
The Dharuk *gunyah* comes up ten times, earliest in NSW in 1830 (<2-040>; twice; once in the combination *giba gunyah*, i.e. a shallow cave used as a dwelling), then in 1835 in Victoria (<2-131>; three times), in NSW in 1847 and 1851 (<2-324> as *ghibber-gunyah*; <3-032>), again in Victoria in 1853 and in 1860 (<3-073>; <3-201>) and finally in Queensland in 1865 <3-242>. This makes *gunyah* the most widespread word for an Aboriginal dwelling.
The reason for words of Dharuk origin to be widely known in Australia is that the whites encountered this language very early in the Cumberland Plain between Parramatta and the Blue Mountains in New South Wales and took the words from this language with them to other places, giving Dharuk words a considerable headstart. However, the speakers of Dharuk did not fare well at all. Almost all of them were killed in a long struggle to protect their lands from white settlers.
In 1853 *gunyah* is first used as a term for temporary shelter for whites:

> I may add that I was the first person who discovered Gippsland, […]. We had not even a tent, but used to camp out and make rough gunyahs wherever we remained for the night. <3-073>

Humpy, derived from the Jagara (QLD) word *yumbi*, is found four times in NSW in 1897 <4-381> and once in Western Australia (!) in 1898 <4-388>. There it is used for one's own dwelling:

> I have been helping Geoff but have finished with him now and am back in my own little humpy, running the wire round my block.

Nyunnoo is a NSW term that only appears once in COOEE and is not mentioned in the Australian National Dictionary (AND, 1988) . Therefore a full quote is given here:

> Looking round they saw their tracks going to where the emu had evidently been; then they saw that they had dragged the emu to their nyunnoo, which was a humpy made of grass. <4-381>

Mia mia (or *miam*) is at eleven instances the most frequent of all, and is from the Victorian languages of Wathawurung and Wuywurung. All tokens are in fact to be found in Victorian writings. Most use of the term is made in the memoirs by Victorian pioneers (written in the 1850s) and edited by Bride (1898). The only other source that mentions *mia mias* is the diary of the fatally failed Burke and Wills expedition (1860; <3-201>). Twice this term is used for buildings in which whites live which hints at the acceptability of this designation, even if *mia mias* are agreed to be rather rough buildings.

I joined A. F. Mollison in 1838. We lived in reed mia-mias and tents comfortably enough for some time. The winters were much drier. <3-077>

I had tame emu chickens performing their strange juvenile antics round my reed mia-mia - yellow-striped and downy little objects, difficult to be recognised as the sources from which future mature emus were to grow. <3-078>

Wurleys are originally found in South Australia (from the Aboriginal language Gaurna) and appear seven times in COOEE. All of these are to be found in only two South Australian sources, namely <3-238> (1865) and <4-330>/<4-369> (1896/97). None of the instances describe *wurleys* as places for whites.

Ceremonies, Daily Life, Pidgin and Creole Terms

Not many words describing daily Aboriginal life can be found in COOEE. Neither *baal* (a negative), nor *bora* (a rite of initiation), nor *budgeree* (good), nor *cobbra* (head), nor bogey (to dive, to bathe), nor *pyalla* (to speak). nor *walkabout* nor *yanerville* (= *corroboree*).

One of the evil spirits believed in by Aborigines in Western Australia is the *gingie.*

They have no religion nor idea of God they are afraid of an evil spirit they called Gingie and they have long flat pieces of wood with a hole at one end for a string this they whirl round their head and it makes a curious moaning sound and they think it keeps off the Gingie. <4-030>

There are two words of Dharuk origin that show a respectable presence in COOEE, namely *cooee* (17 times; Dharuk *guwi*) and *corroboree* (32 times, Dharuk *garaabara*).

Originally, *cooee* is a call used by Aborigines to communicate at some distance. In this meaning it was early adopted by the white settlers. Only six times an Aboriginal call is named a *cooee*, but an astonishing eleven times it applies to a white person, earliest in 1841, although the example is not totally clear. Above that, the term receives an additional explanation in brackets:

At 2 O'clock yesterdy morng we were awoke by a loud coo-ey (the sound always employed here to be heard at a distance): on listening, we perceived the bleating of the lambs was no longer to be heard; they had knocked down the hurdles & got out. All the men ran, nearly without cloaths, & in about an hour they were all safe again. <2-238>

The word is found in COOEE in all Australian states except Queensland. That it had quickly gained acceptance in nineteenth century Australia, although only in non-urban contexts, can be illustrated by the following quotes.

I was cool, and told Sweeney to bring out the revolvers; descended from the tree and got my gun and coo-eyed to Pierre and Kennedy, who came running. <3-305>

In vain I searched the bushes for the prospector's camp. I cooeed twice - got no response. <4-317>

The term *corroboree* for a traditional ceremonial dance of the Aborigines comes up twice as often as *cooee*. But its meaning is so specialized that it is used only once in a figurative sense denoting a social gathering by whites. It is in sources from all states except Queensland, just like *cooee*. An unkown Bulletin author comments on the celebrations of the hundreth anniversary of the landing of the First Fleet in 1888. It clearly shows his reservations about England as a mother country:

We have seen tumultuous revel, we have seen ecstatic display. We have heard wonderful speechmaking, and suffered exquisite headaches from innumerable banquets, feastings, and corroborees. And what does it amount to? Nothing.

The day just celebrated has been the anniversary - for the hundredth time repeated - of the settlement of these lands by England. What is there therein that we should rejoice? <4-156>

There are a few Aboriginal pidgin words like *jerran* (once, to be afraid from Dharuk *jiran*), *murry* (three times; an intensive from Dharuk *mari*) and *yabber-yabber* (three times; to talk, a talk). Only *jerran* and one instance of *yabber-yabber* are by a white person, the rest invariably applies to Aborigines.

As soon as things had got thus far, I began to feel rather "jerran", as the blacks say (i.e. timorous). <2-324>

I left Port Stephens, I intimated to them that I should soon return in a " corbon" (large) ship, with a " murry" (great) plenty of white people, and murry tousand things for them to eat. <2-049>

There was further a great waste of yabber-yabber about the diggers not being represented in the Legislative Council, and a deal of fustian was spun against the squatters. <3-106>

There are some examples in COOEE that show that an Aboriginal Pidgin English existed in different places. This was due to extended contact between Aborigines and Whites and as yet imperfect attempts to communicate with each other. An enlightening example of the beginning stages of such communication is related by George A. Robertson in Van Diemen's Land in his 1829 diary entries:

> 26 May
> Busily engaged in carrying on improvements, conversing with the natives, writing letters &c. Joe very ill. How ridiculous we make ourselves appear when we are only partially acquainted with a language. An instance of this occurred today. Joe called to me to come to his assistance. I found him very weak. He said, 'Mr Robinson TAGGERER WEE Joe TY'. I supposed him to mean that I was to collect WEE (wood) to consume his body when dead and I began to encourage him and said, 'Joe, TIME ME die' (no die). Part of his meaning was made apparent by his little girl coming with a stick to assist him to walk. In the course of the day one of the children evacuated. I immediately asked, 'What name?'; said name TY. The request of this poor man occurred to me, namely, that I was to get him a stick to assist him to walk out for the purpose of evacuation. I certainly was glad of the information as it enabled me to know in case of sickness the state of their body. On another occasion hearing the children repeatedly crying LY or E.LY, and seeing the parents give them water, I concluded that ERY or E.LY meant drink; and used the word in that sense. I soon found they knew not my meaning. <2-035>

A rather unfavourable comment comes from Peter Cunningham in NSW in 1827:

> All the natives round Sydney understand English well and speak it, too, so as to be understood by residents. The Billingsgate slang they certainly have acquired in perfection, and no white would think of competing with them in abuse or hard swearing, a constant torrent of which flows from their mouths as long as their antagonist remains before them. <2-019>

The word *croppy*, which originally denotes a political Irish convict from the 1798 uprising, became the general word for all prisoner for the Aborigines. Thus we can read in the Sydney Gazette in 1832.

> I have observed a remarkable shrewdness in the native tribes; […]; they display a remarkable cunning when they wish to accomplish any object; they make a distinction between free settlers and what they call "croppies" - that is, prisoners; <2-070>

J. Handt, a missionary in NSW, records in his diary in 1833 clear instances of an Aboriginal Pidgin:

> Friday, 25. […] Nerang Jacky, had found honey in the bush, and caught some oppussums besides, so that all of them had plenty to eat. I told them that God was very kind to them in giving them so much food. 'Yes", answered Nerang Jacky, "God that burhery (good) fellow, that gave it mine (my) plenty oppossum." […]
> Thursday, 31. Nerang Jacky went away this evening to fight with another black man by the name of Gentleman Jacky, for a woman, whom each of them called his own. The most singular circumstance is that the women does not live with either; for she has been with white men for some years, and is living with one. Last Monday, Nerang Jacky said he would go to him, and take her away, but when he returned, and I asked him whether he had brought her, he replied, "No, that white fellow mahne it altogedder" (would keep her altogether.) <2-082>

The whites, too, developed a special vocabulary when talking to or about the native population. The following letter by J.G. Paton tells us something about the recruitment of Polynesians for farming purposes in Queensland and the vocabulary used for this:

> A cargo of miscellaneous wares - called "trade" - representing in value £500 or £600, forms the most important part of the equipment of every recruiting vessel […] Arrived at the island, the anchor is generally lowered when two or three miles from the coast. When the natives desire to recruit, or rather the "trade" boat to come ashore, they invariably show themselves on the beach or light a fire. The boatswain on the voyage to the islands becomes the "recruiter" when there, and he is the direct medium on behalf of the planters in securing "recruits." Operations are commenced by lowering two boats - the recruiter's containing one white man and two Kanakas, the Government agent's manned by two white men and three Kanakas. Each man is armed with a Winchester or Snider, the agent and recruiter carrying a revolver, in the first - named boat is the "trade" chest, containing a variety of the articles mentioned. <4-286>

PEOPLE
The Aborigines of Australia did not have a term that included all the natives on the continent. Rather they identified themselves and each other as parts of particular tribes or as people of a certain status.

There is, for example, *kipper* (from Dharuk) for an initiated youth and *koradji* (Dharuk) or *boylya* (Nyungar) for a man experienced in healing. None of these, however, could be found in COOEE.

The term *myall* can mean stranger or an Aborigine living in a traditional way. In the first sense it is found twice in COOEE. Four more times it comes up in place names.

> I asked him why he had been killing cattle; he told me that the Myall fellows, meaning the "Mount Harris tribe," called the Bee Bee Jibbery tribe, had done it; he denied having killed any cattle; <2-225>

> I am assigned servant to Mr. Dangar; I was at his station at Myall Creek, as hutkeeper, for five months, in June, 1838. <2-181>

Warrigal is used in a similar way but favours us with only a single instance.

> On the sixth day after leaving Currawang the blackfellow who accompanied me became so frightened of the Warrigals, or wild blacks, that he tried to leave me, and refused to proceed any further towards the new country. <3-073>

Aboriginal women or wives are called *gins* (45 instances) in Dharuk and *lubras* (12 instances) in a Tasmanian language.

Lubra is used mainly in Victoria (eight times), and only later in South Australia (four times). This marks it as regionally restricted.

On the other hand, *gin* starts out in NSW and is first recorded in Victoria in 1853. It comes to Queensland in 1888, is written in Western Australia in 1895 and finally arrives in South Australia in 1896. This is a remarkable success story, but it can also be argued that it is a sad thing since the general use of *gin* prevented many other Aboriginal words from entering AusE. No figurative or extended use of the word can be found in COOEE. So one example from a letter of the Irish born Daniel Quigley can suffice here.

> There is nothing impossible for a colonial in the back blocks to be up to, but for fear you should think that I married a black princess or an old jin for to colonize myself, I will get Mrs. D. Quigley to send you her photo and that of my gairl, <4-294>

Aboriginal children are called *piccaninny* (five times). This is a truly globalized pidgin word, which has its origins, it seems, in Portuguese *pequeno*, i.e. 'small', and is part of many pidgins world-wide. It was imported by the white settlers and used for the children of the native population. The Aborigines readily accepted the word and used it figuratively as can be seen in the following quote:

> The blacks still say as they said the last flood, 'This is only piccaninny - big one coming.' <2-286>

Now the words used by the white population to designate the Aboriginal population will be discussed. There are various words used, but the frequencies are very telling.

Table 8: Total Frequencies of words denoting Aborigines/Black People

word	#
kafir	2
Jim Crow	4
negro	11
nigger	13
Aborigine/Aboriginal	193
black (various collocates)	650
native	1507

Close inspection reveals that *kafirs*, as expected, are the natives of the African Cape Colony. This word is not used for Australians.

More interesting is the term *Jim Crow*. It originates in the refrain of a plantation song mainly sung by African Americans in the US, where it was later used in a very derogatory way by the whites. The earliest reference in the Oxford English Dictionary (OED) is from 1832. It is not in the AND, but it can be found in COOEE as early as 1853. In the first example *Jim Crow* is used for Aborigines, the second example seems to be a mountain, the third is certainly a toponym and the last, astonishingly, is obviously a particular kind of hat, presumably one whose shape reminds one somehow of Aborigines.

> The aborigines in our neighbourhood (afterwards known as the "Jim Crow" tribe) were from the first peaceable. <3-077>

[…] witness the bunyip and (Mr. Powlett's?) great serpent of the interior, both of which have been accurately described in fifty different shapes; also the volcanic eruptions of Jim Crow, &c., &o.; in short, if leading questions are put to them, as is usually done by enthusiastic inquirers, who are following up their own ideas, they (the natives) may, as I think, be made to say or to describe anything. <3-077>

From the time of my arrival on the Loddon the aboriginal natives were concentrated under the charge of Mr. Parker at Jim Crow Hill (Mount Franklin), and with the exception of murdering a Mr. Allan, who had a small cattle station (which I afterwards purchased) between Mr. Catto and me, committed no depredations of any consequence, […] <3-080>

Were there always so many dead trees, and did Australians always wear such a bewildering variety of hats? There are hard felt and soft felt, broad-brimmed and narrow-brimmed, sailor, Panama, Buffalo Bill, Jim Crow, cowboy, and cavalier; hats puggareed, hats bare, and even the white "Derby" chimney-pots. It is a nightmare of hats. <4-398>

Nigger and *negro* were still relatively neutral terms in nineteenth century Australia. Although they are applied to Aborigines, they are not at all frequent. Some instances also refer to African Americans, some of whom had found a way to come to Australia.

Mr. Hodgson, who had perhaps suffered most by additional fatigues; so that he and Caleb, the American negro, prepared for their return to Moreton Bay. <2-283>
For nearly a week I was camped with some niggers and they having taken a fancy to me gave me a grand corroberie a thing few of the colonials have seen. <4-276>

Three words clearly dominate as designations for the Aboriginal peoples of Australia, namely *Aborigines/Aboriginal* (193 times), *blacks* (650 times) and *natives* (1507 times).
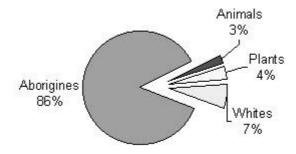*Aborigine*, like *blacks*, is only applied to Australian Aborigines. The first collocates solely with *black*, whereas the latter shows a wide range of collocates.

Table 9: Frequencies of the word black collocating with other words

word	#
black (without collocate)	433
black aboriginal	2
black(-)boy	31
black(-)fellow(s)	168
black gin	8
black girl	4
blackman, -men	35
black native(s)	15
black(-)tracker	4
black(-)woman	14

Native is the most frequent of all the words designating Aborigines. It can be applied to four different *designata*, namely Aborigines (1507 times), native plants (61 times), native animals (44 times) and native-born white Australians (121 times). This is illustrated in Figure 17.
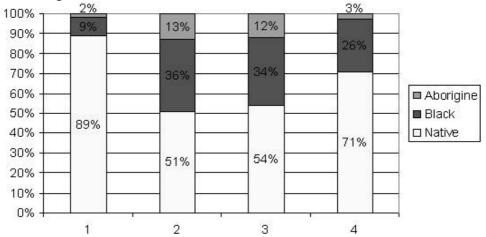
Figure 17: Usage of *native* for Aborigines, animals, plants and whites

The following figure shows the usage of Aborigine/Black/Native over time. Clearly, *native* is the

favoured name altogether, followed by *black* and considerably trailed by *Aborigine*.
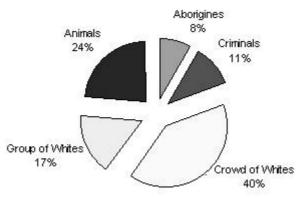
Figure 18: Usage of Aborigine/Black/Native over Time

But the numbers change over the four historical periods 1-4 (= 1788-1825, 1826-1850, 1851-1875 and 1876-1900). The term *native* is predominant in the first and the fourth period. But in the second and third period, there is a sharp increase in the use of *black* and of *Aborigine*. It comes as a surprise that *Aborigine* is relegated to only 3% in the fourth period again, and that *black* also declines by one fourth Obviously, *native* was still unambiguous enough to be used in such high frequencies. The time of the term *Aborigine* had not yet come. The word *black* may have been restricted in usage by its contemporary application to black people in other parts of the world, especially in the US.
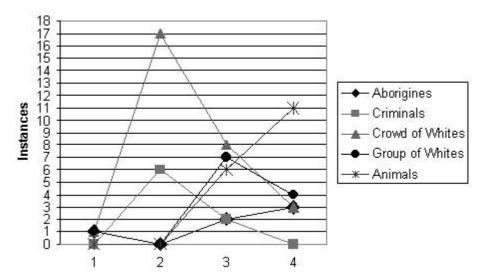
There is one other word that is used to describe a number of Aborigines, *mob*. Of course, *mob* can also have other meanings as Figure 19 shows. Altogether *mob* appears 72 times.

Figure 19: Meanings of the word *mob*

As can be seen, *mob* mostly means a crowd of more or less disorderly whites (a), followed by a mob of cattle (b), a designation for a group of whites without negative overtones (c), a mob of criminals (d) and finally a mob of Aborigines (e). Figure 20 shows the frequencies of meanings a-e in the different periods. Meaning (a) is the original one, (b)-(e) are all unique to AusE. This testifies to the innovative power of the English language in Australia from early on.

Figure 20: Meanings of *mob* over Time

Use of *mob* for Aborigines is low throughout, only six instances are recorded in COOEE. It is not used in a negative sense, but rather is a descriptive element, sometimes condescending, but certainly not demeaning. It is there earliest in 1816 by an unkown writer in Van Dieman's Land.

> Tolo asked Briggs not to go away until they had a dance, The mob of them - about three hundred in number - formed […] three divisions, the men and women forming two of them, and the children another. <1-156>

> The emigrants of 1822 remember a number of natives, who roamed about the district, and were known as the "tame mob." They were absconders from different tribes, and separated from their chiefs. <3-050>

For criminals it can only be found in periods 2 and 3, obviously because this usage took some time to develop and when there were no more groups of convicts it quickly dropped out of use.

> Brady's mob became a terror to all the Colonists - no settler could be secure, for one moment, from their depredations. <2-013>

The original sense is that of a disorderly or even riotous crowd of people. This meaning was prominent in Australia in the first three periods.

> Later in the evening a mob amounting to a thousand, proceeded to the house of Mr Green, the auctioneer in Elizabeth-street, […], and announced their intention of burning or pulling down his house, […] <2-276>

After 1875 meaning (a) as the most frequent meaning is replaced by (b) and (c), Australianisms in the true sense of the word. A *mob of animals* is not found before 1850 in COOEE, but then increases considerably and is applied not only to cattle, but also to kangaroos, mice, etc. This meaning is the most frequent one at the turn of the century.

> There were said to be 10,000 head of cattle on its banks, in various "mobs." <3-079>

> though there are plenty of small fish in the Finke, yet I do not think they are large enough for mobs of pelicans to exist upon; <3-284>

While in Australia, *mob* thus lost more and more of its disparaging overtones and became a very neutral expression describing a number of people or things. This meaning became more frequent than meaning (a) after 1875. An early and very clear example for this is the following, taken from a letter Penelope Selby wrote home in 1851.

> You will have a mob of grandchildren soon. If my Prid had been a girl you would likely have had a great grandchild shortly. <3-018>

Summing up the above, it can said that the Aborigines and their respective languages contributed to the formation of AusE. However, these contributions are mostly very infrequent and apply mostly to very specialized items/situations. Only few words have succeeded in gaining general acceptance, most are lost. In this linguistic history is a true reflection of the history of Aboriginal-white interactions.

7. Conclusion

As could be shown, empirical investigation of early English in Australia helps to answer a whole new variety of questions.

Frequencies of words and their different meanings yield insights into mainstream use. Social data about the authors and factual data about the texts reveal social, temporal and spatial distributions.

This does not invalidate the usefulness of studies dominated by 'reasonings'. What we need is a division of labour and mutual recognition.

Plato and Aristotle must be reconciled.

8. List of Works Cited

Corpus Sources Quoted

Bigge, John Thomas, 1822, *Report of the Commissioner of Inquiry into the State of the Colony of New South Wales*, Adelaide: Libraries Board of Australia, 1966.

Bride, Thomas F., 1898, *Letters from Victorian Pioneers*, South Yarra: Heinemann, 1969.

Burton, E.S., 1833-38, *Letters of Chief Justice Burton*, http://www.law.mq.edu.au/scnsw.

Calvert Expedition, 1902, *Journal of the Calvert Scientific Exploring Expedition, 1886-7*, Victoria Park: Hesperian Press, 1993.

Carboni, Raffaelo, 1855, *The Eureka Stockade*, http://setis.library.usyd.edu.au/ozlit.

Clark, Manning, 1975, *Select Documents in Australian History 1851-1900*, Sydney: Angus & Robertson.

Clark, Manning, 1977, *Select Documents in Australian History 1788-1850*, Sydney: Angus & Robertson.

Clarke, Patricia and Dale Spender, 1992, Life Lines: Australian Women's Letters and Diaries 1788 to 1840, St Leonards: Allen & Unwin.

Compton, George & Ron Manners, 1993, *Kanowna's Barrowman - James Balzano*, Victoria Park: Hesperian Press.

Connell, R.W. and T.H. Irving, 1980, *Class Structure in Australian History: Documents, Narrative and Argument*, Melbourne: Longman.

Dawson, Robert, 1830, *The Present State of Australia, facsimile edition*, Alburgh: Archival Facsimiles Limited, 1987.

Decisions of the Supreme Court of NSW, 1824-41 Decisions of the Superior Courts of New South Wales, 1788-1899, Published by the Division of Law, Macquarie University, http://www.law.mq.edu.au/scnsw.

Felton, Mathew and Sarah, 1829-32, *Journals and Memoirs of Sarah and Mathew Felton*, http://www.users.bigpond.com/narrabeen/feltonmathew/index.htm.

Fitzpatrick, David, 1994, *Oceans of Consolation*, Cork: Cork UP.

Fitzpatrick, Kathleen, 1958, *Australian Explorers*, London: Oxford UP.

Frost, Lucy, 1984, *No Place for a Nervous Lady*, Fitzroy, Vic: McPhee Gribble Publishers Ltd.

Giles, Ernest, 1872, *Diary of Explorations of Mr Ernest Giles in Central Australia*, Adelaide: Sullivan's Cove, 1986.

Hale, John, 1950, *Settlers: Being Extracts from the Journals and letters of early colonists in Canada, Australia, South Africa and New Zealand*, London: Faber & Faber.

Handt, J.C.S., 1831-43, *Wellington Valley, NSW: Letters and Diary of J.C.S. Handt*, http://www.newcastle.edu.au/department/hi/carey/index.html.

Harpur, Charles, 1845, *Poems*, http://setis.library.usyd.edu.au/ozlit.

Harris, Alexander, 1847, *Settlers and Convicts*, Melbourne: Melbourne UP, 1995.

Ingleton, Geoffrey, 1988, *True Patriots All, or News from Early Australia*, Tokyo: Charles E. Tuttle Company, Inc.

Irvine, Nance, 1988, *The Sirius Letters, The Complete Letters of Newton Fowell*, Sydney: The Fairfax Library.

Kramer, Leonie and Adrian Mitchell, 1985, *The Oxford Anthology of Australian Literature*, Melbourne: Oxford UP.

Langloh Parker, K., 1897, *Australian Legendary Tales*, http://digital.library.upenn.edu/books/.

Macmillan, David S., 1966, *Two Years in New South Wales, Peter Cunningham, Surgeon*, R.N. Sydney: Angus & Robertson.

Mann, David Dickenson, 1811, *The Present Picture of New South Wales - 1811*, Sydney: John Ferguson, 1979.

McCrae, George Gordon, 1867, *Mamba The Bright Eyed: An Aboriginal Reminiscence*, http://setis.library.usyd.edu.au/ozlit.

Muskett, Philip E., 1893, *The art of living in Australia*, http://setis.library.usyd.edu.au/ozlit.

O'Farrell, Patrick, 1984, *Letters from Irish Australia*, Sydney: New South Wales UP.

Plomley, N.J.B., 1966, *Friendly Mission. The Tasmanian Journals and Papers of George Augustus Robinson 1829-1834*, Kingsgrove, NSW: Tasmanian Historical Research Association.

Tench, Watkin, 1789, *A Narrative of the Expedition to Botany Bay*, http://setis.library.usyd.edu.au/ozlit.

Tench, Watkin, 1793, *A Complete Account of the Settlement at Port Jackson*, http://setis.library.usyd.edu.au/ozlit.

Ward, Russel and John Robertson, 1969, *'Such Was Life': Select Documents in Australian Social History*, vol. 1: 1788-1850, Sydney: Jacaranda Press.

Webby, Elizabeth, 1989, *Colonial Voices: Letters, Diaries, Journalism and other Accounts of Nineteenth-Century Australia*, St. Lucia: University of Queensland Press.

West, John, 1852, *The History of Tasmania*, Sydney: Angus & Robertson, 1971.

Secondary Writings

Arnaud, René, 1998, "The development of the progressive in 19th century English: a quantitative survey", *Language Variation and Change* 10, 123-152.

Baker, Anthony W., 1984, *Death is a Good Solution,* St. Lucia: University of Queensland Press.

Bernard, John R.L., 1969, "On the uniformity of Australian English", *Orbis* 18, 62-73.

-----, 1981, "Australian pronunciation", In: Delbridge, Arthur, ed, *The Macquarie Dictionary*, Sydeny: Griffin Press, 18-27.

Biber, Douglas, 1988, *Variation across Speech and Writing,* Cambridge: Cambridge UP.

-----, Susan Conrad and Randi Reppen, 1998, *Corpus Linguistics - Investigating Language Structure and Use,* Cambridge: Cambridge UP.

----- and Edward Finegan, 1992, "The linguistic evolution of five written and speech-based English genres from the 17th to the 20th centuries", In: Rissanen, et al., eds, *History of Englishes. New Methods and Interpretations in Historical Linguistics,* Berlin: Mouton, 688-704.

----- and -----, 1997, "Diachronic relations among speech-based and written registers in English", In: Nevalainen, T., L. Kahlas-Tarkka, *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, Helsinki: Societé Neophilologique, 253-75.

-----, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan, 1999, *Longman Grammar of Spoken and Written English,* London: Longman.

Bliss, Allen J., 1979, *Spoken English in Ireland 1600-1740,* Dublin: Cadenus Press.

Borrie, Wilfred D., 1994, *The European Peopling of Australia. A Demographic History 1788-1988,* Canberra: Research School of Social Sciences, The Australian National University.

Clark, Manning, 1957, *Sources of Australian History,* London: Oxford UP.

-----, 1975, *Select Documents in Australian History 1850-1900*, Sydney: Angus & Robertson.

-----, 1977, *Select Documents in Australian History 1788-1850*, Sydney: Angus & Robertson.

Cleverly, John, 1971, *The First Generation: School and Society in Early Australia*, Adelaide: Griffin Press.

Cochrane, George R., 1989, "Origins and development of the Australian accent", In: Collins, Peter and David Blair, eds, *Australian English: The Language of a new Society*, St Lucia: University of Queensland Press, 171-186.

Collins, H.E., 1975, *"The sources of Australian pronunciation", Working Papers of the Speech and Language Research Centre, Macquarie University* 1, 115-128.

Fitzpatrick, David, 1994, *Oceans of Consolation,* Cork: Cork UP.

Fritz, Clemens, 1996, *Early Australian Letters: A Linguistic Analysis*, unpublished MA thesis.

-----, 1998, "Letters from early Australia – linguistic variation and change", *Dialectologia et Geolinguistica* 6, 25-42.

-----, 2000a, "Language, change and identity: The Irish in 19th century Australia", In: Foley, Tadgh and Fiona Bateman, eds, *Irish-Australian Studies. Papers Delivered at the Ninth Irish-Australian Conference* Galway, April 1997, Sydney: Crossing Press, 57-66.

-----, 2000b, "The Irish in Australia: Aspects of linguistic accommodation", In: Hildegard L.C. Tristram, ed, *The Celtic Englishes II*, Heidelberg: Carl Winter, 57-74.

-----, no date, "Inertia and change: Irish English in early Australia", unpublished paper.

Goodwin, Ken, 1986, *A History of Australian Literature,* London: Macmillan.

Görlach, Manfred, 1991, "Australian English: Standards, stigmata, stereotypes and statistics", In: Görlach, Manfred, ed, *Englishes: Studies in varieties of English 1984-1988,* Amsterdam: Benjamins, 144-173.

-----, 1999, *English in Nineteenth-Century England. An Introduction*, Cambridge: Cambridge UP.

Greenwood, Gordon, 1955, *Australia: A Social and Political History,* Sydney: Angus & Robertson.

Gunn, John S., 1972, "Early Australian pronunciation", *Linguistic Communications* 6, 36-40.

-----., 1992, "Social contexts in the history of Australian English", In: Tim W. Machan and Charles T. Scott, eds, *English in its Social Contexts: Essays in Historical Sociolinguistics*, Oxford: Oxford UP, 204-229.

Haan, Pieter de, 1996, "More on the language of dialogue in fiction", *ICAME* 20, 23-40.

Halliday, Michael A.K., 1985, *Spoken and Written Language,* Burwood (Victoria): Deakin UP.

----- and Ruqaiya Hasan, 1985, Language, Context, and Text: Aspects of Language in a Social-semiotic Perspective, Burwood (Victoria): Deakin UP.

Hammarström, Göran, 1980, *Australian English: Its Origin and Status*, Hamburg: Buske.

Hergenhan, L. T., 1976 "English publication of Australian novels in the 19th century: The case of 'His Natural Life'", In: Cantrell, Leon, ed, *Bards, Bohemians and Bookmen: Essays in Australian Literature,* St. Lucia: University of Queensland Press, 56-71.

Hiltunen, Risto, 1990, *Chapters on Legal English,* Helsinki: Suomalainen Tiedeakatemia.

Horvath, Barbara, 1985, *Variation in Australian English: The Sociolects of Sydney*, Cambridge: Cambridge UP.

Ingleton, Geoffrey C., ed, 1952, *True Patriots All, or News from Early Australia as Told in a Collection of Broadsides,* Sydney: Angus and Robertson.

Irvine, Nance, 1992, *Dear Cousin: The Reibey Letters,* Sydney: Hale & Ironmonger.

Jupp, James, 1989, *The Australian People,* Sydney: Angus & Robertson.

Kociumbas, Jan, 1992, *The Oxford History of Australia, vol 2, 1770-1860, 'Possessions',* Melbourne: Oxford UP.

Langker, Robert K., 1980, *Flash in New South Wales 1788-1850,* University of Sydney, Australian Language Research Centre = Occasional Paper 18.

-----, 1981, *The Vocabulary of Convictism in New South Wales 1788-1850,* University of Sydney, Australian Language Research Centre = Occasional Paper 19.

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.

Lonergan, Dymphna, fc., *Irish Influence on Australian English*, PhD thesis.

Mellinkoff, David, 1963, *The Language of the Law,* Boston: Little, Brown and Company.

Mitchell, Alexander G., 1995, *The Story of Australian English*, Sydney: Dictionary Research Centre.

----- and Arthur Delbridge, 1965, *The Pronunciation of English in Australia*, Sydney: Angus and Robertson.

Morris, Edward E., ed, 1896, *Austral English: A Dictionary of Australasian Words, Phrases and Usages, with Those Aboriginal and Maori Words Which Have Become Incorporated in the Language and the Commoner Scientific Words That Have Had Their Origin in Australia*, London: Macmillan.

Newbrook, Mark, 1992, "Unrecognized grammatical and semantic features typical of Australian English", *English World Wide* 13, 1-32.

Newbrook, Mark, 2001, "Syntactic features and norms in Australian English", In: Blair, David and Peter Collins, eds, *English in Australia*, Amsterdam: Benjamins, 113-132.

Nicholas, Stephen and Peter R. Shergold, 1988, "Unshackling the past", In: Nicholas, Stephen, ed., *Convict Workers: Reinterpreting Australia's Past*, Cambridge: Cambridge UP, 3-13.

O'Farrell, Patrick, 1984, *Letters from Irish Australia,* Sydney: New South Wales UP.

-----, 1989, "The Irish in Australia and New Zealand 1791-1870", In: Vaughan, William, ed, *Ireland under the Union 1801-1870, I, A New History of Ireland,* vol. V., 661-681.

-----, 1996, "The Irish in Australia and New Zealand 1870-1990", In: Vaughan, William, ed, *Ireland under the Union, II, 1870-1921, A New History of Ireland*, vol. VI, 703-724.

Palander-Collin, Minna, 1999, "Male and female styles in 17th century correspondence: I think", *Language Variation and Change* 11, 123-141.

Price, Charles, 1987, "Immigration and ethnic origin", In: Vamplew, William, ed, *Australians: Historical Statistics*, Fairfax: Syme & Weldon Associates, 2-22.

Ramson, William S., 1966, *Australian English: An historical Study of the Vocabulary, 1788-1898*, Canberra: Australian National UP.

-----, 1970, "Nineteenth-Century Australian English", In: Ramson, William S., ed, 32-48.

-----, 1972, "Distinctive features of Australian English", In: Turner, George W., ed, 33-45.

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

Robson, Leslie L., 1955, *The Convict Settlers of Australia,* Carlton: Melbourne University Press.

Shaw, Alan G.L., 1966, *Convicts and the colonies. A Study of Penal Transportation from Great Britain and Ireland to Australia and other parts of the British Empire,* Melbourne: Melbourne UP.

Taylor, Brian, 1992, "Otto 988 to Ocker 1988: The morphological treatment of personal names in Old High German and colloquial Australian English", In: Blank, Claudia ed, Language and Civilization: *A Concerted Profusion of Essays and Studies in Honor of Otto Hietsch*, vol II, Frankfurt: Peter Lang, 505-536.

-----, 1997, "Sydney English around 1900: A linguistic critique of Louis Stone's 'Jonah'", In: Heinz Ramisch and K. Wynne, eds, *Language in Time and Space,* Stuttgart: Franz Stiner Verlag, 258-270.

-----, 1998, "Australian English and its interaction with other Englishes", In: Lindquist, Hans, Staffan Klintborg, Magnus Levon and Marai Estlind, eds, *The major varieties of English, Papers from MAVEN 97*, Växjö: Vaxjö University.

-----, 2001, "Australian English in interaction with other Englishes", In: Blair, David and Peter Collins, eds, *English in Australia*, Amsterdam: Benjamins, 317-340.

Troy, Jakelin, 1992, "'Der mary this is fine cuntry is there is in the wourld': Irish-English and Irish in late eighteenth and nineteenth century Australia", In: Dutton, Tom, et al., eds, *The Language Game: Papers in Memory of Donald C. Laycock,* Canberra: ANU, 459-477.

Trudgill, Peter, 1986, *Dialects in Contact*, Oxford: Oxford UP:

Turner, George W., 1960, "On the origin of Australian vowel sounds", *AUMLA* 13, 33-45.

-----, 1994, "English in Australia", In: Burchfield, Robert, ed, *English in Britain and Overseas: Origins and Development, The Cambridge History of the English Language* V, Cambridge: Cambridge UP, 277-327.

Ward, Russel, 1958, *The Australian Legend*, Melbourne: Oxford UP.

----- and John Robertson, 1969, *'Such Was Life': Select Documents in Australian Social History 1788-1850*, Sydney: Jacaranda Press.