Perceptual Vowel Space for Australian English Lax Vowels: 1988 and 2004

Robert Mannell

Speech Hearing and Language Research Centre,
Department of Linguistics,
Macquarie University,
Australia
robert.mannell@mq.edu.au

Abstract

This paper examines changes in the perception of Australian English lax vowels between 1988 and 2004. In each of the years 1988 and 2004, 20 phonetically-trained native speakers of Australian English were presented with 224 different randomised, synthesised hVd tokens. The vowels differed according to their F1 and F2 values (all possible multiples of 100 Hz within a designated "male" vowel space). Half the vowels were "long" vowels (300 ms) and the other half were "short" vowels (150 ms). Subjects were given a forced-choice closed set task. They were told to concentrate on the vowel in each token and to label it with the symbol for the closest Australian English vowel phoneme. The responses for the "short" vowel plane were then plotted on a vowel perceptual contour map in F1/F2 formant space. Visual examination of the two resulting maps revealed apparent changes in the perceptual boundaries between all adjacent pairs of lax vowels. χ^2 analysis of 1988 versus 2004 responses for each vowel at each data point on the vowel space revealed significant changes in perceptual patterns at the perceptual boundaries between four pairs of adjacent vowels (/ɪ/ and /e/, /e/ and /æ/, /æ/ and /ʌ/, /ɒ/ and /ʊ/). Both an independent samples t-test and a MANOVA test examined the pattern of subject response means for each vowel and showed a significant change in the mean response between 1988 and 2004 for /æ/, /ʌ/ and /ɒ/.

1. Introduction

Early impressionistic studies of Australian English (Mitchell, 1946, Mitchell and Delbridge, 1965a, 1965b) described Australian English pronunciation as being characterised by a single dialect continuum which could be divided for descriptive convenience into Broad, General and Cultivated categories. The most important segmental production correlate of these categories was the pronunciation of the vowels /i, u, eɪ, oʊ, aɪ, aʊ/. These categories roughly correlated with socioeconomic status, country versus city and male versus female. They asserted that there were no regional distinctions in pronunciation and that the differences between speakers of Australian English from different regions depended upon the proportion of the speakers found at various points on the continuum. For example, country speakers were more likely to use broad forms than were city speakers and male speakers were more likely to use broad forms than were females. According to this view, regional differences were mainly based upon lexical choice rather than differences in pronunciation. Bernard's (1967) pioneering acoustic study of Australian English mostly confirmed the results of the earlier impressionistic studies although his description of /u/ suggested that a change in the pronunciation of this vowel (loss of a previously distinctive on-glide) may have occurred.

1.1. Change in Australian English Vowel Production

Using Bernard (1967) as a baseline, recent acoustic studies have indicated that the production of certain Australian English vowels have undergone changes since the 1960's. Cox's (1998) re-examination of the Bernard data, which separated the subjects into three age groups, showed significant age differences for the F1 and/or F2 offsets of 7 monophthongs (/e, æ, a, ʌ, ɒ, ʊ, ɜ/), one significant age difference for a monophthong target (/u/ F1), and significant age differences for the

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after full paper review

first target of two diphthongs (/oʊ/, /ɪə/). The target differences for these different age groups were interpreted as possible evidence of change in progress for the production of these three vowels and the offset differences were interpreted as possible evidence of age-related changes in speaking styles that were, perhaps, occurring at the time of the Bernard study. Cox (1996) describes some evidence for change in the vowel /oʊ/ and this process is also suggested in the Harrington, Cox and Evans (1997) study of vowels in the ANDOSL corpus (Millar, Vonwiller, Harrington and Dermody, 1994). Cox (1996) suggested that /æ/ may have become a marker vowel for stylistic variation (cf. the six marker vowels of the Broad - Cultivated dialect continuum).

Cox (1999) compared a matched set of 27 15-16 year old male Sydney-based speakers of General Australian English from Bernard's (1967) data and 60 15 year old Sydney-based speakers of General Australian English from Cox's (1996) data (collected in about 1990). She found significant differences in onset, offset and/or target values for 12 vowels (/i, ɪ, æ, ɒ, ɔ, u, ɪə, eə, aɪ, aʊ, oʊ/). Monophthong targets were found to have undergone the following significant articulatory changes between the 1960's and 1990. The targets of /ɪ, ɒ, ɔ, u/ were raised (lower F1), the targets of /u, ɜ/ were fronted (higher F2) and the target of /æ/ was both lowered (higher F1) and retracted (lower F2).

Cox and Palethorpe (2001) analysed the vowels in the ANDOSL corpus (Millar *et al*, 1994) for synchronic subject age-related evidence of vowel change. The results showed significant lowering of /æ/ (as well as changes to the production of /oʊ, eɪ, aʊ/) for younger male subjects compared to older male subjects.

Cox and Palethorpe (2004) questioned the continuing usefulness of the broadness continuum for the description of Australian English and demonstrated that a new set of vowels are now being used as markers of socio-phonetic variation. They particularly found significant differences in /æ/ (F1), /a/ (F1 and F2) and /ʌ/ (F2) and /aɪ, aʊ/ (target 1, F1 and F2). They described lower open vowels for /æ, a, ʌ/ for a grouping of subjects that they described as "hyper-general" and suggested that this new, much lower /æ/ may have become a new social marker. They also described the way in which this has provided space for the lowering of /e/ which they observed for some speakers in their database and which they hypothesise has "progressed rather rapidly" since data collection in 1998. Cox and Palethorpe (2001) also examined the productions of age-related differences in vowel production for female subjects in the ANDOSL database. For female subjects /æ, a, ʊ/ are lowered and /i, ʊ/ are fronted (as well as changes to the production of /eɪ, aɪ, ɔɪ/) for younger compared to older subjects.

To summarise, it now seems well established that Australian English has undergone some significant changes in vowel production between the 1960s and the present. These changes have affected both monophthongs and diphthongs but since this paper is only concerned with lax monophthongs only changes in those vowels will be summarised here. There is evidence that a significant lowering of /æ/ was emerging (for younger subjects) by the time of the collection of the ANDOSL corpus in the early 1990's and that this lowering has continued, particularly for adolescents. A new very low version of /æ/ appeared amongst adolescent female speakers by about 1998 and this vowel appears to have become a social marker. Retraction of /æ/ has also been noted. Significant evidence has also been found from the early 1990's for the raising of the targets of /ɪ/ and /ɒ/ amongst adolescent speakers. More recently, for data collected in 1998, some differences between broader and general /ʌ/ appear to have developed with the broader form being more raised and retracted. Also, a lowering of /ʊ/ has been recently observed for some female adolescents. Finally, some evidence for a lowering of /e/ seems to be emerging but this has not yet been confirmed.

The question asked in this study is, to what extent have patterns of vowel perception by speakers of Australian English changed over the last 16 years and to what extent have these perceptual changes matched the patterns of lax vowel production changes that have been observed in these production studies.

1.2. Male-female differences in AE vowel production

As well as the expected physiologically-based formant differences between males and females, Cox (1996) found some significant differences in the vowel formant patterns for adolescent male and female speakers of Australian English. These differences, however, were considered to be consistent with the expectation that male speakers of Australian English have a greater tendency to use broader vowels than do female speakers.

Cox and Palethorpe (2001) examined the productions of age-related differences in vowel production for female subjects and found evidence that there were more age differences for females than for males. This suggests that adolescent female speakers of Australian English may be more active initiators of Australian English vowel change than adolescent male speakers.

Accepted after full paper review

In the present study, a model of male vowel production has been used to produce the perceptual tokens, but the majority of perceiving subjects used in this study are female. This study does not attempt to address the question as to whether male and female subjects differ in their perceptual patterns. Whilst there are obvious physiologically determined production differences as well as some differences in the relative pattern of vowel formants, especially for adolescents, it seems likely that adult male and female subjects will not greatly differ in their perceptual behaviour if they are reasonably matched for their pronunciation (eg. if they are all speakers of General Australian English). This is because males and females both have considerable experience at accurately identifying the utterances of speakers of both sexes particularly amongst their own speech communities.

1.3. Regional Differences in AE vowel production

Some evidence is beginning to emerge for regional variation in vowel production that is not readily explained in terms of choice along the broad-cultivated continuum. For example, Cox and Palethorpe (1998) describe evidence for "area specific vowel realisations" of adolescent speech within Sydney. In this study, significant area-specific production differences have been found for F1 for /e, æ, ɔ, ʊ/ and for F2 for /ɔ, ʊ/ as well as target 1 differences for /aɪ, ɔɪ, oʊ, aʊ/. This effect may, however, be specific to adolescents because "... the pressure on adolescents to conform to peer norms may be so strong that it eclipses parental influence. Area effects therefore take precedence because they represent peer effects" (ibid.). Cox and Palethorpe (2003) have also found evidence for regional variation in vowel production between country areas of NSW and Victoria, particularly for /u/ and /æ/ (in /hVd/ context) and /e/ and /æ/ (before /l/).

In the present study, the use of adolescent subjects is avoided. This should hopefully avoid the peer norm behaviour that appears to be behind the suburban patterns of vowel production within the Sydney metropolitan area. Further, most subjects are long-term residents of Sydney or surrounding coastal areas and so the effects of emerging regional patterns of vowel production will be minimised.

1.4. Production and perception of vowels

Lindblom (1986) suggests that changes in the acoustic realisation of phonological systems are the result of adjustments that are motivated by the need to uniformly distribute phonemes in perceptual space and to thus maximise the perceptual distinctiveness or contrast of each phoneme in that space. This explains why, once a disruption occurs in a class of speech sounds, such as the lax vowels, the remainder of the sounds in that class then also shift to re-achieve maximal contrastiveness. This chain reaction is sometimes called chain-shifting. The reason for the initial disruption is not so clear. Possibly, the motivation for such a disruption is sociolinguistic and is related to the maximisation or minimisation of in-group versus out-group vocal differences. Labov (1994) suggests that tense vowels tend to raise and/or move forward, lax vowels tend to lower and/or retract whilst back vowels tend to move forward. Furthermore, since phonologically tense vowels are often peripheral vowels (towards the outer edge of the physiologically possible articulatory vowel space) movements of tense vowels tend to be along the periphery of the vowel space. On the other hand, phonologically lax vowels are often non-peripheral (realised away from the articulatory vowel space periphery) and so their movements tend to maintain this characteristic by shifting along a non-peripheral path. These universal tendencies may, however, be opposed by other (eg. sociolinguistic) forces.

This study examines patterns of lax vowel perception for the years 1988 (Mannell, 1988) and 2004. It attempts to address two hypotheses. The first hypothesis is that the perceptual response of phonetically trained Australian English subjects to synthetic tokens modelling a male production of lax vowels has changed significantly between 1988 and 2004. The second hypothesis is that the changes in the perception of lax vowels by phonetically trained Australian English subjects are related to changes in the production of Australian English lax vowels over the same period.

2. Methodology

2.1. Perceptual Tokens

A vowel formant space was modelled with the following characteristics. Firstly, two planes were defined: a short vowel plane and a long vowel plane with vowel lengths set at 150 ms and 300 ms respectively (approximately following the average short and long vowel lengths for /hVd/ citation tokens in Bernard (1967)). The extreme F1 and F2 values for these planes were derived from the vowel frequency distribution for citation /hVd/ productions by all speakers from Bernard (ibid.). These vowel planes have been shown (Mannell 1988, 1995) to result in stable vowel perceptual patterns that mimic the perception of male vowel data by phonetically trained and phonetically naive male and female perceiving subjects. The tokens were synthesised using a purpose-built software parallel formant speech synthesiser based on the synthesiser described by Clark, Summerfield and Mannell (1986). 112 vowels were defined for each of the long and short vowel planes. These vowels had F1 and F2 values that were all possible multiples of

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after full paper review

100 Hz that fit within the defined male vowel model. In figures 1 to 4 the intersections of the dotted grid lines within the male vowel space boundary (thick unbroken line approximating the vowel triangle), and also the points on the boundary, represent the F1 and F2 values for each of the 112 tokens for that plane. The same vowel tokens, but with different duration, are defined for the other vowel plane, giving a total of 224 tokens. F3 was specified as a simple function of F2 (appropriate for all vowels except, perhaps, /ʊ/) using the average F3 values for citation /hVd/ tokens from Bernard (1967). Appropriate coarticulation for the preceding /h/ and the following /d/ was provided using a locus space ("locus matrix") model, after Clark (1981). The tokens were randomised and presented to subjects at 5 second intervals. A short training sequence preceded the test materials to accustom the subjects to the voice quality. The same tokens and randomisation were used for both the 1988 and the 2004 experiments.

Only lax vowel responses to short vowel plane tokens are analysed in this paper.

2.2. Subjects

In both experiments 20 phonetically trained subjects participated.

The 1988 group consisted of 16 females and 4 males and the mean age was 27.3 (std.dev.=7.38). These are not exactly the same 20 subjects reported in Mannell (1988) as the original response sheets for two subjects have been lost. They were replaced with two phonetically trained subjects who performed a very similar experiment which differed from this experiment only in that it was an unforced choice experiment and this experiment is a forced choice experiment. The two subjects with the fewest blank responses (~5% of their responses) were chosen and added to this group.

The 2004 group consisted of 16 females and 4 males and the mean age was 25.7 (std.dev.=9.39).

The age distribution of the two groups are not significantly different.

Place of origin was not explicitly controlled for this experiment, but in both experiments at least 80% of the subjects were long term residents of Sydney or the surrounding coastal region and only one subject (in the 2004 group) came from outside NSW (he came from Perth).

All subjects were native speakers of Australian English (born here or arrived before 5 years old, and are monolingual English speakers). They were all judged to be speakers of General Australian English.

Subjects all passed a simple hearing screening test and no subjects reported hearing problems.

2.3. Experimental Design and conditions

Subjects were presented the randomised materials in a sound treated room over high quality headphones at 80 dB SPL.

Both experiments were forced choice, closed set experiments. Subjects were instructed to attend particularly to the vowel in each token and to write down the phonemic symbol for the closest Australian English vowel phoneme. Blank responses were strongly discouraged and in both conditions most subjects made no blank responses.

Phonetic rather than orthographic responses were required as even phonetically trained subjects seem to be affected by lexical access issues when responding orthographically (Mannell, forthcoming).

3. Results

Subject lax vowel responses to short vowel plane tokens for each experiment were plotted onto the short vowel plane as 50%, 70% and 85% iso-perceptual contours using specially written software. The 50% contours are equivalent to Nearey's (1977) predominance boundaries and enclose a region where 50% or more of the subjects perceived the tokens as being a particular vowel. Similarly, the 70% and 85% contours enclose similar regions where at least 70% or 85% (respectively) of the subjects perceived a particular vowel (coloured with progressively darker greys from 50 to 85%). Figures 1 and 2 display the results for the 1988 and 2004 experiments respectively.

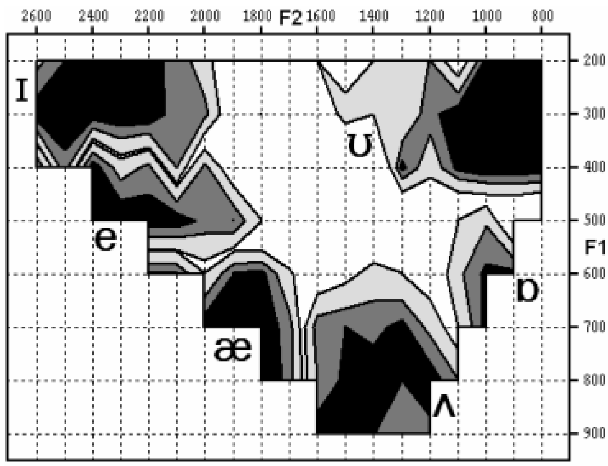

Figure 1. Lax vowel space for1988 short vowel perception data.

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after full paper review

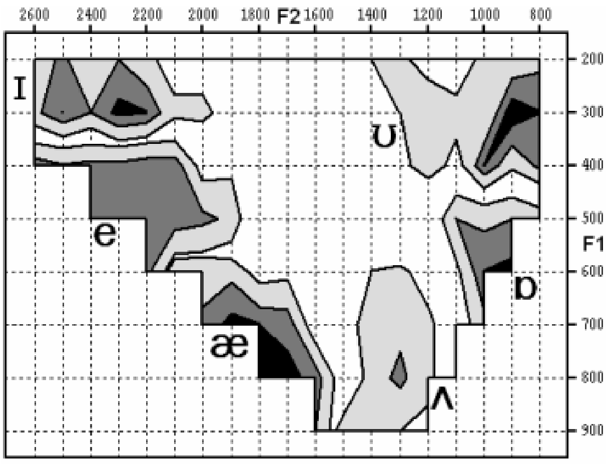

Figure 2. Lax vowel space for 2004 short vowel perception data.

Whilst there are a few types of differences between these vowel spaces, the features of greatest interest are the centres of each of the lax vowel regions and the positions of the boundaries between them. A χ^2 analysis of each data point for each vowel was carried out to determine the points on the maps that are significantly different between the 1988 and 2004 data. Figure 3 displays the points on the map that are significantly different for at least one vowel.

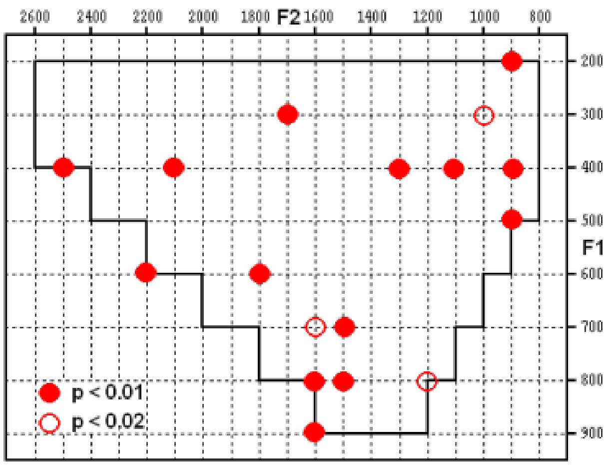

Figure 3. Points on the short vowel plane where the results for the 1988 and 2004 experiments differ significantly for at least one vowel ($\chi 2$ analysis, p<0.01 and 0.02)

The filled in red dots are points that are significantly different at p<0.01 and the unfilled red circles are the points that are significantly different at p<0.02. Note that at p=0.01, one would expect about one point out of the 112 data points to be spuriously significant as a consequence of chance. Note that the number of points (14) greatly exceeds what would be expected by chance for p<0.01. At p=0.02 about two such points would be expected by chance and the number of such points on this plot (3) closely matches this expected number. For that reason, p<0.01 is taken to be the alpha level for this analysis.

In figure 3, the points at F1=400, F2=2100/2500 differ significantly for both /ɪ/ and /e/. The point at F1=600, F2=2200 differs significantly for /e/ and /æ/. The point at F1=600, F2=1800 differs significantly for only /æ/, but if /ɜ/ and /ə/ responses are pooled then this point also differs significantly for /ɜ + ə/. The points at F1=800/900, F2=1600 differ significantly for both /æ/ and /ʌ/, whilst the points at F1=700/800, F2=1500 are only significant for /ʌ/. The point at F1=500, F2=900 is significant for /ɒ/ and /ɔ/ (that is, there is interference from a tense vowel). The points at F1=400, F2=900/1100/1300 differ significantly for only /ʊ/. The remaining two points F1=200, F2=900 and F1=300, F2=1700 are points affected by interference from the tense vowel /ʉ/. With the exception of the last two points, all of these significantly different points are at the boundary of two lax vowels. That is there are significant differences at the boundaries between /ɪ,e/, /e,æ/, /æ,ʌ/ and /ɒ,ʊ/. These can be seen to match visible differences between the two spaces displayed in figures 1 and 2.

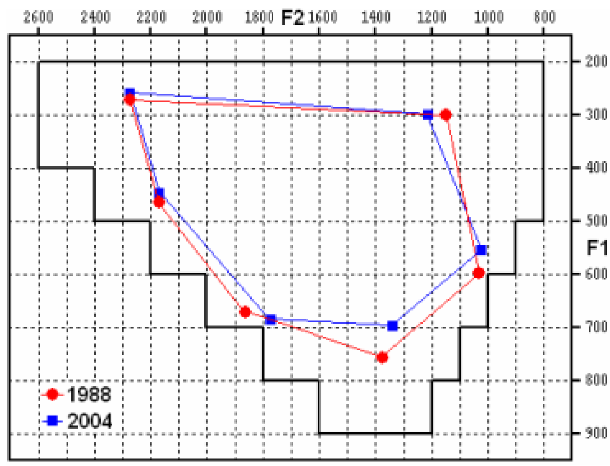

Figure 4. Plot of the grand means for each of the lax vowels for the 1988 and 2004 data

Figure 4 displays the grand means of the subject responses for each vowel for the 1988 and 2004 data. Each individual responded to a number of data points for each target vowel. For each vowel, the responses for each subject are fairly closely clustered in and around the predominance region of that vowel, but also typically spread out a bit into adjacent regions. For each of these 6 lax vowels, all subjects show similar skewed distributions of responses. This is to be expected as the most intense response regions (>85% response rates)

Accepted after full paper review

are usually found near the vowel space boundaries. It is desirable to derive from this data a measure of central tendency for each subject and vowel. As each subject's response pattern for a particular vowel is quite similar to the response pattern of every other subject for the same vowel (ie. approximately the same degree of skew) it seems justifiable to compute and to then compare the individual subject means for each vowel. Whilst the distribution of responses for each subject and vowel combination are skewed, the distribution of the individual subject means for each vowel are not skewed. For each of the lax vowels a t-test was applied to each of the F1 and F2 data for 1998 versus 2004. Significant differences between 1988 and 2004 were found for /æ/ F2 (p=0.007), /ʌ/ F1 (p=0.024) and /ɒ/ F1 (p=0.002). A MANOVA (Wilks' Lambda) test was also applied to each vowel. This test has the advantage of taking into account both F1 and F2 simultaneously. Significant differences between 1988 and 2004 were found for /æ/ (p=0.021), /ʌ/ (p=0.020) and /ɒ/ (p=0.007).

4. Discussion and Conclusions

This paper provides evidence for a shift upward in the perceptual boundary between /ɪ/ and /e/, a shift downward in the boundary between /e/ and /æ/, a retraction of the boundary between /æ/ and /ʌ/ and possible upward movement of the between the boundary of /ɒ/ and /ʊ/. Analysis of the mean responses for each vowel provides evidence for the retraction of the perceptual centre for /æ/, and a raising of the perceptual centres of /ʌ/ and /ɒ/. Whilst the /æ/ perceptual mean has only significantly retracted, the significant boundary shift includes a couple of lower data points which seem to suggest some degree of /æ/ lowering. No significant movement of the /e/ mean response has occurred, but this may be because both the upper and lower boundary of /e/ have moved (cancelling out an overall change in the mean). In any case, these subjects are older and the recent observed production shift downwards for /e/ is for adolescents. Otherwise, the observed vowel perceptual shifts appear to resemble the production shifts observed over the period of this study.

5. References

Bernard, J.R.L. (1967). "Towards the acoustic specification of Australian English", *Zeitschrift für Phonetik* 2/3, pp 113-128

Clark, J.E. (1981). "A low-level speech synthesis by rule system", *Journal of Phonetics*, 9, pp 451-476

Clark, J.E., Summerfield, C.D. and Mannell, R. (1986) "A high performance digital hardware synthesiser", *Proceedings of the First Australian Conference on Speech Science and Technology (SST86)*, Canberra, pp 342-347

Cox, F. (1996). *An acoustic study of vowel variation in Australian English*, Unpublished Doctoral Dissertation, Macquarie University.

Cox, F. (1998). "The Bernard data revisited", *Australian Journal of Linguistics*, 18(1), pp 29-55

Cox, F. (1999). "Vowel change in Australian English", *Phonetica,* 56, pp 1-27.

Cox, F. and Palethorpe, S. (1998). "Regional variation in the vowels of female adolescents from Sydney", *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP98)*, Sydney.

Cox, F. and Palethorpe, S. (2001). "The changing face of Australian English vowels", in Blair, D. and Collins, P. *English in Australia*, John Benjamins, Amsterdam

Cox, F. and Palethorpe, S. (2003). "The border effect" Vowel differences across the NSW - Victorian border", *Proceedings of the 2003 Conference of the Australian Linguistic Society*

Cox, F. and Palethorpe, S. (2004). "A question of broadness", a paper presented at the Australian Linguistic Society Conference, Sydney, July, 2004.

Harrington, J., Cox, F and Evans, Z. (1997). "An acoustic phonetic study of broad, general, and cultivated Australian English vowels", *Australian Journal of Linguistic*s 17, pp 155-184

Labov, W. (1994) *Principles of Linguistic Change, Vol 1. Internal Factors*, Blackwell, Oxford.

Lindblom, B. (1986). "Phonetic universals in vowel systems", in Ohala, J. and Jaeger, J. (eds) *Experimental Phonology*, Academic, London.

Mannell, R. (1988). "Perceptual vowel space of male and female Australian English vowels", *Proceedings of the Second Australian International Conference on Speech Science and Technology (SST88)*, Sydney, pp 22-27

Mannell, R. (1995). "Perceptual mapping and vowel normalisation", *Proceedings of the XIIIth International Congress of Phonetic Sciences (ICPhS95)*, Stockholm, Vol.2, pp 526-529

Millar, J., Vonwiller, J., Harrington, J. and Dermody, P. (1994). "The Australian national database of speken language", *Proceedings of the International Conference on Acoustics Speech and Signal Processing* (Vol 2), Adelaide, pp 67-100

Mitchell, A.G. (1946). *The pronunciation of English in Australia*, Angus and Robertson, Sydney.

Mitchell, A.G. and Delbridge, A. (1965a). *The speech of Australian adolescents*, Angus and Robertson, Sydney.

Mitchell, A.G. and Delbridge, A. (1965b). *The pronunciation of English in Australia*, Angus and Robertson, Sydney.

Nearey, T.M. (1977). *Phonetic feature systems for vowels*, PhD dissertation, University of Connecticut.