A Kinematic analysis of New Zealand and Australian English vowel spaces

Catherine Watson, Jonathan Harrington, and Sallyanne Palethorpe

Speech, Hearing, and Language Research Centre
Macquarie University, Sydney, Australia

ABSTRACT

The study presents a kinematic and acoustic comparison of the vowel spaces of New Zealand English (NZE) and Australian English (AE). Five talkers of NZE (3F and 2M) and 5 female talkers of AE each produced between 9 - 15 tokens of monophthongs in HEAD, HID, HAD, HERD. For each token, measurements of lip aperture, lip protrusion, jaw height, and tongue height and backness were made, in addition to formants and vowel duration being calculated. There were three main findings from the study. Firstly the widely held view that the NZE HID vowel may have lowered as it centralised may be incorrect. In this study results suggest that although the NZE vowel is retracted it remains a high front vowel, produced with highly approximated lips. The second result is HERD is lip rounded in both NZE and AE. The final finding is that there is no significant difference in vowel duration between NZE and AE HEAD and HAD, despite NZE HEAD and HAD being phonetically more raised. The implications of these results are discussed.

1. INTRODUCTION

This study is concerned with an articulatory comparison of the vowel spaces of New Zealand English (NZE) and Australian English (AE). It is a development of earlier investigations in which we compared contemporary NZE vowels with both AE vowels [1] and with NZE vowels recorded from talkers in 1948 [2]. In these studies we found, in compatibility with earlier impressionistic studies, that the contemporary NZE lax vowels in HAD and HEAD are raised, while HID is centralised relative to AE. We also found that the 1948 NZE HID was much less centralised than its contemporary counterpart. This suggests that the NZE HID centralisation, which is one of the most distinctive markers of the NZE/AE accent difference, may have taken place in the last 50 years.

The present kinematic and acoustic analysis was designed to extend our earlier NZE/AE acoustic investigation in three ways. Firstly, we wished to analyse the extent of tongue height differences in the lax front vowels between the accents and to investigate further whether the NZE HID has lowered as well as centralised. Secondly, we aim to test claims in the literature (e.g. [3]) that the NZE HERD is produced with lip-rounding. Finally, following the expected inverse relationship between vowel length and phonetic height, we investigated whether the durations of HEAD and HAD have decreased in NZE to accompany their raising relative to AE.

To avoid confusions that would arise from distinct phonemic transcriptions of NZE and AE vowels we will use a “hVd” lexical set to refer to the vowels of interest.

2. METHOD

2.1.Talkers and Materials

Five talkers of NZE (3F and 2M) and 5 female talkers of AE each produced tokens of /hVd/ monophthongs. For this study we were interested in the four monophthongs HID, HEAD, HAD, and HERD. All NZE subjects produced these monophthongs. Two AE talkers produced all four monophthongs; a further three subjects produced three of these, two produced no HID tokens, and one produced no HAD tokens. The AE data was taken from different corpora recorded in the 1996-1998 period. Between 10 -15 tokens for each monophthongs were produced by the NZE subjects, and between 9-10 tokens were produced by the AE subjects. All words were produced in citation form, and the order of the words was randomised each produced tokens of /hVd/ monophthongs.

2.2 Recordings Conditions

The Subjects were recorded in a sound proof room in the Speech Hearing and Language Research Centre, Macquarie University. A microphone was used to record the acoustic speech signal, and a MOVETRACK electromagnetometer [4] was used to record the movement of the upper and lower lips and the tongue-dorsum. Recording the movement of the four articulators was achieved by measuring the movement of receiver coils through an alternating magnetic field. The first transducer coil was placed 1.75 cm back from the tip of the tongue. The remaining three transducers were then placed at the midpoint of the upper and lower lips on the vermilion border and on the chin (to give jaw position). The x- and y- axis values of the receiver coils were measured relative to fixed transmitters mounted on a helmet behind and above the head. This produced 8 kinematic signals for analysis (for more detail on the experimental setup see [5])

2.3. Digitization, and Acoustic Labelling

The acoustic and kinematic data were digitized directly to a SUN workstation at a sampling frequency of 20 kHz, and 500 Hz respectively, and were quantized to 16 bit numbers. The ESPS/Waves+ system was used for acoustic segmentation, labelling, and formant tracking. The first four formants and their bandwidths were automatically tracked (the settings were 12^{th} order LPC analysis, cosine window, 49 ms frame size, and 39.5 msec frame overlap). The formant tracks were checked for accuracy and hand corrections were made. The onset and offset of the vowel, and the target were marked according to criteria given in [6]. All subsequent analysis were carried out using the EMU system for speech database analysis [7].

3. RESULTS

The nature of kinematic data is such that it is best to analyse the results on a speaker by speaker basis. Due to a lack of space in this study it is not possible to show the results for all speakers. Therefore we selected two speakers (both female), one for NZE (henceforth called nz1) and one for AE (henceforth called au1) and give their results in Figures 1 and 2 respectively. However these results are representative of all the speakers studied. Our discussion of the results will refer to all the subjects.

3.1. Formant Analysis

The most major difference between the NZE and AE subjects was that HID is a central vowel in NZE and a high front vowel in AE A second notable difference is that HEAD and HAD are raised in NZE compared with AE. The final difference is that HERD in NZE is more fronted and raised than in AE. These differences in NZE and AE can be observed by comparing Figures 1 and 2 (a) , the F1/F2 plots of the class centroids of HEAD, HID, HAD and HERD for nz1 and au1 respectively. The formant values were taken at the vowel target. The findings are all in accordance with NZE and AE differences noted in earlier studies (e.g. [1],[3]).

3.2. Vowel Duration

Vowel	**Duration of vowel (msec)**	
	NZE	AE
HID	131±32 (55)	123±29 (29)
HEAD	164±31 (55)	154±33 (49)
HAD	190±36 (55)	196±26 (39)

Table 1: The mean and standard deviations of the vowel durations of HID, HEAD and HAD for 4 of the 5 NZE subjects *(left)* and all the AE subjects who produced tokens the vowel class (*right*). The number of tokens used to calculate the duration statistics is given in brackets.

The vowel duration was calculated to be the difference in time between the vowel offset and onset. Table 1 gives the mean and standard deviation of HID, HEAD, HAD for 4 of the 5 NZE subjects and the AE subjects. The number of tokens used to calculate the duration statistics is given in brackets. The tokens of the fifth NZE subject were removed from this analysis because the duration of their HID, HEAD, and HAD were all significantly longer than the other four NZE subjects ($\alpha > 0.01$).

The durations of tokens of the remaining four NZE subjects, for HID, HEAD and HAD were not significantly different from those produced by the AE speakers ($\alpha > 0.01$). Since the NZE HEAD and HAD are more raised than in AE we would expect that they would have decreased in duration following the expected inverse relationship between vowel duration and height.

3.2. Lip and Jaw Movement

Lip protrusion is measured by the displacement of the upper lip along the x-axis, the smaller the value the more protruded the lips. The results suggested that the lips were protruded in the production of HERD for all the NZE and AE subjects. This suggests lip rounding for HERD. Figures 1 and 2 (b) gives the mean values the upper lip protrusion for nz1and au1respectively. For each vowel class the trajectories have been aligned at the vowel target, averaged, and truncated at the mean distance between the vowel target and onset, and the vowel offset and target. It can be seen for both nz1 and au1 that HERD is clearly separated from HID, HEAD and HAD, and is in the bottom left corner. This is evidence that HERD is lip rounded.

A measure of the lip aperture was taken by estimating the difference between the upper and lower lip in the y -axis (the smaller the value the more compressed are the lips). The results suggested that HERD and HID were produced with more compressed lips than HEAD for both the NZE and AE. In NZE HID was produced with more compressed lips than HERD for 4 out of the 5 speakers. Figures 1 and 2 (c) give the mean lip aperture for nz1 and au1 respectively. The trajectories were calculated in exactly the same way as the lip protrusion trajectories in Figures 1 and 2 (b).

The jaw height data (not illustrated) supports the lip aperture data. In both NZE and AE, HERD and HID had a higher jaw position than HEAD. The vowel HAD was produced with the lowest jaw position for all subjects of both NZE and AE.

Contrary to the evidence in the formant plane plot (Figure 1(a)), NZE HID had a higher jaw height than for HEAD. The kinematic data suggests, therefore, that HID is phonetically closer than HEAD in NZE as in AE. The finding that HERD has a greater jaw height than HEAD in both NZE and AE is probably due to HERD being produced with lip rounding.

3.3. Tongue Position

Figures 1 and 2 (d) give the class centroids of the position of the tongue dorsum extracted at the acoustic vowel target in the x-y plane: lower values along the x axis indicate greater tongue dorsum fronting, higher y-axis values indicate tongue dorsum raising.

The F1/F2 analysis for the NZE subjects (see for example Figure 1(a)) suggests that HID for 4 of the 5 subjects is a centralised vowel with lower height than both HERD and HEAD. However the tongue backness vs. tongue height plots (see for example Figure 1(a)) suggest that, although the tongue is more retracted for HID than for HEAD and HERD (for 4 of the 5 subjects) HID, HEAD, and HERD are of similar height. Thus in NZE HID and HEAD are differentiated mostly on backness and NZE HERD and HEAD on backness, length, and rounding.

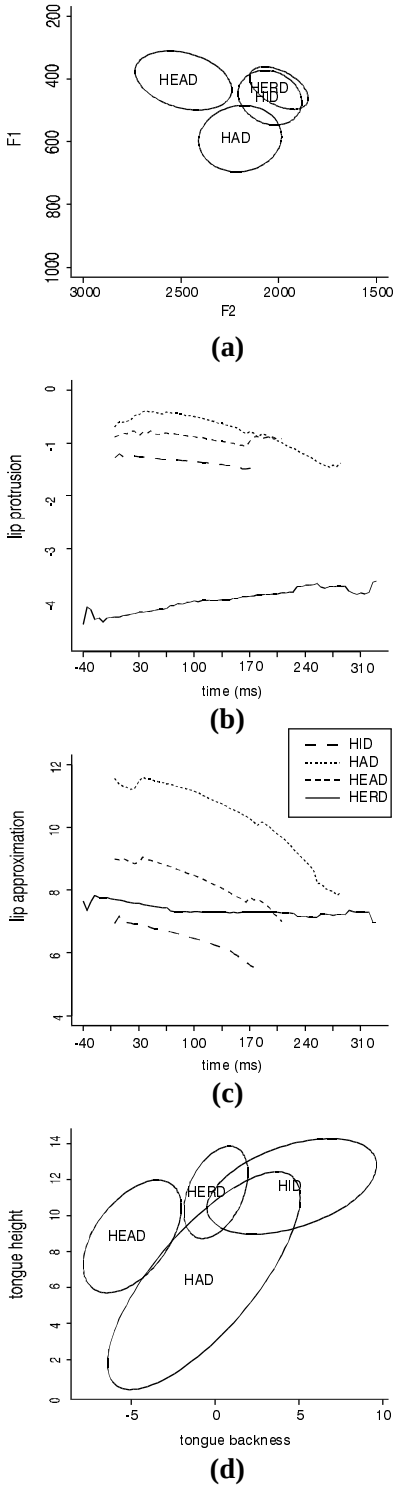

Figure 1 The data for subject nz1 producing the vowels HEAD, HID, HAD, and HERD: (a) the F1/F2 plot of the class centroids of for each vowel class (b) the averaged trajectories of the upper lip protrusion, aligned at the vowel target for each vowel class; (c) the averaged trajectories of the lip aperture, aligned at the vowel target for each vowel class; (c) the tongue backness and tongue height at the vowel target for each vowel class.

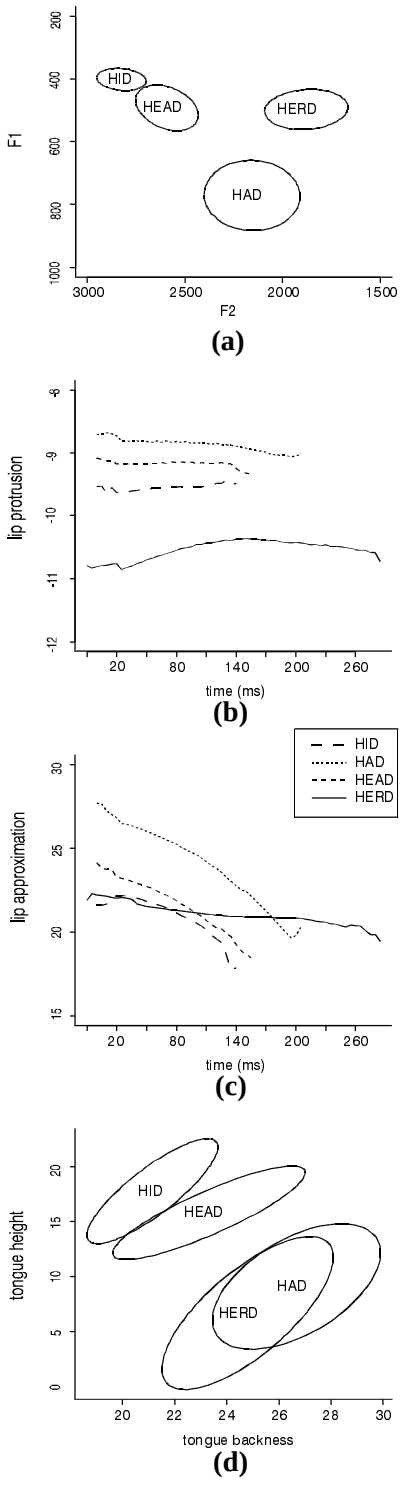

Figure 2 The data for subject au1 producing the vowels HEAD, HID, HAD, and HERD: (a) the F1/F2 plot of the class centroids of for each vowel class (b) the averaged trajectories of the upper lip protrusion, aligned at the vowel target for each vowel class; (c) the averaged trajectories of the lip aperture, aligned at the vowel target for each vowel class; (c) the tongue backness and tongue height at the vowel target for each vowel class.

It was not as easy to draw so many observations from the tongue data of the AE subjects. Only 2 of the 5 subjects yielded sensible data, and of these only one (au1) produced all four tokens of interest. Nonetheless some observations can be made. It can be seen in Figure 2(d) that HID is the most fronted vowel for au1 , and has at least the same tongue height as HEAD. The tongue backness vs. tongue plot for au1 shows a greater compatibility with the F1/F2 plot for the same speaker (c.f. Figures 2(a) and (d)). This was not the case for the NZE subjects.

4. DISCUSSION

The most interesting finding of this study is that the widely view that NZE HID has lowered as it centralised may not be correct. It is perhaps the case that NZE HID has backed but not fallen and has kept the closely approximated lips typically associated with a front high vowel as it was 50 years ago in NZE. Since HID in contemporary AE is a high front vowel produced with closely approximated lips the results of this study suggests that the difference in the tongue height between the contemporary NZE and AE HID is not as great as earlier F1/F2 analyses have indicated. Further, the results suggest that a phonemic transcription for the NZE HID with a barred [i] (cardinal vowel 17) might be more appropriate than schwa (which has been suggested in many early studies e.g [1],[3]) which implies the NZE HID has lowered.

The results of this study also show that HERD is lip rounded in both NZE and AE The high degree of vertical lip approximation, observed in HERD for both NZE and AE may just be a function of lip-rounding since it is nearly impossible to round the lips without approximating them. The kinematic data show that HERD in NZE is almost certainly retracted relative to HEAD, the tongue height data are too ambiguous for any conclusions to be made.

The final finding of this study is that NZE HEAD and HAD have a non-significant difference in vowel duration compared to their AE counterparts and this may imply that HEAD and HAD have not shortened as it raised (a shortening might be expected in view of the expected inverse relationship between vowel height and duration [9]). If NZE HEAD has not shortened but has raised, it may be perceived to be a good deal more tense than forty years ago.

This study has some implications for the extended drag-chain vowel shift proposed for NZE whereby HID moved first, and HEAD and HAD followed [1]. In this study we are suggesting that HID has not lowered , as implied by F1/F2 data. This suggests that the raising of NZE HEAD and HAD is a extended push-chain. But then if that is the case it is not entirely clear why the NZE HEAD and HAD are more raised than the AE HEAD and HAD since HUD and HARD for both accents occupy a very similar place in F1/F2 space.

In summary, the study has shown that further kinematic data are necessary to supplement the formant data and impressionistic analyses on which vowel changes are usually based. This applies in particular to NZE HERD whose formant values due to tongue position are confounded with the effects of lip-rounding (which tends to lower formants). Equally, researchers may have assumed too readily that NZE HID fell as it backed resulting in a modern-day schwa-like quality, not only because the formants are compatible with those that are expected from a central, mid vowel, but also because this is the vowel in English to which NZE HID seems to be closest in quality (i.e. phoneticians perceive NZE HID to be quite similar to a shortened version of Southern British English production of 'bird' and accordingly label it as schwa). However, while the kinematic data provides clear evidence that NZE has backed, there is no evidence that it has fallen, nor that the lip configuration has changed from an approximated position that would mark it as a relatively high vowel.

5. References

1. Watson, C. , Harrington, J. and Evans, Z., "An acoustic comparison between New Zealand and Australian English vowels." *Australian Journal of Linguistics*, in press.

2. Watson, C., Maclagan, M., and Harrington J. "Acoustic evidence for vowel change in New Zealand English" , Paper presented at *Laboratory Phonology 6*, York, July, 1998

3. Wells, J. C. , *Accents of English: Beyond the British Isles*. Cambridge: Cambridge University Press, 1982.

4. Branderud , P. " Movetrack-a movement tracking system", *Perilus IV*: 20-29,1985.

5. Harrington, J., Beckman, M.E., Fletcher, J., and Palethorpe, S. "Kinematic and spectral measures of supralaryngeal correlates of the accent contrast in Australian English vowels", *Journal of the Acoustical Society of America*, submitted.

6. Harrington, J., Cox, F. M., and Evans, Z. "An Acoustic Phonetic Study of Broad, General, and Cultivated Australian English", *Australian Journal of Linguistics* 17: 155-184,1997.

7. Cassidy, S. and Harrington, J. "EMU: an enhanced hierarchical speech data management system", In *Proceedings of the 6th International Conference on Speech Science and Technology*, Adelaide,361-366,1996.

8. Lehiste, I. and Peterson, G. "Transitions, glides and diphthongs",*Journal of the Acoustical Society of America*: 33 268-277,1961.