ACOUSTIC MEASUREMENTS OF THE DIPHTHONGS OF WOMEN SPEAKERS OF GENERAL AUSTRALIAN ENGLISH

L. Penny

Department of Speech Pathology
Flinders University of South Australia

ABSTRACT — Formant and duration measures of the general variety of Australian English, as spoken by healthy young native-born women speakers are presented. The findings are discussed in relation to Bernard's (Bernard & Mannell, 1986) data for male speakers, and the implications for views on regional and social varieties of Australian English are discussed. The data may be used to confirm Clark's (Clark, 1989) revised transcription system for Australian vowels.

INTRODUCTION

The acoustic specification of Australian speech is of interest to linguists concerned with the description of Australian speech varieties, and also to speech pathologists, concerned with the diagnosis and remediation of disordered speech. In the first case, questions concern the history, development and differentiation of Australian English and its regional and social varieties, while the latter involves the acceptability and effectiveness of speech communication. Vowel data are of considerable interest because it is in particular vowel colour, movement and duration that a good deal of regional and social variation resides (Clark, 1989). Speech pathologists may be confronted with inappropriate pitch, intonation or tempo, significantly involving vowels. Quantitative acoustic vowel analysis is essential to confirm and refine auditory descriptions on the one hand, and to diagnose and monitor the status of individuals on the other. Diphthongs are of special interest since they are recognized as carrying considerable weight in the perception of Australian language varieties, and a number of studies (Mitchell, 1946, Mitchell and Delbridge, 1965, Bernard, 1989, Osawa, 1989, Clark, 1989) suggest that they may be among the most variable of Australian speech sounds.

Diphthongs are vocalic nuclei containing two distinct targets connected by glide. This latter is occasioned by the articulators moving rapidly from the steady-state position of the first target to assume the position appropriate for the second. Diphthongs are distinguished on the one hand from relatively pure, single target vowels and in the other from sequences of two vowels, each the nucleus of its own syllable, as in "being", /biɪŋ/. Diphthongs are single segments conventionally represented by two symbols indicating their assumed auditory targets. Eight diphthongs are traditionally recognized in Australian English, three gliding to /ɪ/, /eɪ, aɪ, ɔɪ,/, two to /ʊ/, /aʊ, oʊ/ and three to /ə/, /ɪə, ɛə, ʊə/.

Normative data of the acoustic measures of the diphthongs of healthy young women speakers of known social and regional backgrounds (South Australia) are therefore of interest. Bernard (Bernard and Mannell, 1986) has presented comprehensive data for male speakers and Penny (Penny 1991) has previously offered single target data for female speakers.

METHOD

Forty-eight young women (mean age 22 years and 6 months), native-born speakers of Australian English each supplied three tokens of each of 19 vowels (single target and diphthongs) in citation form, in /h__d/ frame. Judgments as to their speech being of the "general" variety were made by Penny, following the descriptions of Mitchell and Delbridge (Mitchell and Delbridge, 1965). Qualified speech pathologists had already assessed the subjects as having normal speech, and their hearing had been tested as within normal limits. The recordings were made in a sound-treated room using a Sony digital audio-tape recorder TCD-D10, maintaining a constant mouth-to-microphone distance. The recordings were analysed using a Kay DSP5500 spectrograph with both wide and narrow bands, and an amplitude trace available. Formant centres and component durations were estimated by eye, with the aid of cursors. The problem of identifying formant centres in women's speech was addressed (in part) by having two judges independently take measures in those cases where greatest difficulty is to be expected. some 34% of readings were checked in this way. The targets were identified with the aid of the time cursor which was manipulated to section the different components into steady state or transition areas.

RESULTS

Figure 1(a) shows (top) the vowel space of the women's single target vowels and (bottom) the positions of the two targets and the glides of their diphthongs. Figure 1(b) shows the corresponding vowel spaces for Bernard's male general speakers (Bernard & Mannell, 1986), the areas of each target are 2 standard deviations in size. Figure 2 is a highly schematic plot of the first three formants and the glides of each diphthong, for both women and men. Vowel numbers follow Bernard's sequence.

The women's data

Glides to /ɪ/

/eɪ/ The present study finds a somewhat variable first target between /æ/ and /a/, while the second target is neatly in the region of /ɪ/. The two targets are of almost the same duration, and both are shorter than the glide. There is little variability in the length of this phoneme. Bernard's results are very similar.

/aɪ/ The initial, steady state target for the women is slightly retracted from /a/ towards /ɒ/, and is quite consistent. A rapid glide covering a long distance takes the expression towards /ɪ/, but slightly lowered. The duration of this second target is short and variable, but the over-all duration of this phoneme is consistent. It is the longest of all phonemes for the women (but not for men). The men's data show a similar initial position, and a glide running towards /ɪ/ but falling short of it. Bernard also found the second target to be inconsistent, and the glide to carry considerable perceptual clout.

/ɔɪ/ This phoneme moves, for the women, from a tightly defined initial target near /ɔ/, but a little lower, via a rapid glide, over a long distance, to a position only slightly retracted from /ɪ/. The three components contribute almost equally to the total duration, and it is the shortest of all the women's diphthongs. Bernard's data show a very similar pattern. This is a strong, consistent phoneme.

In sum, the results for these three phonemes in the present study are not very different from Bernard's (Bernard & Mannell, 1986), though Bernard's glides tend to be longer in duration and the second target shorter. In all cases the nominal second target is approached, but slightly undershot. The women, however get closer.

Glides to /ʊ/

/aʊ/ In the present study this phoneme has a low and variable first target between /æ/ and /a/. A pronounced glide moves to a position coinciding fairly closely with /ɒ/. This second target is relatively short, but is strongly expressed. In comparison, Bernard's subjects start from a target yet more fronted, and also proceed in the direction of /ɒ/ but fall short of it. This target is much lower, for the men than for the women, so the slopes of the two glides are rather different.

/oʊ/ The first target, held for a short but fairly consistent time, is somewhat fronted from /ɒ/. A strong glide moves up and forward to a consistent target fronted from /u/. The over-all duration of /oʊ/ is somewhat variable. Bernard's subjects start from a position fronted and higher than /a/ and move to a glide also fronted from /u/. Durations are similar in the two groups. Bernard may be correct in supposing that it is the /h__d/ frame that fronts the second target, but certainly that the nominal second target symbol is hardly accurate, is a finding of both studies.

Glides to /ə/

/ɪə/ The first target is consistent and in the position of /ɪ/. It is quite stable, and the longest of the three components. The glide moves only slightly, to a second target in the direction of /ɛ/. Bernard's subjects also start coincident with their single target /ɪ/. The second target moves towards their /ɛ/ position. The glide in both cases is short, the shortest of all glides. Several cases in the women's data in fact had no glides at all, and the realization of this phoneme is as a long, rather unstable /i/.

/ɛə/ Again, the women's data suggest that this phoneme is not always realized with two distinct targets. The targets over-lap, occupying an enlarged /ɛ/ position. In some cases the only sign of a second component was a region of distinctly reduced amplitude. The duration plot shows 3 sections

but this conceals the fact that there are 2 realizations of this phoneme - a diphthong, or (in about 30% of cases) a long, single but complex target.

/ʊə/ The women's expressions start from a target which may be described as a fronted /ʊ/ with a variable second formant. Movement is to a central position, just retracted from /ɜ/. This is, for women, the most variable phoneme in length. The male data show a similar but less variable starting position, and after a short glide (in both distance and time) the finish is central.

In sum, the so-called centring diphthongs show, in both studies, a weak and variable second target. The diphthongal status of the three phonemes is questionable, though /ʊə/ in the women's data is holding to this more strongly than the other segments.

Over-all, the women appear to use a greater vowel space than the men, and this appears in both single and two-target segments. The greater dimension of difference is front-to-back. There is a general tendency for women to realize both targets of diphthongs more evenly than men, and for women to "undershoot" the second target less. A trading-relationship may explain the generally shorter duration of the women's glides. On-glide measures are not reported since almost all first targets start well on target and it was not possible to break the glide into sections corresponding to an off-glide from the first and an onglide into the second target. The glides were smooth transitions. The final off-glide is taken to belong more to the final stop than to the second target, and is not reported.

DISCUSSION

The differences between Bernard's results and the present study lie mainly in the more determined expression of both targets by the women, and their greater vowel space. The over-all impression of the women's speech is of greater care. Perhaps the women have been misidentified as "general" speakers? Perhaps there may be inherent differences in the speech of men and women? There may be regional differences within the general variety, or there may have been changes with time in the expression of this variety, for Bernard's samples and the present are separated by some 25 years.

The subjects were classified by Penny, as general speakers, but they all also independently identify their own speech as of the general variety. It is true that they display some of the markers of cultivated speech - if care, being on target, and vowel space size are considered such markers, but they also display features of the general part of the speech continuum, especially their shorter more rapid glides and their realizations of /aɪ/ and /eɪ/.

Women speakers are often regarded as conservative in the sense of maintaining pressure towards the cultivated end of the spectrum. The women here regard cultivated speech as a charming relic, to be appreciated (when genuine) but not emulated. These subjects are not phonetically naive (as were Bernard's) and were concerned to produce speech samples as natural as is possible, given the rather un-natural contexts of the study; they are general speakers. It is perhaps unfortunate that Mitchell and Delbridge identify "care" with cultivated speech. It well might be, but at the same time there seems no reason why people with general variety target locations and dynamics could not also be careful, starting on target (as the women do) and mostly reaching the second target, excepting in the case of the weak centreing /ə/ where the diphthong, as with Bernard's finding, appears to be changing to a long, complex single target. What is suggested here is a *careful* general variety.

The question of regional differences is hard to determine. Oasa's (Oasa, 1989) data for Adelaide women are based on N=7, and perhaps this small sample produced atypical results, for the present study does not find the differences between reports. The data are difficult to compare though, for the contexts are different. The evidence for regional variety must be regarded as sketchy and uncertain.

It is possible to use the present data to confirm the auditory basis of Clark's proposals for a revised transcription of Australian English. Clark abandons the traditional classificatory system and opts for a system employing either simple or complex target types, further classified into long or short. In this system, the 8 traditional diphthongs are all classed as long vowels with complex targets.

/eɪ/ is replaced by [æe] for broad and [ɛe] for cultivated - the traditional symbols fit rather better than Clark's suggestions, but neither is very good.

/aɪ/ is replaced by /ɐe/, /ɑe/, /ɔe/, /ɑə/ The third variant, which Clark finds in some broad speakers, is very close to our data.

/ɔɪ/ becomes (broad) /oɪ/ or /oə/ before /l/. The traditional symbol is better since the first target is closer to /ɔ/ than /o/.

/aʊ/ becomes /æɔ/ or /oə/ (broad) Clark's suggestion is much better for our data. The first target is a little lower, and retracted from /æ/ and the glide is to /ɔ/.

/oʊ/ becomes /əʉ/ or /ɐʉ/ (Broad), /ɔo/ (before l). Our findings are much better symbolized by Clark's second variant than by the traditional notation.

/ɪə/ is /ɪᵊ/ or /ɪə/ or /ɪ:/ Clark's second suggestion fits our data very well. The first target is dominant, and the second target weak and marginal.

/ɛə/ becomes /ea/ /eᵊ/ /e:/ We confirm Clark's observation that this segment has forms paralleling thos of /ɪə/, and that there is a tendency to trade the second target for length.

/ʊə/ is /ʊə/, /ʉə/ /o:/ The first variant is confirmed in our study; although the starting position is a fronted /ʊ/ it is certainly not /ʉ/. There is little danger in our data of this segment losing contrast with /ɔ/ (/o:/ in Clark's system).

CONCLUSION

The vowel data submitted here raise as many questions as they answer, in part because of the limited context in which they were elicited. It is hoped that they allow, at least, some comparisons to be made from a common base-line. It is entirely possible, indeed probable that diphthongs are realized differently in closed syllables compared with open syllables, and that in the latter they are more robust.

NOTE

Thanks are due to Kate Wildy for checking the formant estimates and to Ralph Richardson for designing the figures.

BIBLIOGRAPHY

Bernard, J.R. (1989) *Quantitative aspects of the sounds of Australian English*, in Collins, P. & Blair, D. (eds) Australian English, 1989, (St Lucia: University of Queensland Press).

Bernard, J.R. (1989) *Regional variation in Australian English: a survey*, in Collins, P. & Blair, D. (eds) Australian English, (St Lucia: University of Queensland Press).

Bradley, D (1989) *Regional dialects in Australian English*, in Collins, P. & Blair, D. (eds) Australian English, (St Lucia: University of Queensland Press).

Clark, J.E, (1989) *Some proposals for a revised phonetic transcription of Australian English*, in Collins, P. & Blair, D. (eds) Australian English, (St Lucia: University of Queensland Press).

Collins, P & Blair, D (1989) (eds), *Australian English*, (St Lucia: University of Queensland Press).

Horvath, B. (1985) *Variation in Australian English: the sociolects of Sydney*, (Cambridge: Cambridge University Press),

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

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

Oasa, H. (1989) *Phonology of current Adelaide English: a survey*, in Collins, P. & Blair, D. (eds) Australian English (St Lucia: University of Queensland Press).

Penny, L. (1991) *Acoustic measures of the single target vowels of Australian women general speakers*. 8th Australian Language and Speech Conference, Melbourne.

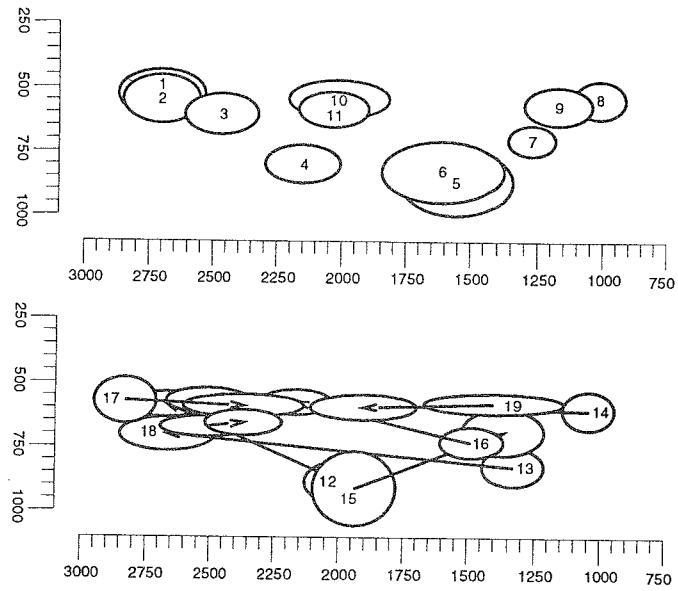

figure 1(a). Women's vowel spases, (top, single target; bottom, diphthongs)

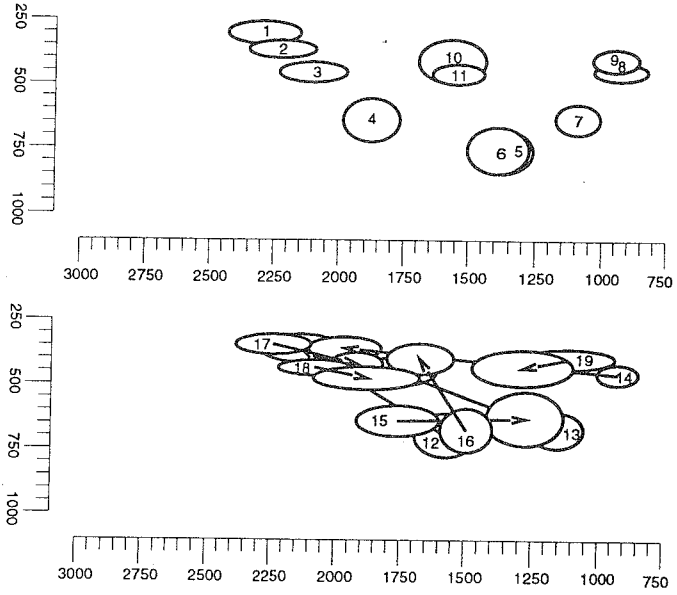

figure 1(b). Men's vowel spaces, Bernard's data (top, single target; bottom, diphthongs)

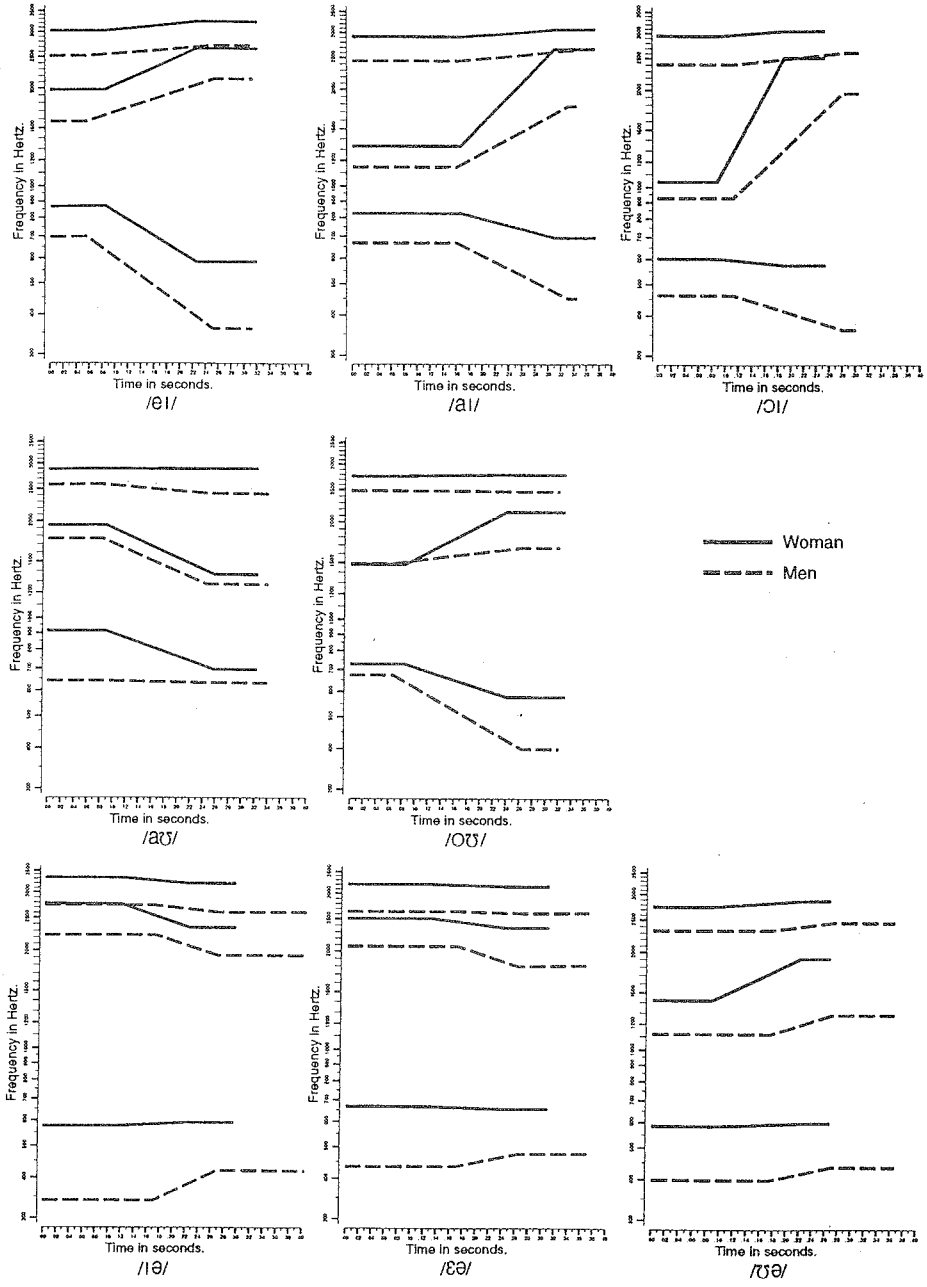

figure 2. Schematic plot of the first three formants and the glides of each diphthong.