AN ACOUSTIC-PHONETIC COMPARISON OF ENGLISH DIPHTHONGS PRODUCED BY NATIVE SPEAKERS OF TAMIL AND AUSTRALIAN ENGLISH

Santha Sampath, David Slater and Michael Wagner

Department of Computer Science
University College
Australian Defence Force Academy
University of NSW

ABSTRACT - This paper describes a comparative study of diphthongs produced in two different dialects of English, namely that spoken by native speakers of Tamil and that spoken by native speakers of Australian English. Similarities as well as some notable differences between these speaker groups have been found in the production of certain diphthongs. Further differences have also been observed between diphthongs produced by Tamil speakers who have been resident in Australia for a short time and those who have been resident for a long time. The small number of diphthongs occurring in Tamil is suggested as a possible explanation for some of the contrasts reported herein.

INTRODUCTION

Some significant differences in the vowel distribution in the F1 - F2 space were described in a previous study of English monophthongs produced by South Indian migrants whose native language is Tamil and by native speakers of Australian English (Sampath, Slater, Oasa & Wagner, 1991). Speaker and dialectal differences are expected to be also pronounced in diphthongs, since these complex vowels are characterised by transitions to and from two distinct vowel regions (Clermont, 1988 & 1991). The aim of this work is therefore to compare the two groups of speakers in terms of their respective formant patterns of diphthongs.

There are eight informants of whom four are native speakers of Tamil (these form a group of non-native speakers of English) and the other four are native speakers of Australian English. Each group consists of two males and two females. Among the Tamil informants, two have been resident in Australia for more than twenty years and the other two have been resident in Australia for two years. It is worth noting that the Tamil language has only two diphthongs, /ai/ and /ou/.

EXPERIMENT

Speech material

Eight English sentences, which include most of the phonemes in Australian English, were carefully chosen for the experimental data. These sentences are listed below. The words used in this paper, taken from seven of these sentences, are underlined.

1. People say that things could be better.
2. It is good to choose well.
3. You should start now, but that may not be possible.
4. He came here for a change of job.
5. Please buy the right toy for the boy.
6. It was a pleasure to chair the meeting.
7. There is no reason to hurt her feelings.
8. She came to the end of her tour of duty.

Each of the eight sentences was typed on a card, each informant was given these sentences in a random order and each sentence was spoken and recorded. This process was carried out ten times so that each informant recorded eighty sentences in a single recording session. The sentences were recorded in a quiet environment. The Interactive Laboratory System (ILS) software package was used to digitise the speech data with a sampling frequency of 10000 Hz. From these eight sentences, eight monosyllabic words containing diphthongs were chosen for analysis. Table 1 gives a summary of the target diphthong types and the words used, together with the IPA representation for each word. Diphthongs can be classified in terms of the second vowel target. In English, this classification divides diphthongs into three types, which are referred to here as target diphthong types.

Target Diphthong Type	Word	IPA Representation
/ɪ/ diphthongs	buy	/baɪ/
	boy	/bɔɪ/
	say	/seɪ/
/ʊ/ diphthongs	no	/noʊ/
	now	/naʊ/
/ə/ diphthongs	chair	/tʃɛə/
	here	/hɪə/
	tour	/tʊə/

Table 1. Words used in this paper

ACOUSTIC ANALYSIS

Formant extraction

Using a visual display of the waveform of each word, the diphthongal nucleus boundaries were marked, and then confirmed by listening to each token several times. Within these boundaries acoustic parameters such as the fundamental frequency (F0), the first four formants (F1, F2, F3 and F4) and the bandwidth associated with each formant were automatically extracted using the Linear Predictive Coding (LPC) autoregression coefficients method available within ILS. For the analyses, the data were obtained at eleven equidistant points in time throughout the diphthong. These points are termed percentage points of diphthong duration, with the 0% points of diphthong duration being the estimated beginnings of the diphthong nuclei and the 100% points of diphthong duration being the estimated endings of the diphthong nuclei. The automatic extraction of formants for some of the female informants was not always successful. In these cases the formant frequencies were obtained by visual inspection of the spectrum. Occasionally, the formant frequencies near the start and end of the diphthong nuclei had to be discarded because unreliable values were obtained due to possibly incorrect LPC analyses.

Formant charts

The analysis of variations in formant frequencies is especially important as it is expected to give some indication of the vowel-to-vowel movement which characterises diphthongs. Some examples of these variations of formant frequencies for different diphthongs are shown in figures 1 to 5 below.

The values of F1, F2 and F3 for the percentage points of diphthong duration shown in these figures are obtained by averaging over the ten tokens. It is found that for native speakers of Australian English, the temporal variations of F2 for all the diphthongs are similar to those found in the literature (Bernard, 1989 and Clermont, 1991). Variations in the F2 patterns reflect either an increase or decrease in frequency, with respect to time, dependant upon the target diphthong type. In most cases a transition seems to be taking place at about the 50% point of diphthong duration. This is also true, on the whole, for the Tamil speakers although there are some marked differences which are discussed below.

DISCUSSION OF THE RESULTS

Similarities between the two groups

The formant patterns of the diphthong /ɔɪ/ in the word "boy" are found to be similar for both groups. Figure 1 gives an example from each group. The Tamil speaking informant, whose formant pattern is shown in figure 1, has been a resident of Australia for only two years and one can assume that the influence of Tamil is maximal in her case. It should be pointed out that, for the native speaker of Australian English, formant frequency values at the 90% and 100% points of duration had to be discarded as the LPC analysis gave erroneous results. As this occurred for all ten tokens recorded, it is possible that the actual end of the diphthong nucleus is closer to the 80% point of duration shown on the graph, thus suggesting that the manually selected offset boundaries were too generous. If this is the case, then the second vowel target of the diphthong /ɔɪ/ in "boy" has not been realised by this speaker.

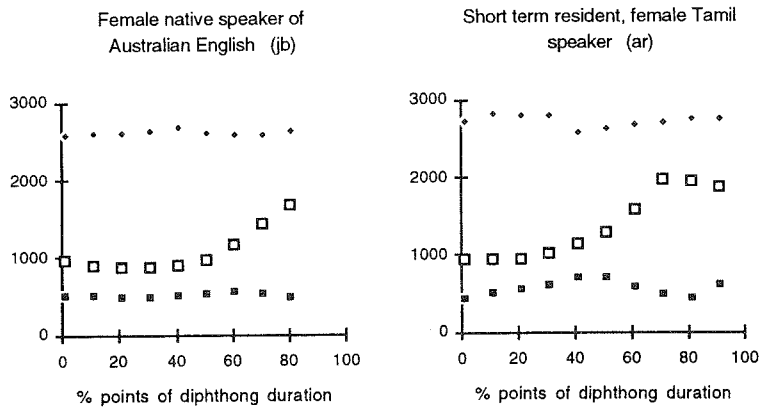

Figure 1. F1, F2 and F3 formant patterns (in Hz) for the diphthong nucleus of the word "boy"

The formant patterns of the diphthong /aɪ/ extracted from the word "buy" are found to be similar for both groups. However, the transition seems to take place at an earlier point for three of the Tamil informants when compared with native speakers of Australian English. This can be seen, for one Tamil speaker, in figure 2.

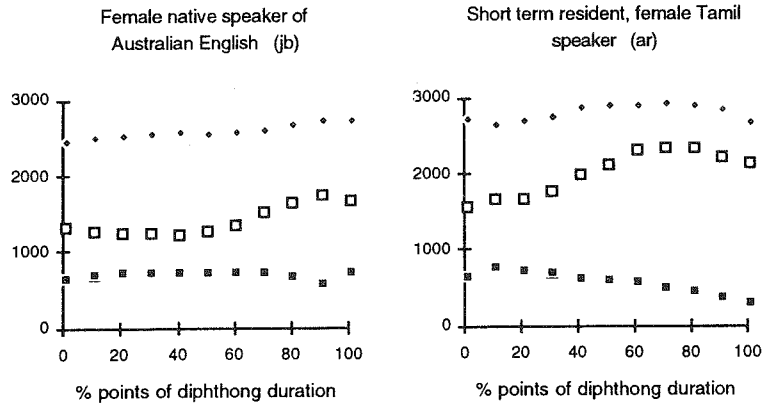

Figure 2. F1, F2 and F3 formant patterns (in Hz) for the diphthong nucleus of the word "buy"

The temporal variation of formant patterns for the diphthongs in the words "chair", "now" and "tour" are similar for both groups. In the case of the vowel /ʊə/ in the word "tour", there is no clear indication of a transition between one vowel to another for both groups of speakers.

Differences between the two groups

The variation of the slope of the F2 pattern with the duration of the diphthong in the word "here" is found to be different for the two groups, as can be seen from representative informants graphed in figure 3. In the case of native speakers of Australian English, F2 is relatively constant up to the 50% point of diphthong duration and then decreases slowly. On the other hand, for the Tamil speakers there is a continuous decrease in F2 values over the duration of the diphthong.

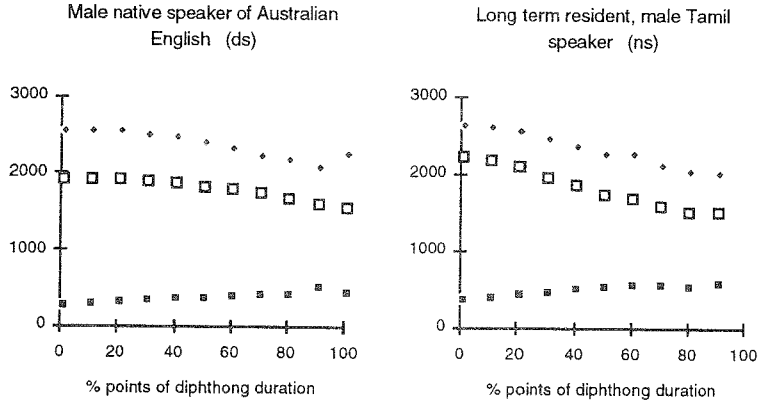

Figure 3. F1, F2 and F3 formant patterns (in Hz) for the diphthong nucleus of the word "here"

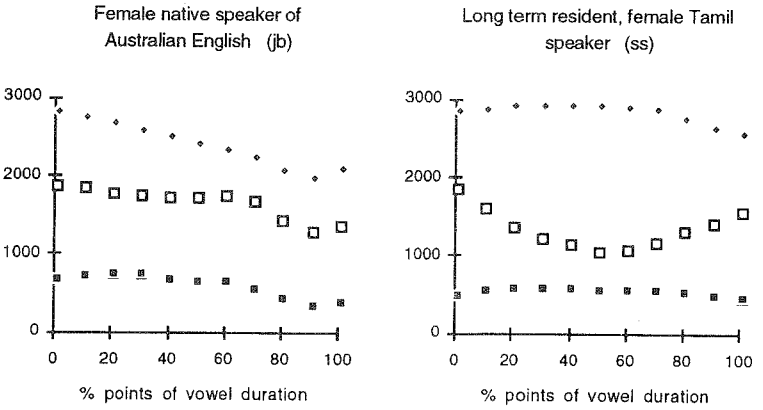

Figure 4. F1, F2 and F3 formant patterns (in Hz) for the vowel nucleus of the word "no"

An interesting difference is found in the formant patterns for the vowel in the word "no". This difference is shown in figure 4. It seems to be a characteristic of all the Tamil speakers that in the word "no" the vowel is not diphthongised. This could be a result of the influence of the vowel /ɔ/ in Tamil.

Another interesting difference is found in the formant patterns of the diphthong /eɪ/ in the word "say". This is shown for four speakers in figure 5. One graph shows a sample from the group of native speakers of Australian English and the other three are from the group of Tamil speakers. The formant patterns are very similar between the male native speaker of Australian English and the male Tamil speaker who is a long term resident of Australia. However, for the female Tamil speaker, who is also a long term resident of Australia, the pattern is clearly distinct. This pattern is also observed for the male Tamil speaker who is a short term resident of Australia. Figure 5 indicates that for both the long term resident female and the short term resident male Tamil speakers, the diphthong /eɪ/ in the word "say" is characterised by a transition from a front towards a central vowel.

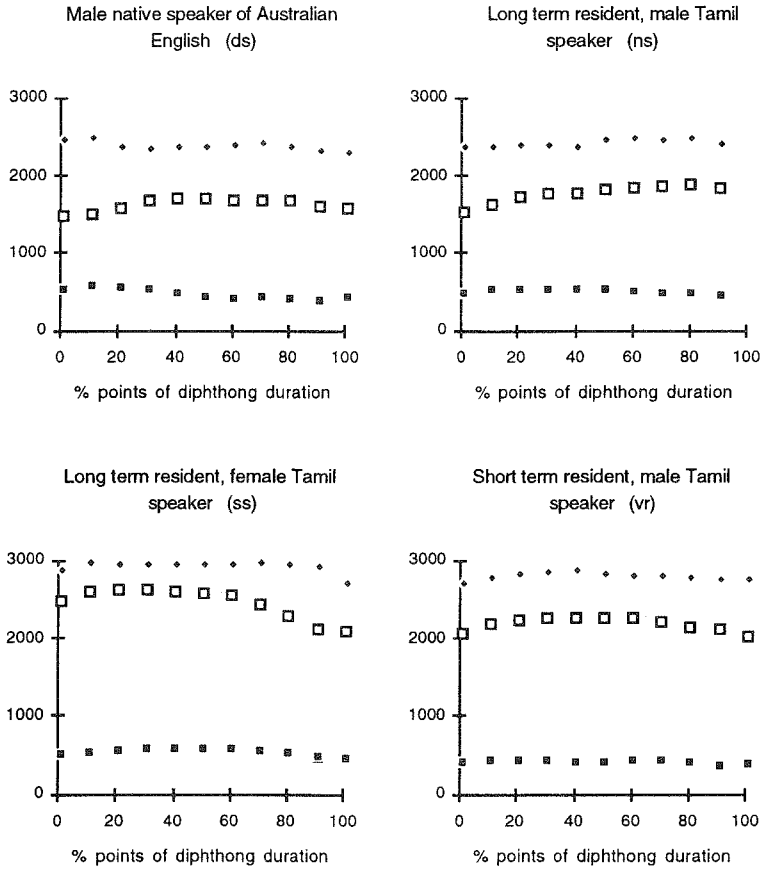

Figure 5. F1, F2 and F3 formant patterns (in Hz) for the diphthong nucleus of the word "say"

The expected diphthongal transition does not occur for the diphthong in the word "say" for the other two Tamil speakers. The downward slope in F2 pattern at the end of the diphthong could be due to the influence of the next word "that" which follows the word "say" in the chosen sentence. The formant values shown in figure 5 for the long term resident female and the short term resident male speakers of Tamil are similar to the monophthong vowel /ɛ/. The long term resident male Tamil speaker appears to be influenced by Australian English and this is reflected in the formant patterns of other

diphthongs, which are very similar to those of native speakers of Australian English except for the word "no". The transition in the diphthong /aɪ/in the word "buy", for example, for this informant is at the 50% point, which is similar to where the transition occurs for native speakers of Australian English. The variation in the slope of the F2 pattern for the diphthong /ɪə/ in the word "here" is very gentle.

FURTHER WORK

Although our current research on the comparison of the diphthongs produced by the two groups considered in this paper is only at a preliminary stage, it is possible to say that there does seem to be some influence from their native language on the English spoken by speakers of Tamil, particularly for those who have resided in Australia for a relatively short period of time. From the discussion above, it appears that some English diphthongs which Tamil does not have are rendered as monophthongs by native speakers of Tamil. It would be interesting to see how widespread this phenomenon is amongst Tamil speakers of English, and it is intended to analyse diphthongs spoken by other native speakers of Tamil.

In the experiment reported here, the words are extracted from sentences and hence there is the possibility of some influence from the initial consonant on the diphthong, as well as effects from succeeding words. For this reason, it is intended that further study be done to examine the same words uttered in isolation to compare the results of such words uttered in isolation with the same words uttered in continuous sentences. A controlled experiment is also planned for more quantitative modelling of the production of diphthongs in these two dialects. Also, more data have been collected for other Tamil non-native speakers of English in order to develop a quantitative model that will explain the differences in these two dialects of English. Such a model may then be expected to advance our knowledge about how a speech recognition system needs to be modified to recognise these two dialects of English.

REFERENCES

Bernard, J. R. (1989), *Quantitative aspects of the sounds of Australian English*, edited by Peter Collins and David Blair, p. 187.

Clermont, F (1988), "A dual exponential model for formant trajectories of diphthongs", Proc. Second Australian International Conference on Speech Science and Technology, pp.146-151

Clermont, F (1991), *Formant-contour models of diphthongs: a study in acoustic phonetics and computer modelling of speech*, Ph.D. Thesis, The Australian National University.

Gottfried, M (1989), "Some acoustical properties of diphthongs", J. Acoust. Soc. Am., Suppl. 1, Vol. 86, Fall, 123.

Sampath, S., Slater, D., Oasa, H & Wagner, M (1991), "A preliminary analysis of English spoken by Tamil speakers", Paper presented at the 8th Annual Conference on Language and Speech, Melbourne.

ACKNOWLEDGMENTS

The authors would like to thank Dr Frantz Clermont for invaluable discussions, and Alison Hume and Peter West for the data collection and segmentation of speech material.