ACOUSTIC CORRELATES OF ACCENT CHANGE: VIETNAMESE SCHOOLCHILDREN ACQUIRING AUSTRALIAN ENGLISH

J. Ingram and J. Pittam

Department of English
University of Queensland

ABSTRACT: Changes in vowel formant trajectories in four vocalic nuclei for 10 Vietnamese immigrant schoolchildren acquiring Australian English are reported. Changes are in conformity with predictions based on a comparative phonetic/phonological analysis of the two languages.

INTRODUCTION

This study investigates the auditory and acoustic correlates of accent change in a group of Vietnamese immigrant schoolchildren in the process of acquiring fluency in Australian English. Phonetic and phonological accommodation to L2 will involve progressive loss of a 'foreign accent', a process which will be accomplished at various rates and with various degrees of completeness by different individuals.

Australian English and Vietnamese have many potentially important differences of sound pattern. Both have a comparatively rich system of vowel contrasts, offering numerous points of close correspondence. However certain vocalic nuclei which are highly distinctive of Australian English (AE) provide potentially salient sources of 'foreign accent' colouring. The vocalic nuclei /u/, /ai/, /ei/, and /i/ were selected for this reason, embedded in the carrier words 'two', 'five', 'eight', and 'teeth'.

When considering the likely mapping between AE vocalic nuclei and those of Vietnamese (V), each vowel must be considered separately. Vietnamese has two high back vowels, rounded /u/ and unrounded /ɯ/ (Thompson, 1965). Neither corresponds very closely to the phonetic target for AE /u/, though impressionistically V /u/ seems closer than /ɯ/. Assuming that the first strategy of Vietnamese speakers would be to substitute the closest L1 equivalent and then progressively accommodate to the L2 target, the predicted accent change would be towards a more centralized /u/.

In Vietnamese phonology, there is no case for a phonemic /ai/ diphthong, although /aj/ constitutes a common syllable terminus (Thompson, 1965). An acceptable phonetic approximation of the AE /ai/ diphthong may be achieved by a low central-back Vietnamese vowel followed by a /j/ offglide. But Vietnamese phonotactic constraints preclude an obstruent following an /aj/ off-glide. It might therefore be predicted that the diphthongal character of AE /ai/ will be preserved and even enhanced by Vietnamese speakers' insertion of a /j/ glide, but at the cost of deleting any final obstruent in the word such as 'five' (AE /faiv/ -> V /faj/), particularly if the obstruent in question does not belong to the very limited set which may occur in coda position in Vietnamese. The onset of the /ai/ diphthong might be expected to undergo some backing with accommodation to (broader forms of) AE.

Vietnamese /e/ is monophthongal is closed syllables and allophonically lowered and centralized before apical stops (Thompson, 1965; Nguyen, 1970). As in the case of AE /ai/, Vietnamese speakers could phoneticially approximate its diphthongal character by adding a /j/ off-glide. For a word like 'eight', the predicted Vietnamese phonetic realization and the direction of accent change may depend upon how the item is lexicalized. Two possibilities suggest themselves: V /et/, or V /ej/. In the case of the former, the vowel will likely be retracted, and shortened by the closed syllable environment. In the case of the latter, the vocalic nucleus will likely be lengthened and diphthongized.

The consonantal features relevant for Vietnamese accented English in the present study involved word initial /t/, and the final obstruents in the items 'five', 'eight', and 'teeth'. Vietnamese has two contrastive dental stops in initial position. One is described as aspirated and lenis, and the other as unaspirated and tense with a sharp release gesture (Thompson, 1965). Nguyen (1970) reports lack of aspiration as the most common error on initial /t/.

The only obstruents which may appear in syllable final position in Vietnamese are /p/, /t/, and /k/, which are unreleased. Consequently, the production of final fricatives in English is usually much reduced, omitted, or replaced by an unreleased stop. Also, closed syllables are often accompanied by a prominent 'checked' quality, which may derive from a Vietnamese allophonic rule of vowel shortening (Nguyen, 1970), or from glottal tension associated with the high rising tone that accompanies Vietnamese syllables which are closed by a final stop. These coda characteristics appear to be transferred to closed syllables in English.

THE SPEAKERS

Ten Vietnamese immigrant schoolchildren ranging in age from 6 to 12 years and of 1 to 22 months residence in Australia at first testing were interviewed at school in July 1984 and again in May 1985. The subjects' speech was assessed with the Edinburgh Articulation Test (EAT) (Anthony, 1971), a naming task consisting of 44 pictures of common objects generally within the vocabulary of the preschool child. Subjects were also asked to recite the days of the week and numerals one to ten.

Ten Australian-born schoolchildren matched in age to the Vietnamese, were used as a reference group against which to evaluate anticipated vowel formant changes on the part of the Vietnamese speakers. The Australian speakers were drawn from the data base of a larger study of speech variation in Brisbane schoolchildren (Ingram, Pittam, & Newman, 1985), in which the same speech assessment had been used. These subjects, however, were recorded on one occasion only.

ACOUSTIC MEASURES

Broad band spectrograms (using a Kay 6061B sound spectrograph) were obtained for the four items elicited from the Vietnamese speakers at each of the two interviews and once from the Australian reference group. Because of the relatively high fundamental frequencies of the young speakers, half speed recording was used to obtain more densely packed harmonics in relation to the analysing filter bandwidth. A scale expansion of 50% (0-8KHz) was used to print the spectrograms.

The apparent centre frequencies of the first two formant bands were estimated by eye and measurements were taken from the broad band displays using a frequency scale calibrated in 100 Hz steps, which is approximately the optimal accuracy achievable for children's voices by spectrographic means (Monsen & Engebretson, 1983). Measurements were made by JI and JP, both of whom have considerable experience in spectrographic analysis, and with the assistance of a trained research assistant. Independent measurements by JI and JP for the Australian reference group showed that the estimated formant frequencies rarely differed by more than a single scale value (100 Hz).

Five measurement points were taken to represent the F1-F2 trajectories of each vowel or diphthong. The first and fifth measurements were taken at vowel onset and offset, and the other three of equidisant intervals in between. Formant frequency values originally expressed in Hz were converted into Bark units using the simple approximation attributed to Traunmuller (1983):

$$B = \frac{26.81f}{1.96+f} - 0.53$$

with frequency f in kHz.

The Bark transformation was used to provide a perceptually relevant scale for assessing any observed changes in formant trajectories which were represented in an F1-F2 graph.

An index of F1-F2 trajectory change reflecting movement towards or away from the AE reference trajectory was required to evaluate vowel changes that might be observed in the Vietnamese speakers over the two testing occasions. We will refer to such an index as the formant change score (FCS). The 'closeness' of any two formant trajectories can be expressed by the vector sum of their corresponding interpoint distances in the F1-F2 space. To compute interpoint distances, the following Euclidean distance metric was used:

$$D_{ij} = \sqrt{\sum_{1}^{n} (F_{ik} - F_{jk})^2}$$

where, D_{ij} = distance in Bark units between two points (i,j) in F1-F2 space.
F_{ik} = Bark value of Formant k for point i.

VOWEL FORMANT CHANGES

Changes in Bark-scaled F1-F2 trajectories are shown in Figure 1, which includes the Australian reference trajectories for comparison. Substantial movement is apparent for two vocalic nuclei: /u/ and /ai/.

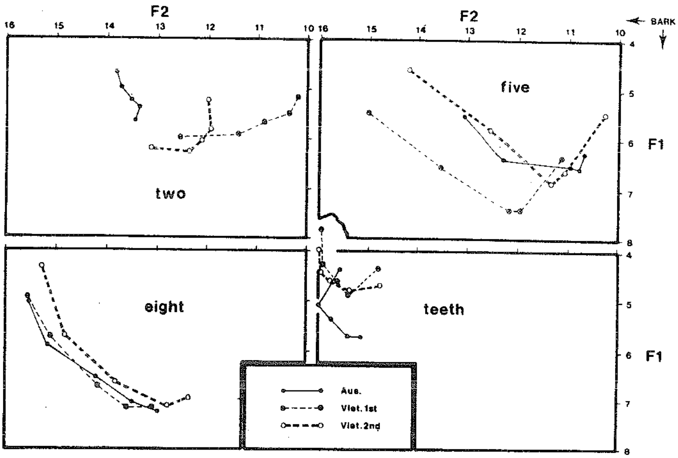

Figure 1. Changes in vowel formant trajectories by Vietnamese speakers in relation to AE targets.

The target position for /u/ in AE is notably more centralized than in most English dialects. The Vietnamese speakers showed substantial, but not complete, movement towards the Australian target. The primary acoustic change involves a raising of the second formant, which could be due to a more fronted tongue-body position, some loss of lip rounding, or a combination of these features, both of which are characteristic of AE /u/.

DISCUSSION

The centralization of /u/ was the most prominent of the vocalic changes. It is also one of the salient features distinguishing vowel sounds of Australian English from other English dialects (one exception being Cockney speech). Although stronger in Broad AE, all three varieties have a more centralized /u/ than has been reported for American or standard British English (Bernard, 1970). This feature has apparently exerted a strong influence upon the Vietnamese childrens' vowel target for /u/, shifting it substantially in the direction of [ʉ].

It was anticipated that in their pronunciation of the AE /ai/ diphthong in 'five', the Vietnamese speakers would show a shift on the follow-up recording towards a more retracted tongue body position for the initial target and less diphthongal movement overall. These expectations were only partly fulfilled. Some retraction was evident, but the degree of diphthongization remained unchanged. The results suggest that the children may be continuing to analyse the AE diphthong as a /j/ off-glide, but are modifying their pronunciation of the initial vowel towards the more retracted AE target [ɒɪ].

Two vocalic nuclei did not show significant accent change: AE /ei/and /i/. The reference group target was, as anticipated, slightly more centralized for /i/ and remained so over the two testing occasions. It is possible that the dialect differences for this vowel are too slight to have a perceptual impact upon the Vietnamese speakers.

The apparent lack of change for the /ei/ diphthong was, however, unanticipated. As mentioned earlier, there are two plausible phonemicizations that Vietnamese speakers may make in their initial perception of the AE word 'eight', one primarily influenced by the diphthongal character of the AE vowel nucleus, the other, by the final alveolar stop. As also noted earlier, a combination of these vocalic and consonantal features is incompatible with the phonotactics of Vietnamese. Hence, speakers at an initial stage of second language learning, where L2 words are simply incorporated into the phonological structures of L1, will be obliged to choose one or the other phonological representation for the 'foreign' word. The initial choice of phonological representation for the item in question will have consequences for a host of interdependent features of its phonetic realization and for its subsequent phonetic change towards a more standard L2 pronunciation. The item 'eight' is the only one of the four for which competing Vietnamese phonemicizations are likely. The proposed phonological constructions and their corresponding phonetic realizations for the four AE test items by the Vietnamese speakers are:

Australian English target		Vietnamese approximation	
Orthographic form	Phonemic form	Phonemic form	Phonetic form
two	/tu/	/tu/	[t̪ɯ]
five	/faiv/	/faj/	[faj]
teeth	/ti /	/tit/	[t̪iˀt̚]
eight	/eit/	/ey/	[ej]
		/et/	[eˀt̚]

The availability of competing L1 phonemicizations for the L2 form is suggested as the reason why no consistent trend in formant trajectory change was observed for the item 'eight'.

CONCLUSIONS

In a further study that exigencies of space prevent us from reporting here (but see Ingram & Pittam, forthcoming), it was shown that the formant trajectory changes were more important for the prediction of perceived accent change than were the consonantal features mentioned above.

A comparison of the performance of speakers over a wider age range would have bearing on the long-standing question of developmental factors in second language phonological accommodation. Extending the age range to include both children and adults would, of course, raise the problem of an appropriate vocal tract normalization for the FCS's. This in turn raises an interesting possibility that currently proposed, competing procedures for vowel normalization (Syrdal, 1985) could be evaluated using accent change data of the kind reported here, where it is possible to separate acoustic variability due to inherent vocal tract features from that which is attributable to phonetic change.

Vowel quality shifts due to L2 accommodation seem to be highly perceptible when they occur within a single vocal tract. Perceptual judgements of the degree of vowel shift from a sample of second language learners, where there is sufficient variation in vocal tract size and in the degree of phonetic accommodation to L2, could provide a useful data base for evaluating competing vowel normalization schemes. Hindle (1978) provided a similar test of three vowel normalization procedures by assessing their ability to reduce inherent speaker variance in the recovery of sociolinguistically conditioned phonetic variation in vowel quality. However, accent change measures offer the double advantage of being able to obtain finer perceptual measures of vowel quality differences unconfounded by inherent vocal tract variation, and being able to separate this source of variation from that which is purely phonetic and which is of primary interest to linguistics and sociolinguists.

REFERENCES

ANTHONY, A. (1971) "The Edinburgh Articulation test: a textbook", (Edinburgh: Livingstone).

BERNARD, J.R. (1970) "Toward the Acoustic Specification of Australian English", Zeit. fur Phon. Sprach. und Komm. 23, 113-128.

HINDLE, D. (1978) "Approaches to Vowel Normalization in the Study of Natural Speech", In D. Sankoff (Ed.) "Linguistic Variation: Models and Methods", (New York: Academic Press).

INGRAM, J. & PITTAM, J. (forthcoming) "Auditory and Acoustic Correlates of Perceived Accent Change: Vietnamese Schoolchildren Acquiring Australian English".

INGRAM, J., PITTAM, J. & NEWMAN, D. (1985) "Developmental and Sociolinguistic Variation in the Speech of Brisbane Schoolchildren", Aus. J. Ling. 5, 233-246.

MONSEN, R.B. & ENGEBRETSON, A.M. (1983) "The Accuracy of Formant Frequency Measurements: a Comparison of Spectrographic Measurements and LPC", J. Speech and Hear. Res. 26, 89-97.

NGUYEN, D.L. (1970) "A Contrastive Phonological Analysis of English and Vietnamese", Pacific Linguistics: Series c. 8 (Canberra: ANU).

SYRDAL, A.K. (1985) "Aspects of a model of the auditory representation of American English vowels", Speech Communication 4, 121-135.

THOMPSON, L.C. (1965) "A Vietnamese grammar", (University of Washington Press: Seattle).

TRAUNMULLER, H. (1983) "Analytical Expressions for the Tonotopical Sensory Scale", Unpublished PhD Thesis, Stockholm University.