Curing the Goat's Mouth

Laurie Bauer & Paul Warren

School of Linguistics and Applied Language Studies
Victoria University of Wellington University
New Zealand
laurie.bauer/paul.warren@vuw.ac.nz

Abstract

Prompted by an overlap in our transcriptions of the MOUTH and GOAT vowels in New Zealand English (NZE), we conducted an acoustic analysis of these diphthongs in samples of NZE for three age groups and two speaker sex groups. In addition, we considered the realization of the CURE vowel, since the start and end points of this diphthong appeared to overlap with the end and start points of MOUTH and GOAT. We conclude that although there is overlap in MOUTH and GOAT, our initial 'confusion' of these vowels in transcription reflects our reliance on monophthongal phoneme target values for start and end points of the diphthongs, and that a closer phonetic analysis allows better discrimination of the vowels, which should be reflected in our choice of transcription symbols. We conclude also that while the endpoint of the CURE vowel does indeed overlap with the start point of MOUTH and GOAT, the start point of CURE lies between the endpoints of the other two.

1. Introduction

In this paper we investigate the realisations of three diphthongs in New Zealand English, which we will refer to as the MOUTH, GOAT and CURE vowels, using Wells' (1982) labels for lexical sets. In a recent study of New Zealand English phonology (Bauer & Warren, 2004), we offer [ɐʉ] and [ɐɨ] as transcriptions of GOAT. At the same time, while we offer [æʉ] and [ɛʉ] as transcriptions for MOUTH, the possibility arises of a more retracted starting point, giving [ɐʉ] as a possible transcription for MOUTH as well as for GOAT. A related issue concerns the acoustic structure of the CURE vowel. When listening to recordings for Bauer and Warren (2004), we were struck by the openness of the end-point of this diphthong in comparison with that of NEAR/SQUARE (so likely nearer [ɐ] than [ə]). This seems to imply that CURE could be a reverse version of either GOAT or MOUTH (so [ʉɐ]).

The main issue for this paper is that in our initial analysis of recordings for Bauer and Warren (2004) we arrived at impressionistic transcriptions that seemed to support an overlap of the MOUTH and GOAT diphthongs. However, since there is no discussion of a merger of these vowels either in the linguistic literature or in comments to the press (contrast this with the situation for NEAR and SQUARE) it seems extremely unlikely that the merger is particularly widespread. Nevertheless, our initial observation suggested that further investigation was warranted.

In apparent contradiction of our auditory impressions of these tokens, our preliminary acoustic analysis of the same recordings shows in fact that there is no evidence for a merger of MOUTH and GOAT (see Figure 1). This suggests that in our auditory analysis of these recordings we were being misled by something (or are simply just not cut out to be phoneticians). Note that the figure does indeed show similar starting points – the small differences between them may be due to preceding phonetic context: [m] for MOUTH, [g] for GOAT). The end points are however quite different, particularly for F2, which shows that MOUTH has a closing point further back than GOAT.

2. Further data from NZSED

To provide further data for the MOUTH and GOAT vowels, as well as for CURE, we analysed a larger set of recordings from the New Zealand Spoken English Database (NZSED: Warren, 2002). The complete data

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

Accepted after abstract only review

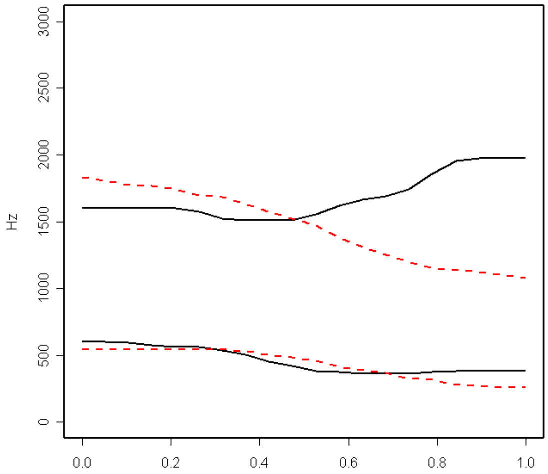

Figure 1: formant tracks for GOAT (solid lines) and MOUTH (dotted lines) as recorded for Bauer and Warren (2004).

set for NZSED is based on the ANDOSL database for Australian English (Millar, Dermody, Harrington, & Vonwiller, 1990), and includes tokens of vowels read mostly in a hVd context. The exceptions to this context happen to include the diphthongs of interest here – the MOUTH and GOAT vowels have no final [d] in their context, being in the carrier words *how* and *hoe*, and the CURE vowel is in the word *tour*. Note that although the latter may of course not have the CURE vowel at all, it was regularly pronounced with CURE in the NZSED lists.

NZSED is a project in progress, and speech data are most completely available for informants from the Wellington region, recorded over the period 2000-2001. More recently, recording sessions have been started in Hamilton in the Waikato region. Although it is commonly acknowledged that there is little regional variation in NZE, recordings from Hamilton may be of interest as a comparison set to the Wellington region for a number of reasons – the Waikato region has a more obvious Maori presence than Wellington and NZE spoken in that region may therefore be more clearly influenced by Maori English (for a description of this variety see Warren & Bauer, 2004); Hamilton is a smaller centre than Wellington, and may therefore experience changes in NZE in different ways or at different times from the capital; research on playground vocabulary (Bauer & Bauer, 2000, 2003) indicates that for lexical items in use at least in this restricted domain, the Waikato region is in a different dialect area from Wellington. However, it turns out that for the vowels in question (though not quite so clearly for some of the monophthongs), there is little discernible difference between the Wellington and Hamilton data, and so in our discussion we group the speakers from each region together, while maintaining distinct groupings for speaker sex and age.

The NZSED data recorded for Wellington speakers contains 6 speakers in each cell defined by the combination of speaker sex and three age ranges (18-30, 31-45, 46-60). For Hamilton, at time of writing we have available data from 5 young, 7 mid-age and 6 old female speakers, and 6 young, 5 mid-age and 3 old male speakers. All speakers considered here self-identified as Pakeha (New Zealanders of European descent), and data from recordings made of Maori speakers of English from the Wellington region are excluded from consideration. (The sex × age range sampling is not yet complete for that group. Recordings of Maori informants in Hamilton will be starting shortly.)

The data presented below also include trajectories for the NEAR and SQUARE vowels (using the word list items *hear* and *hair*), to serve as reference points chiefly for the endpoint of CURE. We also include in our figures ellipses of the STRUT, TRAP, GOOSE, FOOT and FORCE vowels.

For each token from the NZSED word list, the digitised recording was used to generate formant tracks in *ESPS/xwaves*™. *Emu* (http://emu.sourceforge.net/) was then used to view speech files, spectrograms and formant tracks, and to produce phoneme annotations for the speech files, using the following criteria. For segmentation between /h/ and the vowel, we chose a point where the frication of the /h/ (which was invariably voiceless) had clearly given way to voicing with formant information for the vowel. The end of the vowel was marked as the point at which formant information was no longer clearly identifiable, giving way either to silence or to vocal murmur during the stop closure for the final /d/. In the case of *tour*, the /t/-vowel transition was taken as the point at which aspiration of the /t/-release had given way to voicing with formant information. For open syllables (*how, hoe, cure, tour, near, square*) the vowel offset was taken as the point at which the vowel had died away to the extent that formant information was no longer reliably present.

The *Emu* library in the *R* statistical programming package (http://cran.r-project.org/) was then used to scan the phoneme annotation files for occurrences of each vowel of interest and to extract formant data for these tokens. The formant data were analysed to produce average tracks (for diphthongs) or target values (monophthongs) for the examples in each sex × age grouping. The formant tracks used for the diphthongs were from the beginning to the end of the vowel, marked as indicated above. Each track was smoothed to remove minor perturbations in the formant data, using a running mean smoother. Twenty temporally equally spaced measures were then taken from each smoothed F1 and F2 track and used to compute average trajectories for each diphthong. For

Accepted after abstract only review

the monophthongs in each set, average F1 and F2 values were taken over the vowel stretch from 0.25 to 0.50 of the total duration of each vowel. An analysis starting point of 0.25 of the way into the vowel allowed us to avoid any interference in the formant track from the initial /h/ or arising from segmentation error during labelling. The monophthongal vowels were in all cases followed by /d/, and an analysis end-point at 0.50 of the way into the vowel ensured that formant tracks were not affect by movement of the tongue towards the alveolar ridge. The average F1 and F2 values for each vowel were then used to determine average and standard deviation values in each sex × age set, leading to ellipsoid plots of the vowel's distribution.

3. Results

As noted above, comparisons of age- and sex-matched subgroups of Wellington and Hamilton data show similar patterns for the diphthongs of interest for the two regions, so the data to be presented will put the two regional groups together.

The panels in Figure 2 show formant data from female participants on the left, and from male participants on the right, with the oldest speakers in the top panels and the youngest speakers in the bottom panels. Three of the diphthong trajectories are for centering diphthongs (NEAR, SQUARE, CURE). The two trajectories on the left of each panel are for the NEAR and SQUARE vowels, which are clearly very close to one another, particularly for the younger speakers and particularly for the females, as has previously been shown for NZE (e.g. Gordon & Maclagan, 2001). The trajectory with a clear elbow is the formant track for the CURE vowel. The trajectory before the elbow indicates the movement away from the initial /t/ consonant in the word used to elicit this vowel (*tour*). The main diphthongal movement of interest is after the elbow. Contrary to our initial observation, the end-point of this diphthong does not appear to be markedly more open than that of NEAR or SQUARE in the wider dataset. One interesting unexpected finding for the CURE vowel is that the starting point of the main diphthongal movement appears to become progressively further back for younger speakers (compare in particular the old females with the mid-age and young females, and the old and mid-age males with the young males).

Figure 2 also shows ellipse distributions of STRUT, TRAP, GOOSE, FOOT and FORCE (in order, following the ellipses clockwise from the lowest ellipse). We have little to say about differences in these distributions for the different speaker groups, although there are some interesting patterns of change that we will be exploring elsewhere. These distributions are shown in order to provide reference points within the vowel system for the diphthongs of interest.

Note that the starting points for both GOAT and MOUTH fall within or fairly close to distribution for STRUT. Note also, however, that these two starting points correspond effectively to different realisations of STRUT. That is, they are not phonetically identical, and the GOAT starting point is both closer and backer than that of MOUTH. More dramatic, though, is the difference in the glide targets for these two diphthongs. The target for GOAT appears quite clearly to be in the neighbourhood of the GOOSE vowel, while that for MOUTH is closer to the FOOT vowel. This latter observation is confused somewhat by the apparent forwards shift in the FOOT vowel for the younger speakers (noticeable also for the mid-age females), so that the end-point of MOUTH becomes closer to FORCE than to FOOT, most notably for the young females. It is also evident that the extent of the diphthongal movement for MOUTH is much greater for the younger speakers.

As far as the data we have used are concerned, we have a fairly robust finding for a suggested transcription for MOUTH and GOAT in NZE. GOAT should be transcribed as [ʌʉ] or possibly as [aʉ] in that the start point is frequently slightly retracted from core exemplars of the STRUT vowel. MOUTH should be [ʌo]. Even where [o] is not reached, the trajectory appears to be towards [o]. However, we note that broader versions of this vowel have a much more centralised (and sometimes unrounded) target, for which [ʊ] or even [ə] might be better. Where CURE is concerned our data give regular evidence in favour of a transcription [ʊʌ]. However, this was in the word *tour*, which is unusual in NZE in having CURE where it is not preceded by a /j/. Native speaker intuition and general New Zealand practice in the universities seems to be in favour of equating the starting point of this vowel with GOOSE (which would lead to a [ʉʌ] transcription, or, if we wanted to be provocative, a [yʌ] transcription). This should probably be explained in terms of a palatalisation-effect from the preceding palatal in a variety where GOOSE is regularly a front vowel.

Accepted after abstract only review

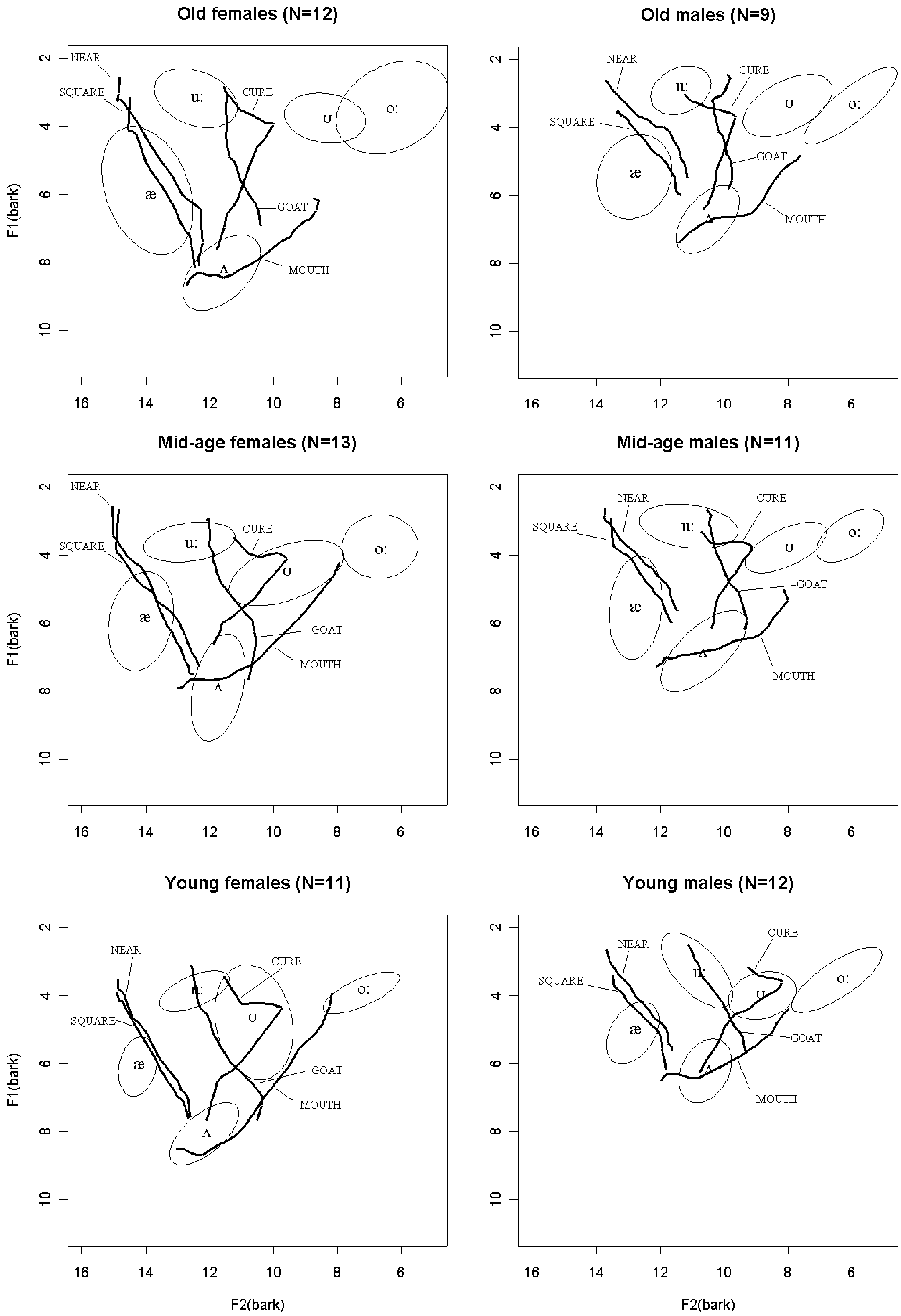

Figure 2: formant tracks for diphthongs (SQUARE, NEAR, CURE, GOAT, MOUTH) and ellipses for monophthongs (STRUT, TRAP, GOOSE, FOOT, FORCE) for selected vowels from the word-list recordings for NZSED (see text for details).

Accepted after abstract only review

4. Discussion

Given the extent of the differences between the MOUTH and GOAT diphthongs, how was it then that we (in preparation of Bauer & Warren, 2004) felt that there was potential overlap between these two vowels? One important issue is that the transcription system that we were working towards described the diphthongs in terms of the expected phonemic structure of the monophthong system. That is, if we assume the starting points of MOUTH and GOAT belong to one of the monophthongs then a transcription with identical (STRUT) starting points would be accurate, but would miss the additional detail of where in the distributional area for this vowel the diphthong actually starts. So there were two factors working against us here: the variability of the realizations of a particular phoneme, and our predisposition to hear sounds in phonemic categories.

When it comes to end points, the same is not quite true. In English diphthongs, it is important to know which direction the tongue is moving, not to know how far it moves, so that [ai], [ae] and [aɛ] will all be perceived as variants of the PRICE vowel, for example. Similarly, with the MOUTH vowel it is sufficient for identification of the vowel to establish a closing movement towards something not front and spread, and the endpoint is not crucial. All of the possible phonemes we could identify as end points (if we wish to identify the end points of diphthongs with phonemic areas) are extremely variable (as is illustrated partly by the plots in Figure 2, but also by the fact that our own — British — pronunciations of these vowels add an extra dimension to this variability). Accordingly, even accuracy at what we might term a phonemic level, let alone at a phonetic level, is hard to achieve. A good phonetic training should be able to overcome these problems, of course, and indeed we believe it can and that it helped us here, just not enough.

If we consider this on a larger scale, we can say that there are in NZE two monophthongal short vowel phonemes whose quality overlaps with the starting point of the diphthong in the MOUTH vowel: the TRAP vowel and the STRUT vowel. Given that the START vowel and the STRUT vowel overlap in quality for many speakers of NZE (Easton & Bauer, 2000; Watson, Harrington, & Evans, 1998), we should add that to make a third possibility. If we look for monophthongal phonemes whose realizations overlap with the realizations of the approximate end point of this diphthong, we find at least GOOSE, FOOT and THOUGHT, and possibly KIT in its unstressed variants. If we say three possible notations for the starting point and three possible notations for the finishing points, there are nine possible notations of the MOUTH vowel any one of which would be a reasonable representation for the value produced by some speakers of NZE at the current time, and these exclude notations with diacritics or with vowel symbols which are not part of the monophthongal series of vowels for NZE. The possibilities for GOAT are probably slightly less numerous, but it is not all that surprising that some overlapping transcription should be available from all of this.

Labov et al. (1972) are worried about how vowels appear to be able to merge and subsequently de-merge. The examples considered there are historical ones where we have neither formant values nor detailed descriptions of the amount of variation there was in the vowel sound at the time. The lack of information on variation means that we tend to assume a relatively homogeneous pronunciation for the vowels concerned: but our own experience of variation in present-day English, along with the application of the Uniformitarian Principle (Labov, 1972: 275), show this to have been unlikely. Because the different vowels involved in the merger had different parameters of variation, the two would not necessarily have been confused just because some of the variants came close to each other. The fact that we have no acoustic description is important, because neither did the people describing the situation at the time. They were dependent on their own ears, just as we were before we moved on to acoustic analysis, and they quite possibly made similar errors to the type we made. With no phonetic theory of vowel description to guide those commentators, they must have been even more likely to have used phonemic categories than we were. Thus we can see that reports of the same starting point and same finishing point for a diphthong need not, in fact, mean that the two vowels were phonetically identical.

5. Acknowledgements

We are grateful for comments from participants at the 2004 Language and Society Conference of the Linguistic Society of New Zealand and from the anonymous reviewers of our SST-2004 abstract submission.

6. References

Bauer, L., & Bauer, W. (2000). Nova Zelandia est omnis divisa in partes tres. *New Zealand English Journal, 14*, 7-17.

Bauer, L., & Bauer, W. (2003). *Playground Talk*. Wellington: Linguistics and Applied Language Studies, Victoria University.

Bauer, L., & Warren, P. (2004). New Zealand English: phonology. In B. Kortmann, E. W. Schneider, K. Burridge, R. Mesthrie & C. Upton (Eds.), *A Handbook of Varieties of English: A Multimedia*

Accepted after abstract only review

Reference Tool (Vol. 1, pp. 580-602). Berlin: Mouton de Gruyter.

Easton, A., & Bauer, L. (2000). An acoustic study of the vowels of New Zealand English. *Australian Journal of Linguistics, 20*(2), 93-117.

Gordon, E., & Maclagan, M. (2001). Capturing a sound change: A real time study over 15 years of the NEAR/SQUARE diphthong merger in New Zealand English. *Australian Journal of Linguistics, 21*(2), 215-238.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia PA: University of Pennsylvania Press.

Labov, W., Yaeger, M., & Steiner, R. (1972). *A Quantitative Study of Sound Change in Progress*. Philadelphia: US Regional Survey.

Millar, J. B., Dermody, P., Harrington, J., & Vonwiller, J. (1990). *A national database of spoken language: concept, design, and implementation.* Paper presented at the International Conference on Spoken Language Processing (ICSLP-90), Kobe, Japan.

Warren, P. (2002). NZSED: building and using a speech database for New Zealand English. *New Zealand English Journal, 16*, 53-58.

Warren, P., & Bauer, L. (2004). Maori English: phonology. In B. Kortmann, E. W. Schneider, K. Burridge, R. Mesthrie & C. Upton (Eds.), *A Handbook of Varieties of English: A Multimedia Reference Tool* (Vol. 1, pp. 614-624). Berlin: Mouton de Gruyter.

Watson, C. I., Harrington, J., & Evans, Z. (1998). An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics, 18*(2), 185-207.

Wells, J. C. (1982). *Accents of English*. Cambridge, England: Cambridge University Press.

Accepted after abstract only review