DIVERGENCE OF FACE AND TRAP IN AUCKLAND ENGLISH: A POTENTIAL REGIONAL SOUND CHANGE IN NEW ZEALAND ENGLISH

Catherine I. Watson[1], Elaine Ballard[1], Brooke Ross[2], Helen Charters[1]

The University of Auckland, Victoria University of Wellington
c.watson@auckland.ac.nz, e.ballard@auckland.ac.nz, bros138@aucklanduni.ac.nz,h.charters@auckland.ac.nz

ABSTRACT

Earlier acoustic studies in New Zealand English suggest that the first target of diphthongs FACE and MOUTH may have raised in line with the monophthong TRAP. A recent study of the monophthongs in Auckland English indicates that TRAP is lowering for younger speakers. In the current study we investigate whether FACE and MOUTH have also lowered. We present the preliminary results from a formant analysis for over 2000 diphthongs from 40 Auckland-based speakers; 17 women and 16 men aged between 18 and 25 and 7 older women aged between 45 and 70. We show that whilst the first target of MOUTH has moved in the same direction as TRAP, this is not the case for FACE. The first target of FACE is both fronted and raised for younger speakers suggesting that FACE is no longer shifting in line with TRAP. We finish discussing the extent of this sound change.

Keywords: New Zealand English diphthongs sound change acoustic analysis.

1. INTRODUCTION

Acoustic investigations on New Zealand English (NZE) vowels show that they differ to their counterparts in Northern Hemisphere varieties of English. NZE monophthongs are characterized by raised TRAP, DRESS, THOUGHT, NURSE and LOT, the centralization of GOOSE and the lowered and centralized FOOT and KIT. The monophthongs START and STRUT are no longer qualitatively distinct and are distinguished solely by length [1-5].

While there has been considerable research on the NZE monophthongs [1-5], much less attention has been given to the diphthongs. In the latter case the focus has almost exclusively been on the merger of the falling diphthongs NEAR and SQUARE ([1-4, 6, 7]). The little research on the other NZE diphthongs shows that these diphthongs also differ from other varieties of English spoken in the Northern Hemisphere. Of note are the raised first target in monophthongs FACE and MOUTH [2, 8, 9]. The first target of these diphthongs is close to NZE TRAP and it has been tentatively suggested that these raised diphthongs have moved in line with the raised TRAP.

It has long been assumed that barring Southland, there is no regional variation in NZE [10]. However, a recent study into Auckland English has suggested that there are regional differences within the NZE monophthongs [11]. In that study findings show the short front vowels DRESS and TRAP in the speech of young Aucklanders are significantly lowered and differ to the front vowels of older speakers in the same city. Given the earlier finding suggesting a link between TRAP and diphthongs FACE and MOUTH, it seems appropriate to consider whether these diphthongs have also lowered in line with TRAP. In the following study we consider whether this is the case with an acoustic analysis of FACE and MOUTH as spoken in Auckland English.

2. METHOD

2.1 Speakers

As part of the Auckland Voices project [11,12], 33 speakers from three suburbs in Auckland (Mt Roskill (8 women, 6 men), Papatoetoe, (6 women, 7 men) Titirangi, (3 women, 3 men)) were recorded reading a 390 word passage. The participants were aged between 18 and 25 years. Our speakers were either New Zealand born or arrived in the country under the age of seven. A further 7 older female speakers from Titirangi (between 45-70yrs) were also recorded, these were all New Zealand born.

2.2 Data Preparation.

Speakers' recordings were digitised and transcribed in ELAN [13]. WebMAUS (NZE option) [14] was used for the automatic phonetic labelling, then phonetic boundaries and labels were checked and hand corrected where necessary in PRAAT [15]. PRAAT text grids were converted into the EMU-webApp [14, 16] format for formant calculation and analysis. Formant tracks were subsequently checked and hand corrected when necessary. Finally the first and second formants were extracted at the vowel targets marked according to the criteria in [2]. The remainder of the formant analysis was done in R [17] using EmuR [18]. Only vowels that carried phrase stress were studied in this analysis. Over 7500 monophthongs and 4000 diphthongs were analysed. In this study we focus on PRICE, FACE, MOUTH and GOAT, of which there were 2088 tokens. Table 1 gives the total number of diphthongs.

Table 1 Total number of diphthongs in the study

No. Diphthong Tokens by Suburb & Gender		
Suburb	Male	Female
Papatoetoe	367	315
Mount Roskill	272	418
Titirangi	153	154
Titirangi - Older		409

2.3. Data transformation

We combined the female and male data, using the data transformation process outlined in [11]. This involves using a simple translation, rotation and stretching of the formant plane of one gender to the other so that the three point vowels /i:. a:, ɔ:/ match. The transformation is done using

$$\begin{bmatrix} M_1^F \\ M_2^F \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \quad (1)$$

where (M_1, M_2) are the original male formant values, and (M_1^F, M_2^F) are the transformed male values converted to be typical for a female speaker, and μ_1, μ_2, A_{11}; A_{12}; A_{21}; A_{22} are the set of six numerical constants that enable the transformation.

Figure 1 shows the male and female monophthong data, untransformed, and transformed. In this case we transformed the male data to the female space, the female data remains unchanged. It can be seen that the centroids of the point vowels in Figure 1 (a) do not match (black for men, green for women), but in Figure 1 (b) they do. Six constants are generated in the monophthong space transformation process. In the diphthong analysis the first and second targets of the male speakers are transformed in the female space using (1) and the same set of numerical constants obtained from the monophthong analysis.

3. RESULTS

3.1 Diphthong mean Target values and Trajectories.

Diphthong trajectories for all diphthong data were acquired by extracting the tracks of the first and second formants (F1 and F2) between the first and second target times. The trajectories were then time normalized, and then averaged for each phonetic label. The averaged formant trajectories for the older female speakers from Titirangi and for the young female speakers from Papatoetoe are plotted on an F1/F2 space in Bark in Figures 2 and 3 respectively. The trajectories are superimposed over the relevant point vowel ellipses /i:. a:, ɔ:/ for each speaker group. The first target is marked with the Wells key word [19] associated with the diphthong. These plots are from [12] and are used with permission.

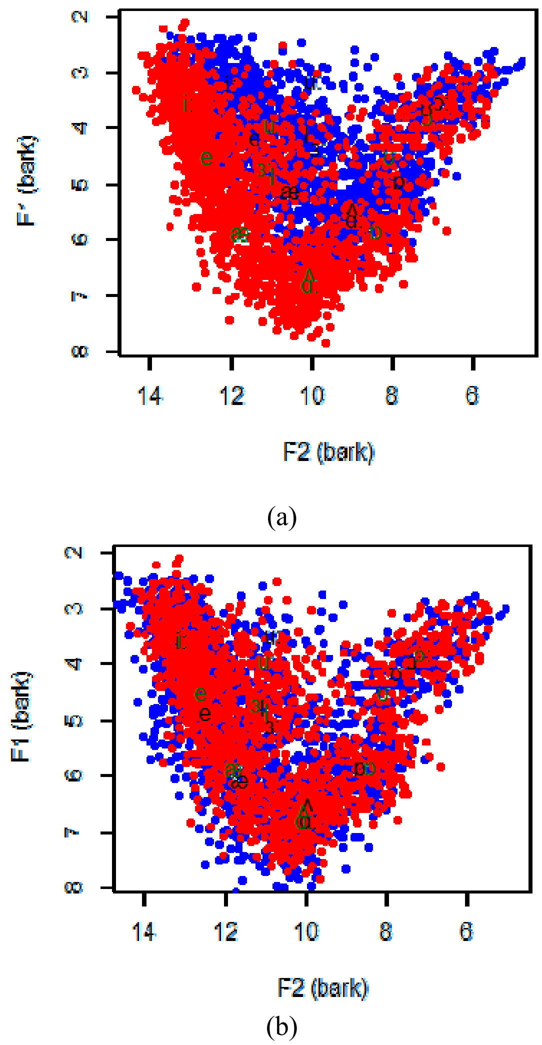

Figure 1: The (a) untransformed and (b) transformed formant data. The male data is transformed to the female space.

The diphthong trajectories from the older women in Titirangi (Figure 2) are characteristic of NZE, and are very similar to those found in [2]. The first and second targets are also very similar to those found for young bilingual Māori speakers of NZE [20]. The first target of PRICE is also lowered and retracted, which is standard for NZE. The first targets of FACE and MOUTH are close to each other and finally, the first target of GOAT is lowered as is common in Southern Hemisphere dialects.

The diphthong trajectories of the young female Papatoetoe speakers (Figure 3) differ from the older cohort. It can be seen that the first target of FACE is raised, as is the entire diphthong. The first target of

GOAT is also raised, and PRICE is fronted. The entire diphthong trajectory of MOUTH appears to have backed. These findings could be seen for all the young speakers, however due to space limitation we are only showing the diphthong trajectories for the young Papatoetoe females.

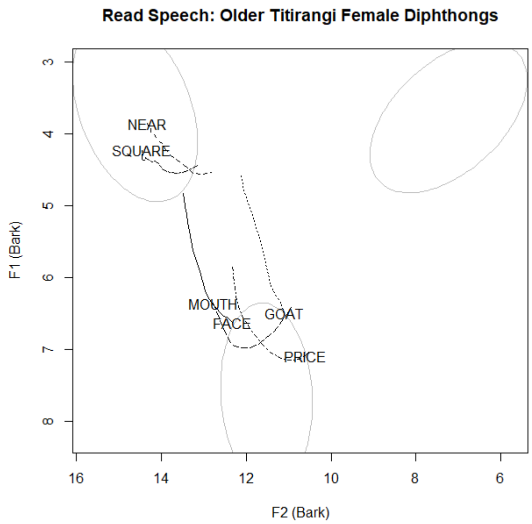

Figure 2: The time normalized average trajectory of diphthongs for older Titirangi female speakers. Formant tracks are extracted between the first and second target and superimposed over the relevant point vowel ellipses. Plot from [12], used with permission.

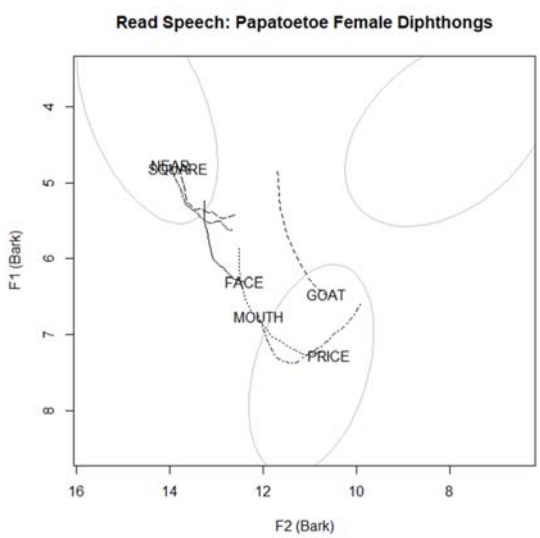

Figure 3: The time normalized average trajectory of diphthongs for the young Papatoetoe female speakers. Formant tracks are extracted between the first and second target, and superimposed over the relevant point vowel ellipses. Plot from [12], used with permission.

3.2 Statistical Analysis

To investigate whether the differences observed between the diphthongs of the young and older speakers are significant we performed statistical analysis on the formant values at the vowel targets. Table 2 gives the mean formant values given in Bark for the first and second targets for each diphthong for the four speaker groups. The mean is calculated from all the women speakers and the transformed men speaker. Similar observations to those made from the visual inspections of the trajectories can be made from these means: such as the raised first target of FACE and GOAT for the young speakers, the fronted first target of PRICE.

Table 2: Mean values of F1 and F2 given in Bark at the first and second target (T1 and T2) for PRICE, FACE, MOUTH, and GOAT for Mount Roskill (Mt R) Papatoetoe (Papa), Young Titirangi (Ti Y) and Old Titirangi (Ti O).

Vowel	Group	T1F1	T2F2	T2F1	T2TF
PRICE	Mt R.	6.77	9.99	5.36	11.31
	Papa.	6.77	9.76	5.42	11.15
	Ti. Y	6.54	9.67	5.39	11.08
	Ti O	6.42	9.42	5.28	10.87
FACE	Mt R.	5.34	11.68	4.49	12.24
	Papa.	5.68	11.32	4.72	12.03
	Ti. Y	5.48	11.41	4.61	11.85
	Ti O	6.02	10.87	4.42	11.86
MOUTH	Mt R.	6.32	11.16	6.19	9.29
	Papa.	6.40	10.80	6.24	9.09
	Ti. Y	5.84	10.96	5.83	9.06
	Ti O	5.76	11.30	5.80	9.69
GOAT	Mt R.	5.46	9.44	4.18	9.90
	Papa.	5.77	9.53	4.41	10.06
	Ti. Y	5.42	9.24	4.22	9.78
	Ti O	5.91	9.85	4.21	10.70

For the statistical analysis we modelled F1 and F2 values simultaneously in the 2-D formant space. This allows for increased statistical power in which changes in any direction in the plane can be detected. For each target and diphthong we built three linear mixed effects models for observations of F1 and F2. Each model had speaker as a random effect and the fixed effects investigated were type (F1, F2), sex (female. male) and group (Young Titirangi, Old Titirangi, Young Mount Roskill, Young Papatoetoe). All statistical analysis was performed in R [17], the linear mixed effect models were calculated using the `lme()` function in the `nlme` package [21] and the comparison between models was done using the `anova()` function implemented in the `nlme` package. The null model had type as the fixed effect.

It is still possible to find significant differences between the female and male speakers regardless of the transformation of the male formant data. However in this situation there was no significant difference, so the remainder of the study focuses on group effects. Although not ideal we conflated suburb and age effects, because we only had one group of older speakers. The statistical analysis found that there were significant differences between the null model and the model with fixed effects of type and group for the first target of FACE and MOUTH, and the second target of MOUTH, see Table 3. Despite the suggestions from the visual inspections, there were no significant findings for PRICE and GOAT.

Table 3: Significant differences when null model is compared to model with fixed effect of type and group.

Target	Vowel	Degrees of Freedom	AIC Difference	Log Likelihood Value	P Value
T1	FACE	2	9.7	-1243.370	p<0.01
	MOUTH	2	3.9	-895.1300	p<0.01
T2	MOUTH	2	2.4	-968.9283	p<0.05

Using a post-hoc t-test we found that for the first target of FACE the older speakers' F1 was significantly higher than each group of younger speakers ($p<0.05$) and the older speakers' F2 was significantly lower ($p<0.05$) than each group of young speakers. For the first target of MOUTH the older speaker's F1 was significantly lower than for the young speakers from Mount Roskill and Papatoetoe ($p<0.05$), and the older speaker's F2 was higher than for the young speakers from Mount Roskill and Papatoetoe ($p<0.05$). For the second target of MOUTH the F2 for older speakers was significantly higher than for the young speakers from Mount Roskill and Papatoetoe ($p<0.05$).The Titirangi young speakers followed the same trend as the young Mouth Roskill and Papatoetoe speakers for both T1 and T2, however their results were not significantly different to those for their elders.

4. DISCUSSION

Young Auckland speakers in our database seem to be rejecting some typical NZE diphthong realizations. Our findings show that for these speakers the first target of MOUTH has significantly lowered and backed in a similar fashion to TRAP. By contrast, the first target for FACE has fronted and raised for the same speakers. A correlation analysis shows that the movements of TRAP and the first target of MOUTH are significantly correlated ($r=0.342$, $p<0.05$), but there is no correlation with the movement of TRAP and the first target of FACE. It would then seem that for our young Auckland speakers FACE is no longer caught up in the short front vowel shift, and their FACE is considerably more raised than the young NZE speakers in [20].

For all young speakers FACE is raised to some extent and its trajectory is shortened. While these changes may no longer be in line with the front vowel system, there appears to be a link with other diphthongal shifts. Though not significant, we note a fronting of PRICE and raising of the first target of GOAT. For young speakers from Mount Roskill, and Papatoetoe MOUTH, is lowered and retracted compared with the older speakers. In contrast to FACE the movement of MOUTH is in line with TRAP.

What is motivating these changes in the diphthongs? If these changes are indeed part of a system, what might the process of change be? Evidence from Auckland Voices speakers, suggests diphthong shifting may have started with PRICE. Despite the movement of FACE being the most obvious change, PRICE is consistently fronted for all young speakers whereas this diphthong, by contrast, is always retracted for the older cohort. A retracted PRICE would be in line with findings from other studies on NZ English speakers [1-3].

It unclear what may cause PRICE to front. One possibility is the influence of other English dialects through social media and travel. Varieties such as General American English have both raised variants of the FACE diphthong and fronted variants of PRICE [19]. A more American colouring may reflect a more global outlook for young speakers similar to that suggested for young Capetonian speakers of South African English [22]. Or the move towards a more global English could be due to a conscious move away from commonly stigmatized pronunciations of NZE diphthongs such as MOUTH [10].

Another factor to consider would be the contact with newer English or L2 English varieties. Of note here is the fact that the young speakers who have spearheaded these diphthongal changes reside in ethnically diverse suburbs where a variety of English with fronted PRICE and a raised and often monophthongal variants for FACE is spoken (Indian English) [23]. A scenario where the desire to communicate between speakers of different varieties of English can lead to dialect contact and levelling has been found in other studies on English spoken in multi-diverse communities in urban centres. Similar diphthongal changes are reported in London English where the young non-Anglo speakers of English are having an impact on the variety of English spoken in the metropolis [24].

It is not clear at this stage whether just one or a combination of these factors has an impact on the changing face of Auckland English. At present, our results are preliminary and future planned acoustic and perceptual studies will deepen our understanding.

ACKNOWLEDGEMENTS

We thank the NZ Royal Society Marsden for funding Auckland Voices project, the participants, the research assistants, and Miriam. Meyerhoff.

REFERENCES

[1] Maclagan, M. (1982). An acoustic study of New Zealand vowels. *New Zealand Speech Therapists' Journal*, *37*(1), 20-26.

[2] Watson, C., Harrington, J. & Evans, Z. (1998). An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics,* 18(2), 187-287.

[3] Watson, C., Maclagan, M. & Harrington, J. (2000). Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change,* 12(1), 51-68.

[4] Warren, P. (2018). Quality and quantity in New Zealand English vowel contrasts. *Journal of the International Phonetic Association*, 48(3), 305-330.

[5] Maclagan, M. & Hay, J. (2007). Getting *fed* up with our *feet*: contrast maintenance in the "short" front vowel shift. *Language Variation and Change,* 19 (1), 1-25.

[6] Maclagan, M. & Gordon, E. (1996). Out of the AIR and into the EAR: Another view of the New Zealand diphthong merger. *Language Variation and Change,* 8, 125-147.

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

[8] Britain, D. (2008). When is a change not a change? A case study on the dialect origins of New Zealand English. *Language Variation and Change*, *20*(2), 187-223.

[9] Maclagan, M., Gordon, E. & Lewis, G. (1999). Women and sound change: conservative and innovative behavior by the same speakers. *Language, Variation and Change*, 11 (1), 19-41.

[10] Bauer, L. (1994). English in New Zealand. In Burchfield, R (Ed.), *The Cambridge History of the English Language: Volume 5*. Cambridge: Cambridge University Press. pp.382-429.

[11] Watson, C., Ross, B, Ballard, E., Charters, H., Arnold, R., Meyerhoff, M. (2018) "Preliminary Investigations into Sound Change in Auckland. *The Proceedings of the 17th International Australasian conference on Speech Science and Technology*, Sydney 5-7 December, 2018 pp. 17-20.

[12] Ross, B. (2018). *An Acoustic Analysis of New Zealand English vowels in Auckland.* Unpublished Master's thesis submitted to Victoria University, Wellington in fulfilment of the requirements for the degree of Masters of Arts.

[13] Sloetjes, H. & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation* (LREC 2008)

[14] Kisler, T. Reichel U. D. & Schiel, F. (2017). Multilingual processing of speech via web services, *Computer Speech & Language*, 45, pp. 326–347

[15] Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345

[16] Winkelmann, R., & Raess, G. (2014). Introducing a web application for labeling, visualizing speech and correcting derived speech signals. In the *9th International Conference on Language Resources and Evaluation (LREC 2014)*, pp. 4129-4133.

[17] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

[18] Winkelmann, R., Jaensch, K., Cassidy, S. & Harrington, J. (2018). emuR: Main Package of the EMU Speech Database Management System R package version 1.0.0

[19] Wells, J. C. (1982). *Accents of English 1: An Introduction.* Cambridge: Cambridge University Press.

[20] Watson, C., Maclagan, M., King, J., Harlow, R. & Keegan, P. (2016). Sound change in Māori and the influence of New Zealand English. *Journal of the International Phonetic Association*, 46(2), 185-218.

[21] Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2018). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-137, https://CRAN.R-project.org/package=nlme.

[22] Chevalier, A. (to appear). Internal push, external pull: the reverse short vowel shift in South African English. In Hickey, R (ed.) *English in Multilingual South Africa*. Cambridge, Cambridge University Press.

[23] Maxwell, O. & Fletcher, J. (2010). The acoustic characteristics of diphthongs in Indian English. *World Englishes* 29 (1), 27-44.

[24] Kerswill, P., Torgensen, E. & Fox, S. (2008). Reversing "drift': innovation and diffusion in the London diphthong system. *Language Variation and Change* 20 (3), 451-491.