Variation in FACE and FLEECE trajectories in Australian English adolescents according to community language diversity

Joshua Penney, Felicity Cox, Andy Gibson

Department of Linguistics, Macquarie University, Sydney
joshua.penney@mq.edu.au; felicity.cox@mq.edu.au; andy.gibson@mq.edu.au

ABSTRACT

Australia is an increasingly multicultural society with high levels of linguistic diversity found within its larger cities. Nevertheless, research on Australian English is often based on an Anglo-centric, monolingual model. In this paper, we explore variation in the production of FACE and FLEECE vowels produced by adolescent speakers from areas of Sydney that differ in their level of language diversity and the dominant non-English languages spoken within them. A dynamic vowel formant analysis demonstrates that speakers from more linguistically diverse areas produce a raised first element of FACE and less onglide of FLEECE, compared with speakers from a less linguistically diverse area, whose production of these vowels more closely resembles previous descriptions of 'mainstream' Australian English. Productions from the linguistically diverse areas pattern similarly, indicating the changes are not direct transfer effects.

Keywords: Australian English; diphthongs; sound change; language contact; sociophonetics

1. INTRODUCTION

Australian English (AusE) exhibits a number of phonetic characteristics that distinguish it from other varieties of English. The majority of the described differences occur within the vowel system [8]. In this paper, we focus on two vowels that have long been considered markers of an Australian accent: FACE and FLEECE [10]. FACE, generally transcribed as /æɪ/ in AusE, can be described as a front closing diphthong comprising an open first element in the region of TRAP and a gliding component pointing towards KIT [17]. There is evidence that after a period of historical lowering and fronting described as anticlockwise diphthong shift by [25] and 'drift' by [23], the first element of this vowel may now be in the process of raising towards the vicinity of DRESS [7, 12, 21] in a clockwise reversal of the previously described diphthong shift.

Although FLEECE is generally described as a long monophthong, it shows a delayed target with a long onglide in AusE in which F1 decreases and F2 increases over the course of the vowel from a more centralised location towards its position at the high front extremity of the space, giving this vowel a diphthongal quality [17, 11]. This onglided FLEECE, which may be transcribed as [ᵊiː], is considered characteristic of AusE [9, 21]. The diphthongal nature of AusE FLEECE is made possible by the lowered first element of FACE ensuring separation between the two vowels.

Contemporary Australian society is increasingly multicultural, particularly in larger cities such as Sydney. According to the 2016 national census, more than two thirds of Sydney's residents have at least one overseas-born parent, and nearly 40% of households speak a language other than English at home [2]. Despite such high levels of linguistic diversity, and evidence that this has driven changes to AusE in the recent past [18, 15], until relatively recently much of the research on AusE has been based on what is often referred to as 'mainstream' AusE (MAusE), based on an Anglo-centric, monolingual/monocultural model [6, 13] (though see [5] for a review of studies examining ethnocultural variation in AusE). As part of the Multicultural Australian English project, we are currently building a corpus of audio recordings and survey data from adolescent AusE speakers from areas of Sydney that vary in their levels of community language diversity and the dominant non-English languages spoken within them, including very heterogeneous areas with high proportions of linguistic diversity through to a relatively homogeneous, English-dominant area. This project will enable us to create a more inclusive model of variation in AusE and to examine the effects cultural and linguistic diversity have on sound change.

Previous research has demonstrated that speakers from linguistically diverse communities may propel changes in vowel realisations. For example, [4] found that second-generation AusE speakers with Italian and Greek heritage advanced change away from overtly local (broad) forms that had been adopted by their first-generation parents. More recently in London English, speakers from linguistically diverse communities have been shown to advance change to the diphthong system (among a raft of other changes). This includes the raising and fronting–and in some cases monophthongisation–of FACE [3, 4, 14]. Such changes are likely due to effects of language contact in diverse communities that are then spread through social interaction and accommodation to innovative forms [4, 16].

We here present an analysis of FACE and FLEECE vowels produced by AusE speakers from areas of differing linguistic diversity. We consider the entire shape of the trajectory to account more fully for dynamicity. We may expect to observe that speakers from areas with higher levels of linguistic diversity will produce vowels that differ from those produced by speakers from a more homogeneous, English-dominant area, who we would expect to produce vowels that represent traditional AusE descriptions.

2. METHODS

2.1. Participants and locations

Speech data from 59 adolescent male speakers of AusE aged 15-18 (mean age: 15.8) are included in this analysis. All participants were either born in Australia or migrated to Australia at a very young age. Participants were recruited from high schools throughout Sydney. The schools were selected due to their locations in areas of Sydney that differ in the level of community diversity and the dominant non-English languages spoken within these areas. Five areas of Sydney were selected: the Northern Beaches (NB, n=20) was selected as it has a high level of residents who speak English only and low levels of community language diversity. The other areas were selected as they have high proportions of residents who speak the four most common non-English languages: Arabic in Bankstown (BK, n=9); Chinese (Mandarin and Cantonese) in the Inner West (IW, n=14); Hindi (and other Indo-Aryan languages) in Parramatta (PM, n=5); Vietnamese in Cabramatta (CB, n=11). Note that participants were recruited based on their residence/schooling in these areas, not their language background. Therefore, speakers within each area do not necessarily speak these languages but nonetheless live in the areas in which the dominant non-English languages are spoken. For each of the linguistically diverse areas the majority of participants had a non-English speaking background; for the low diversity area the majority of participants had monolingual English-speaking backgrounds.

2.2. Data collection

41 of the speakers were recorded in a face-to-face setting in a quiet room of their school (n=36), a local library (n=4), or their home (n=1); 18 speakers were recorded remotely via supervised video call while at school during periods of restriction on face to face contact due to COVID-19. Face-to-face participants were recorded to a Zoom H6 recorder through a RODE HS2 headset microphone with a sampling rate of 44.1kHz and 16 Bit quantisation. Remote participants were recorded through an online recorder [https://mmig.github.io/speech-to-flac/] with a sampling rate of either 44.1kHz or 48kHz and 16 Bit quantisation. All files were subsequently resampled to 44.1kHz.

The data for this analysis were collected as part of a picture naming task, in which participants were prompted to produce 225 words and short phrases through the presentation of images on a computer screen. Participants were also recorded while taking part in a spontaneous conversation with a peer and a research assistant; however, these conversational data are not included in this analysis.

For this analysis, we focus on items with FACE and FLEECE vowels produced in single words with /(C(C))VC/ syllable structure, where the coda /C/ was either a voiced or voiceless obstruent (e.g. *spade*, *seat*, etc.). Each word was produced once by each speaker with the exception of one of the words containing FLEECE (*seed*), which was elicited three times for the purposes of formant normalisation. This resulted in 764 items for analysis (FACE: 353; FLEECE: 411).

2.3. Data processing and acoustic analysis

Audio files were segmented with [19] utilising an AusE model. Segment boundaries were subsequently hand corrected by trained annotators. The data were imported into an emu database using [26]. Formant frequencies were estimated with [1] utilising default settings for male speakers (5 formants; 5000Hz ceiling; 0.025s window length) and time normalised. Values in the first and last 15% of the vowel were excluded to minimise effects of neighbouring segments. F1/F2 trajectories were visually inspected and obvious outliers hand corrected. F1/F2 values were *z*-score normalised within speakers [20].

2.4. Statistical analysis

The data were analysed with generalised additive mixed models (GAMMs) using the mgcv [27] and itsadug [24] packages in R [22]. Separate models were fitted for F1 and F2 for each of the two vowels. Each model included a parametric difference term for the five areas (to model constant difference between areas), a smooth over normalised vowel duration (with NB as reference), a smooth over normalised vowel duration by area (an ordered factor modelling non-linear differences from the NB reference for each of the other areas), and a random smooth over normalised vowel duration by participant. For each model basis functions were set to ten (i.e. k=11) and an AR1 error term was included to account for autocorrelation in the models.

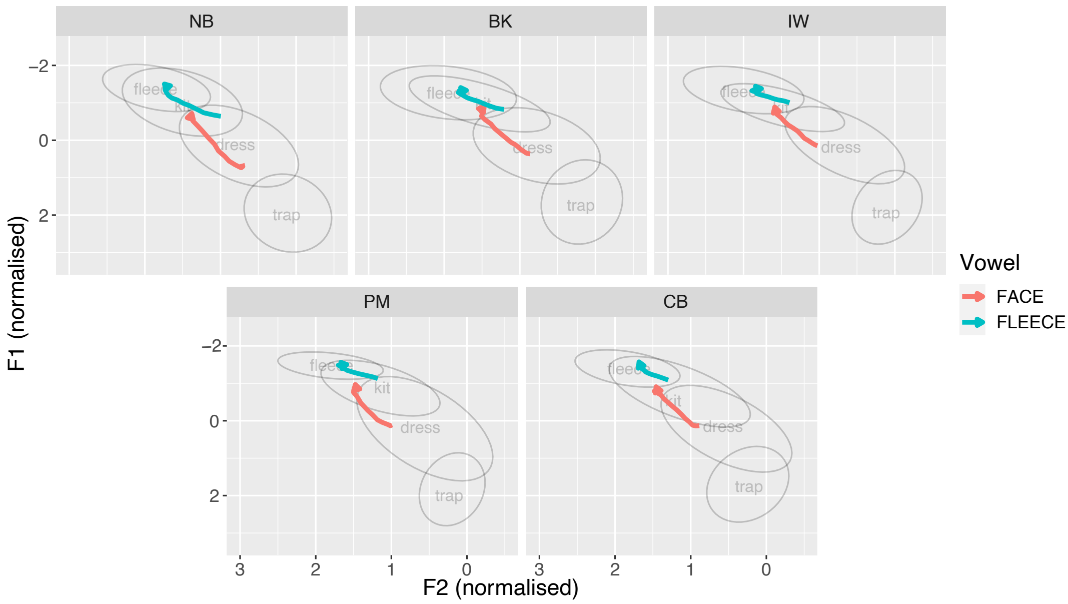

Figure 1. Mean normalised F1/F2 formant trajectories for FACE (red lines) and FLEECE (blue lines) vowels with reference to mean monophthongal vowel space for speakers from five areas of Sydney.

3. RESULTS

Figure 1 illustrates the mean F1/F2 trajectories of FACE and FLEECE for each of the five areas with reference to vowel spaces based on mean target values of monophthongal vowels for each area. (Note that the monophthongs were elicited in the same task as the FACE and FLEECE vowels outlined above; however, their inclusion here is for illustrative purposes only and a complete analysis of the monophthongs is not included in this paper.) From the figure it can be seen that the NB speakers exhibit a lowered and retracted onset of FACE compared to other speaker groups beginning from below the mean of DRESS. This is consistent with recent reports of MAusE FACE decoupling from TRAP that has recently been described for this vowel [12, 21]. In contrast, the speakers from each of the other areas show an even more raised and fronted onset. Additionally, the NB speakers exhibit a clear onglide in FLEECE, as expected according to past descriptions of AusE [17, 9]. The speakers from all of the other areas appear to show a reduced onglide relative to that observed in the NB group, although the BK group appears to produce an onglide that is intermediate between the extremes shown by the NB group and the other three groups.

The results showed significant parametric (i.e. constant) and non-linear effects for the GAMM analyses for both F1 and F2 of FACE for each of the areas relative to the reference NB area (see Table 1 for a summary). This indicates that speakers from all four of the areas showed lower F1 values (i.e. higher realisations) and higher F2 values (i.e. fronter realisations) in general compared to the NB group, but that they also showed differences in the shape of the vowel trajectory. Figure 2 shows the predicted values of the GAMMs for F1 and F2 of FACE and illustrates these differences between the areas.

Area/Vowel	BK	IW	PM	CB
FACE F1 (parametric)	*	*	*	*
FACE F1 (non-linear)	*	*	*	*
FACE F2 (parametric)	*	*	*	*
FACE F2 (non-linear)	*	*	*	*
FLEECE F1 (parametric)	-	*	*	*
FLEECE F1 (non-linear)	*	*	*	*
FLEECE F2 (parametric)	*	*	-	-
FLEECE F2 (non-linear)	-	*	*	*

Table 1: Summary of significant differences from the reference EN group for GAMMs for F1 and F2 in FACE and FLEECE. Significant effects (at $\alpha = 0.05$) are marked by asterisks.

For FLEECE, the F1 model showed significant parametric differences for all of the groups except BK, and significant non-linear differences for all areas relative to the reference NB group. For F2, there were significant parametric differences for BK and IW relative to the NB reference group, and significant non-linear differences for all areas apart from BK. This indicates that all of the groups, apart from BK, showed overall lower F1 values (i.e. higher realisations) in addition to differences in trajectory shape of F1. For F2, the BK and IW groups showed overall higher values (i.e. fronter realisations) than the NB group; however, the PM and CB groups did not show overall fronter realisations. With the exception of the BK group, all of the groups produced F2 trajectories that differed in shape from the NB group. Figure 3 shows the predicted values of the GAMMs for F1 and F2 of FLEECE and illustrates the differences between the areas.

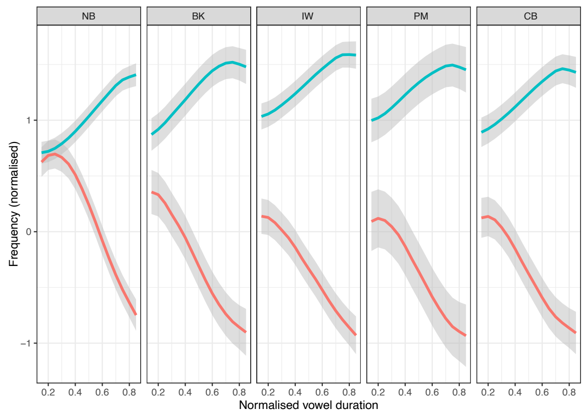

Figure 2: Model predictions of normalised F1 (red lines) and F2 (blue lines) of FACE by normalised vowel duration for speakers from five areas of Sydney. Error ribbons show 95% CIs.

4. DISCUSSION

The results described above demonstrate that there is variation in the production of these two vowels in AusE according to the area a speaker comes from. We assume that the trajectories of the NB area group can be taken to represent the productions of AusE speakers with reduced influence from speakers with language backgrounds other than English compared to the other groups. Correspondingly, as seen in Figure 1, their productions conform to previous descriptions of 'mainstream' AusE: for FACE, we observe a lower first element compared to the other groups with evidence of the recently reported raising change to this vowel, and for FLEECE, we observe the characteristic onglide.

On the other hand, compared to descriptions in the AusE literature, the four other groups appear to produce innovative realisations of these vowels: FACE and FLEECE are both produced with markedly shorter trajectories by all of these groups (see Figure 1) – and, for the most part, with higher and fronter onsets – and for FACE, a higher offset is also evident. As these areas represent communities with high levels of linguistic diversity, we interpret these results as indicating change to these vowels due to language/dialect contact within these communities.

Interestingly, speakers from the linguistically diverse areas seem to pattern together and produce similar realisations of these vowels, although the BK group appear to produce realisations that lie somewhere between the NB group and the other groups. This would seem to indicate that the changes are not due to direct transfer effects from one of the non-English languages spoken in these areas, or that, if they originally were due to such transfer effects, these have now spread through the linguistically diverse areas of Sydney beyond those with a particular language background [3, 4]. Additionally, at least with regard to FACE, the innovative realisations appear to be an advancement of the raising change that has been noted for MAusE speakers. While we cannot speculate as to the initiation of this change, it would appear that it is now being driven by speakers in the linguistically diverse areas, with the speakers from the homogeneous area lagging behind. Taken together, these findings may suggest a shift in speakers from linguistically diverse areas away from overtly local, or mainstream, AusE forms, similar to the observations made by [18].

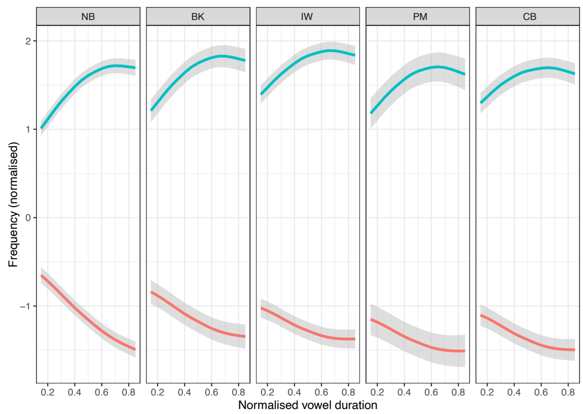

Figure 3: Model predictions of normalised F1 (red lines) and F2 (blue lines) of FLEECE by normalised vowel duration for speakers from five areas of Sydney. Error ribbons show 95% CIs.

Another important aspect of this analysis is that the vowels in FACE and FLEECE appear to pattern together in that greater onglide for FLEECE (non-linear F1 effect for NB compared to the other groups) occurs with a lowered first element of FACE in the NB speakers (both parametric and non-linear differences between NB and other groups in F1). Conversely, for the speakers from areas of high diversity, reduced FLEECE onglide patterns with a more raised first element of FACE. The relationship between these two vowels in chain shifting of the type described by [25] and [23] is a fruitful area for future research.

Finally, we note that this analysis is preliminary, and that a limitation is that some of the groups have fewer speakers, and hence fewer items, than others. Data collection for this project is ongoing, and in our future research we aim to broaden our scope to include both female and male speakers, and to extend the analysis to the entire AusE diphthong system.

5. ACKNOWLEDGEMENTS

This research was supported by Australian Research Council Future Fellowship Grant FT180100462 to the second author.

6. REFERENCES

[1] Albin, A. 2014. PraatR: An architecture for controlling the phonetics software "Praat" with the R programming language. *J. Acoust. Soc. Am 135*, 2198.

[2] Australian Bureau of Statistics. 2016. *Census 2016*.

[3] Cheshire, J., Fox, S., Kerswill, P., Torgersen E. 2008. Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica 22*, 1–23.

[4] Cheshire, J., Kerswill, P., Fox, S., Torgersen E. 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. *J. Sociolinguistics 15*, 151–196

[5] Clothier, J. 2020. Ethnolectal variability in Australian Englishes. In: Willoughby, L, Manns, H., (eds.), *Australian English reimagined*. Routledge, 155–171.

[6] Clyne, M., Eisikovits, E., Tollfree, L. 2001. Ethnic varieties of Australian English. In: Blair, D., Collins, P. (eds.), *English in Australia*. Benjamins., 223–238.

[7] Cox, F. 1999. Vowel change in Australian English. *Phonetica 56*, 1–27.

[8] Cox, F. 2020 Phonetics and phonology of Australian English. In: Willoughby, L, Manns, H., (eds.), *Australian English reimagined*. Routledge, 15–33.

[9] Cox, F., Palethorpe, S. 2007. Illustrations of the IPA: Australian English. *J. Int. Phon. Assoc. 37*, 341–350.

[10] Cox, F., Palethorpe, S. 2012. Standard Australian English: The sociostylistic broadness continuum. In: Hickey, R. (ed.), *Standards of English: Codified Varieties around the World*. Cambridge, 294–317.

[11] Cox, F., Palethorpe, S., Bentink, S. 2014. Phonetic archaeology and fifty years of change to Australian English /iː/. *Aust. J. Ling. 34*, 50–75.

[12] Cox, F., Palethorpe, S., Harrington, J. 2015. *Diphthong shift in Australian English,* presented at Australian Linguistic Society Conference, Sydney.

[13] Cox, F., Penney, J., Palethorpe, S. 2022. Fifty years of change to prevocalic definite article allomorphy in Australian English. *J. Int. Phon. Assoc, 1–31. doi:10.1017/S002510032200007X*

[14] Gates, S. M. 2018. Why the long FACE? Ethnic stratification and variation in the London diphthong system. *U. Penn. Working Papers in Linguistics 24*, 39–48.

[15] Grama, J., Travis, C. E., Gonzalez, S. 2020. Ethnolectal and community change ov(er) time: Word-final (er) in Australian English. Aust. J. Ling. 40, 346–368.

[16] Harrington, J. Gubian, M., Stevens, M., Schiel, F. 2019. Phonetic change in an Antarctic winter. *J. Acoust. Soc. Am. 146*, 3327–3332.

[17] Harrington, J., Cox, F., Evans. Z. 1997. An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Aust. J. Ling. 17*, 155–184.

[18] Horvath, B. 1985. Variation in Australian English: The sociolects of Sydney. Cambridge.

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

[20] Lobanov, B. 1971. Classification of Russian vowels spoken by different speakers. *J. Acoust. Soc. Am. 49*, 606–608.

[21] Purser, B., Grama, J., Travis, C. 2020. Australian English over time: Using sociolinguistic analysis to inform dialect coaching. *Voice and Speech Review 14,* 269–291.

[22] R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

[23] Trudgill, P. 2004. *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh.

[24] van Rij, J., Wieling, M., Baayen, R. H., van Rinj, H. 2022. Itsadug: Interpreting time series and autocorrelated data using GAMMs. https://CRAN.R-project.org/package=itsadug.

[25] Wells, J. C. 1982. *Accents of English*. Cambridge.

[26] Winkelmann, R., Harrington, J., Jänsch, K. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language 45*. 392-410.

[27] Wood, S N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. Royal Stat. Soc. 73,* 3–36.