Spectral contrast reduction in Australian English prelateral vowels

Tunde Szalay, Titia Benders, Felicity Cox, Michael Proctor
Macquarie University

Coarticulation causes predictable variation in speech, with the potential to affect phonological contrast [1]. Vowel-lateral coarticulation in the rime is known to reduce or neutralise phonemic vowel contrast in several varieties of English [2, 3, 4, 5]. In Australian English (AusE), the F1-F2 vowel space is reduced in this context due to the phonetic backing of prelateral front vowels [3, 5]. In particular, spectral contrast reduction for long/short vowel pairs may occur between /ʉː-ʊ/ (*pool-pull*), /əʉ-ɔ/ (*dole-doll*), /æɔ-æ/ (*vowel-Val*), and /iː-ɪ/ (*heel-hill*) [3, 5, 7]. However, these observations have been primarily based on impressionistic analysis or visual representation of formants. Therefore, we aimed to systematically analyse acoustic contrast reduction in AusE lateral-final rimes. We hypothesised that spectral contrast would be reduced between the pairs /ʉː-ʊ, əʉ-ɔ, æɔ-æ, iː-ɪ/ compared to other prelateral vowels [6, 5].

29 female native monolingual speakers of AusE produced 16 stressed vowels in the /hVl/ paradigm. 16 unique /hVd/ words contrasting the same vowels were recorded as a baseline. Each word was repeated 3 times. Formant trajectories from the vowel in the pre-/d/ context and from the rime in the pre-/l/ context were extracted. Dynamic formant trajectories were modelled using the first three discrete cosine transformation (DCT) coefficients of the first three formants; i.e. 9 coefficients (3 coefficients×3 formants) characterised each token [7]. Vowel duration in the /d/-context and rime duration in the /l/-context were also measured.

DCT coefficients and duration values were used to train two random forest models [8]: one for classifying pre-obstruent vowels, and one for lateral-final rimes. Random forest is a supervised classification algorithm consisting of a training- and a testing phase. During the training phase, random forest learnt the 16 vowel categories based on DCT coefficients, duration values, and category labels in each coda condition, using bootstrap samples from 75% of the data [8]. Out-of-bag error rate indicates the accuracy of category learning in the training phase. In the testing phase, random forest classified the remaining data into 16 vowel categories in each coda condition using DCT coefficients and duration values. Comparing the random forest classification to the original vowel labels yielded two confusion matrices (Fig. 1).

Out-of-bag error rates show that vowel categories were learnt more accurately in the /d/- than in the /l/-context (3.6% vs 24.1%). In the /d/-context, seven vowels were classified with 100% accuracy, including /ɪ, ʉː, əʉ, æɔ/ (KIT, GOOSE, GOAT, MOUTH); the least accurately identified vowels were /ɜː/ (NURSE) and /ɔ/ (CLOTH) (Fig. 1). In the /l/-context, only /e/ (DRESS) was identified with 100% accuracy. Members of the pairs /ʉː-ʊ, əʉ-ɔ, æɔ-æ, iː-ɪ/ were confused with high frequency: 26% and 28% of /ʉː/ and /ʊ/ were confused with each other. /ɔ/ was more often misidentified as /ɐ/ (STRUT) (32%) than /əʉ/ (16%); however, /əʉ/ was misidentified as /ɔ/ (52%). /æɔ/ and /æ/ tokens were confused with each other (30%). 19% of /iː/ tokens were misidentified as /ɪ/, and 5% of /ɪ/ as /iː/. These pairs were not confused in the /d/-context (Fig. 1).

The results support our hypothesis: spectral contrast was reduced between the vowels within the pairs /ʉː-ʊ, əʉ-ɔ, æɔ-æ, iː-ɪ/ in the /l/-, but not in the /d/-context. Contrast reduction can be attributed to coarticulation with the dorsal gesture of /l/ and the dorsal gesture's similarity to /ɔ/ [9]. As /ʉː/ and [ʉ] in /əʉ/ are backed [10] before coda /l/, F2 is lowered, becoming similar to /ʊ, ɔ/. /æ/ followed by an /ɔ/-like /l/ can be spectrally similar to /æɔ/ (Fig. 2). Articulatory similarity between /ɔ/ and /l/ can cause /l/ to encroach on /əʉ, æɔ/, leading to the loss of the second target (Fig. 2). /iː-ɪ/ are less frequently confused as both vowels are backed to a similar extent, so that /iː/ remains more peripheral (Fig. 2). Acoustic contrast reduction between the members of the pairs /ʉːl-ʊl, əʉl-ɔl, æɔl-æl, iːl-ɪl/ has implications for a potential ongoing vowel change in prelateral environments in AusE [1].

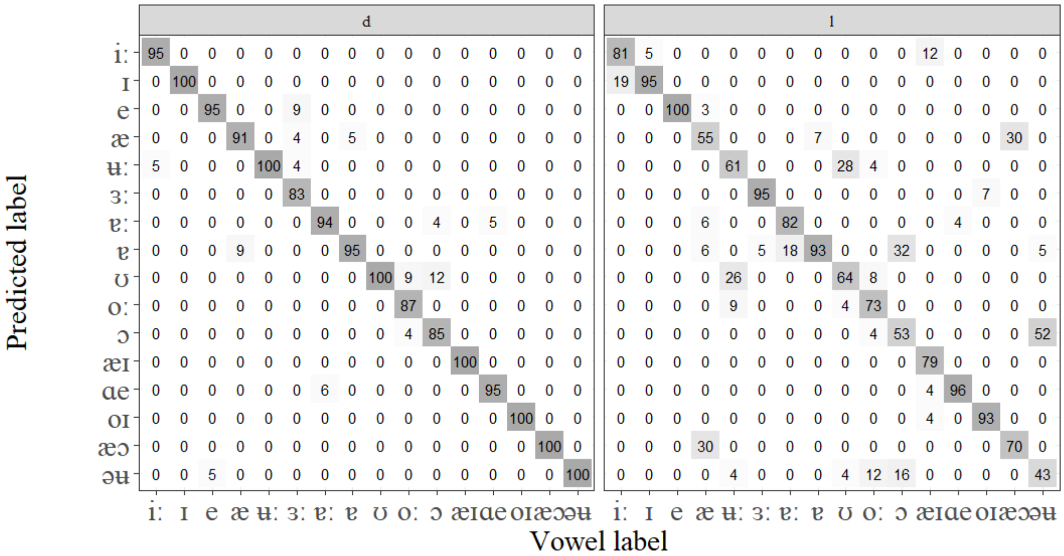

Fig. 1. *Confusion matrices produced by random forest classification.*

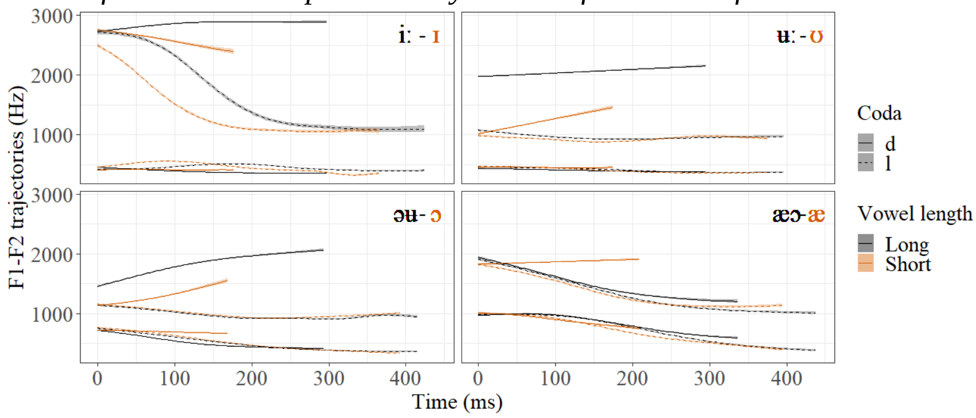

Fig. 2. *Mean F1 and F2 trajectories by coda and vowel pair.*

[1] Garrett, A., & Johnson, K. (2013). "Phonetic bias in sound change," in *Origins of Sound Change*, ed. Yu, Alan, C.L. (OUP Oxford, UK), 51–97.

[2] Altendorf, U., & Watt, D. (2008). "The dialects in the South of England: phonology," in *Varieties of English*, ed. B. Kortmann & C. Upton, (Mouton: Berlin), 194–222

[3] Cox, F., & Palethorpe, S. (2004). "The border effect: Vowel differences across the NSW-Victorian border," in *Proc Conf of Australian Linguistics Society*.

[4] Labov, W., Ash, S., & Boberg, C. (2008). *The Atlas of North American English, Phonetics, Phonology and Sound Change* (Mouton: Berlin).

[5] Palethorpe, S., & Cox, F. (2003). "Vowel modification in pre-lateral environments," in *Int Seminar on Speech Production*, Sydney.

[6] Bradley, D. (2004). "Regional characteristics of Australian English: phonology," in *Varieties of English: The Pacific and Australasia*, ed. Burridge, K. & B. Kortmann, (Mouton: Berlin), 645–655.

[7] Harrington, J. & Cassidy, S. (1994). "Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English," *Lang Speech* 37(4), 357–373.

[8] Liaw, A., & Wiener, M. (2002). "Classification and regression by randomForest", *R News* 2(3), 18–22.

[9] Gick, B., Kang, M. A., & Whalen, D. H. (2002). "MRI evidence for commonality in the post-oral articulations of English vowels and liquids," *J Phon* (30), 357–371.

[10] Lin, S., Palethorpe, S. & Cox, F. (2012). "An ultrasound exploration of Australian English /CVl/ words," in *Proc 14th SST*, Sydney, 105–108.