FEBRUARY 18 2021

Spectral contrast reduction in Australian English /l/-final rimes

Tünde Szalay ; Titia Benders ; Felicity Cox ; Sallyanne Palethorpe ; Michael Proctor

J Acoust Soc Am 149, 1183–1197 (2021)
https://doi.org/10.1121/10.0003499

JASA ARTICLE

Spectral contrast reduction in Australian English /l/-final rimes

Tünde Szalay,[a)] Titia Benders,[b)] Felicity Cox,[c)] Sallyanne Palethorpe,[d)] and Michael Proctor[e)]
Department of Linguistics, Macquarie University, New South Wales, Australia

ABSTRACT:
Vowel contrasts may be reduced or neutralized before coda laterals in English [Bernard (1985). *The Cultivated Australian: Festschrift in Honour of Arthur Delbridge*, pp. 319–332; Labov, Ash, and Boberg (2008). *The Atlas of North American English, Phonetics and Sound Change* (Gruyter Mouton, Berlin); Palethorpe and Cox (2003). *International Seminar on Speech Production* (Macquaire University, Sydney, Australia)], but the acoustic characteristics of vowel-lateral interaction in Australian English (AusE) rimes have not been systematically examined. Spectral and temporal properties of 16 pre-lateral and 16 pre-obstruent vowels produced by 29 speakers of AusE were compared. Acoustic vowel similarity in both environments was captured using random forest classification and hierarchical cluster analysis of the first three DCT coefficients of $F1$, $F2$, and $F3$, and duration values. Vowels preceding /l/ codas showed overall increased confusability compared to vowels preceding /d/ codas. In particular, reduced spectral contrast was found for the rime pairs /iːl-ɪl/ (*feel-fill*), /ʉːl-ʊl/ (*fool-full*), /əʉl-ɔl/ (*dole-doll*), and /æɔl-æl/ (*howl-Hal*). Potential articulatory explanations and implications for sound change are discussed.
 https://doi.org/10.1121/10.0003499

(Received 17 September 2020; revised 8 January 2021; accepted 20 January 2021; published online 18 February 2021)
[Editor: Ewa Jacewicz] Pages: 1183–1197

I. INTRODUCTION

Coarticulation, the influence of adjacent sounds on each other, causes predictable variation in speech with the potential to affect phonological contrast (Garrett and Johnson, 2013; Hyman, 2013). Vowel-lateral coarticulation in particular may reduce or neutralise phonemic vowel contrast in several varieties of English, including Australian English (AusE) (Altendorf and Watt, 2008; Cox and Palethorpe, 2004; Labov *et al.*, 2008; Palethorpe and Cox, 2003; Wade, 2017). In AusE, vowel-lateral coarticulation has been shown to compress the $F1$–$F2$ vowel space due the phonetic backing of front vowels in the pre-lateral environment (Bernard, 1985; Cox and Palethorpe, 2004; Palethorpe and Cox, 2003). For instance, contrast reduction is regularly observed between *pool* and *pull* (Bernard, 1985; Cox and Palethorpe, 2004; Palethorpe and Cox, 2003). However, carefully controlled and systematic analysis of AusE vowels is required to further our understanding of how coda laterals influence preceding vowels and reduce vowel contrast.

Both acoustic and perceptual vowel contrast reduction in pre-lateral environments have been reported in several dialects of English. The contrast between the high vowels in *feel-fill* is reduced in some Southern dialects of American English and in Standard Southern British English, through the phonetic lowering of the tense vowel /iː/ in the pre-lateral environment (Altendorf and Watt, 2008; Harris, 1994; Labov *et al.*, 2008; Turton, 2014). The *pool-pull* contrast is reduced in Pennsylvanian and Southern British due to the phonetic lowering of the vowel in *pool* (Altendorf and Watt, 2008; Labov *et al.*, 2008). The same contrast is also reduced in South AusE, through a different mechanism: phonetic backing and lowering of the tense vowel /ʉː/ in the pre-lateral environment (Butcher, 2006; Oasa, 1989). The acoustic *pool-pull-pole* contrast is reduced in Ohio, as the vowels in *pool* and *pole* shift towards *pull* in the vowel space (Arnold, 2015; Wade, 2017). A perceptual merger between the mid and low front vowels /e/ and /æ/ has been observed in the pre-lateral environment (*hell-Hal*) in New Zealand English (Thomas and Hay, 2005) and in Melbourne English (e.g., Loakes *et al.*, 2014; Loakes *et al.*, 2012). Collectively, these findings suggest that different, dialect-specific mechanisms may be involved in vowel-lateral interactions in different varieties of English. These findings may be consistent with potential sound change, as acoustic and perceptual contrast reduction caused by coarticulation is often implicated in the initiation of sound change (Blevins, 2006; Harrington *et al.*, 2018; Ohala, 1989, 1993); however, only a few apparent-time studies have explored such mergers in pre-lateral contexts (e.g., Kleber *et al.*, 2012; Strycharczuk and Scobbie, 2017).

A. Pre-lateral vowels in AusE

1. AusE vowel inventory

In AusE, coda /l/ has been shown to influence vowels in ways that can potentially reduce perceptual and acoustic vowel contrast, especially between the pairs /ʉːl-ʊl, əʉl-ɔl,

a)Electronic mail: tunde.szalay@mq.edu.au, ORCID: 0000-0002-6982-5639.
b)ORCID: 0000-0003-0143-2182.
c)ORCID: 0000-0001-8479-7624.
d)ORCID: 0000-0002-2351-5539.
e)ORCID: 0000-0002-3083-6859.

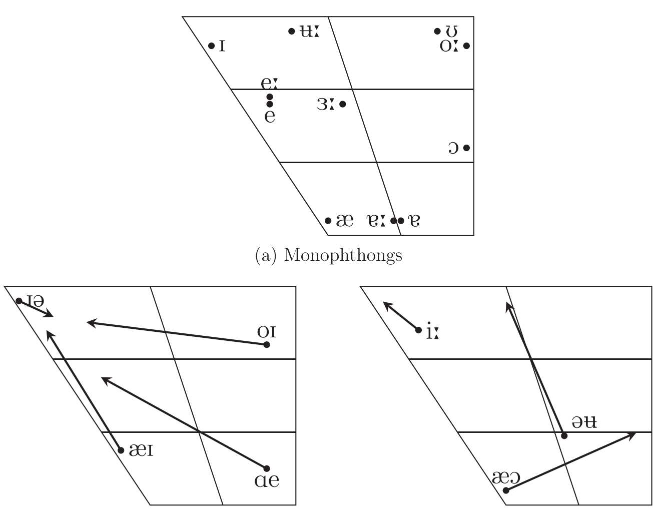
https://doi.org/10.1121/10.0003499

æɔl-æl/ (Bernard, 1985; Bradley, 2004; Palethorpe and Cox, 2003; Szalay *et al.*, 2018). AusE has a large vowel inventory consisting of 18 stressed vowels and schwa (Fig. 1), utilising both spectral and durational contrast (Cox and Fletcher, 2017). Duration is contrastive for spectrally similar vowels; for instance, the vowel pairs /ɐː-ɐ, eː-e/ (e.g., *card-cud, shared-shed*) contrast mostly in duration, whereas the pairs /ʉː-ʊ/ (e.g., *Luke-look*) contrast both in duration and in spectral quality (Cox and Fletcher, 2017).

The AusE vowel system includes both diphthongs (/æɪ, ɑe, oɪ, æɔ, əʉ, ɪə/) and monophthongs (front: /iː, ɪ, e, eː, æ/, central: /ʉː, ɜː, ɐː, ɐ/, and back: /oː, ɔ, ʊ/)[1] (Harrington *et al.*, 1997; Watson and Harrington, 1999). The vowels /iː/ and /ʉː/ are classified as monophthongs but are characterised by an onglide (Bernard, 1970; Cox and Palethorpe, 2007; Cox *et al.*, 2014; Harrington *et al.*, 1997). /ɪə/ is classified as a diphthong, but can also be realised as a monophthong [ɪː] or as disyllabic [ɪːə] (Cox and Palethorpe, 2007; Harrington *et al.*, 1997). Some of the diphthongs, having similar first or second target characteristics to a monophthong, form pairs with monophthongs (Cox, 1999). For example, /æɔ/ and /æ/ (e.g., *loud-lad*) share the first target of the diphthong, whereas /əʉ/ shares the location of the second target with the nucleus of /ʉː/ (e.g., *code-cooed*) (Cox, 1999). These vowels are considered pairs, as the members have moved in parallel in sound change: /æɔ/ lowering, was accompanied by /æ/ lowering, and the fronting of /ʉː/ took place in parallel with the fronting of the second element of /əʉː/ (Cox, 1999).

2. Effect of coda /l/ on monophthongs

AusE pre-lateral vowels differ from their non pre-lateral counterpart in many ways: front and central vowels are phonetically lowered and backed and some low and back vowels are phonetically raised (Bernard, 1985; Cox and Palethorpe, 2004; Palethorpe and Cox, 2003). Front and central vowels exhibit phonetic lowering before /l/ shown by increased $F1$ in /iː, ɪ, e,/ and /ʉː, ɜː/ (Bernard, 1985; Cox and Palethorpe, 2004; Palethorpe and Cox, 2003). Front /iː, ɪ, e, æ/ and central /ʉː, ɜː/ are also characterised by lowered $F2$ representing phonetic retraction before /l/ (Bernard, 1985; Cox and Palethorpe, 2004; Palethorpe and Cox, 2003). Among the low and back vowels only /ɔ, ʊ/ and /ɐ/ are influenced by coda /l/, as the former two exhibit phonetic backing and the latter phonetic raising (Bernard, 1985; Cox and Palethorpe, 2004; Palethorpe and Cox, 2003).



(a) Monophthongs

(b) Centring diphthongs.

(c) Rising diphthongs.

FIG. 1. General AusE vowel inventory. Figure reproduced from Cox and Fletcher (2017).

3. Acoustic and auditory contrast reduction of some vowel pairs

Spectral contrast reduction between the members of the pairs /ʉːl-ʊl/, /əʉl-ɔl/, and /æɔl-æl/ has been shown through auditory-impressionistic observations and visual representations of formant trajectories (Bernard, 1985; Bradley, 2004; Palethorpe and Cox, 2003). Spectral contrast between /ʉːl-ʊl/ and /əʉl-ɔl/ is reduced due to the $F2$ lowering in pre-lateral /ʉː/ and in both the first and second target of the diphthong /əʉ/ in pre-lateral context (Palethorpe and Cox, 2003). Bernard (1985) observed that the second target of the diphthongs /əʉ, æɔ/ is frequently lost before /l/ and commented on the lack of observable change between the end of the vowels /æɔ, əʉ/ and /l/, which can potentially contribute to spectral contrast reduction between the members of the pairs /əʉl-ɔl/ and /æɔl-æl/. While the acoustic targets of /iːl-ɪl/ are distinct, the loss of the onglide of /iː/ and the schwa-like offglide of both vowels may increase spectral similarity (Palethorpe and Cox, 2003). However, duration contrast between the members of these pairs is maintained (Palethorpe and Cox, 2003).

In line with the acoustic contrast reduction, perceptual contrast reduction between the members of the pairs /ʉːl-ʊl/, /əʉl-ɔl/, and /æɔl-æl/ has been noted (Loakes *et al.*, 2012; Szalay *et al.*, 2018). As spectral contrast is reduced between the members of these pairs, listeners rely on duration cues: listeners who maintain a larger duration contrast in their own production perceive the members of these pairs more accurately if the speaker maintains a larger duration contrast too (Szalay *et al.*, 2018).

4. Regional differences

The Victorian dialect of AusE shows contrast reduction between /el-æl/ and /ɐlC-ɔlC/ in production and in perception (Bernard, 1985; Cox and Palethorpe, 2004; Lewis and Loakes, 2012; Loakes *et al.*, 2010c). The $F1$ of pre-lateral /e/ is increased towards /æ/ (e.g., *hell, Hal*) (Cox and Palethorpe, 2004). The acoustic contrast reduction is reflected in a perceptual near-merger between /el/ and /æl/, as Victorian English listeners misperceive /el/ as /æl/ in the pre-lateral, but not in a pre-obstruent position when distinguishing minimal pairs (e.g., *telly-tally, pellet-palate*) (Loakes *et al.*, 2010a,b,c, 2011). Acoustic and perceptual contrast reduction has also been shown between /ɐlC-ɔlC/ *gulf-golf*) (Bernard, 1985; Lewis and Loakes, 2012).

JASA https://doi.org/10.1121/10.0003499

Phonetic backing of pre-lateral /ʉː/ is more prominent in South Australia than in New South Wales (NSW), despite pre-lateral /ʉː/ being backed in NSW as well. There is also some evidence for a potential merger between pre-lateral /iː/ and /ɪ/ in Adelaide and Hobart (Bradley, 2004), but not in NSWs, where this pair appears to be differentiated by length (Palethorpe and Cox, 2003).

Collectively, these findings suggest that coda /l/ impacts the preceding vowel in various ways depending on vowel quality and speaker dialect and might lead to potential mergers. However, the impact of coda /l/ on preceding vowels and the potential for a loss of contrast has not been systematically examined. Several of the observations on the effect of coda /l/ and on apparent contrast reduction were made only on the basis of impressionistic observations or visual representations of formants (Bernard, 1985; Palethorpe and Cox, 2003).

B. Aims and hypotheses

The aim of the present study was to systematically characterise the spectral properties of AusE vowels produced in pre-lateral environments and determine the impact of vowel-lateral coarticulation on vowel contrast. We hypothesised that in the pre-lateral context (1) front vowels would have a higher $F1$; (2) front vowels would have a lower $F2$; and (3) spectral contrast would be reduced between /ʉːl-ʊl, əʉl-ɔl, æɔl-æl/, and possibly also between /iːl-ɪl/.

To test hypotheses (1) and (2), and also to systematically characterise the effect of coda /l/ on the spectral properties of non-front vowels, we examined the effect of /l/ on $F1$, $F2$, and $F3$ of monophthong targets in /l/-final rimes compared to monophthong targets in /d/-final rimes. To test hypothesis (3) and to systematically characterise spectral contrast reduction in the pre-lateral vowel space, we modelled the dynamic properties of pre-/d/ monophthongs and diphthongs and each of the entire lateral-final rimes using discrete cosine transformation (DCT, see Sec. II C) of the first three formants. We quantified spectral contrast and similarity using random forest classification and agglomerative hierarchical cluster analysis of AusE vowels based on duration values and the first three DCT coefficients of $F1$, $F2$, and $F3$.

II. METHODS

A. Participants

Data from 29 female native monolingual speakers of AusE, born in NSW to Australian-born parents (year of birth: 1981–1992, age at recording: 18–27, mean age at recording = 20.2 years) were analysed. None of the participants reported any speaking, hearing, or reading, difficulties.

B. Material and procedure

Sixteen stressed vowels of AusE were elicited in two monosyllabic paradigms: hVd and hVl (Table I). All phonotactically legal words and non-words were elicited in these two contexts. The vowels /ɪə/ and /eː/ were not elicited in the /l/-context as /ɪəl/ and /eːl/ are phonotactically illegal. The elicitation items varied in lexical frequency and included seven non-words in the /l/ context: *hal, hule, harl, hooll, holl, hile, hoil*, and seven in the /d/ context: *hude, hud, hod, hade, hoyd, howd, hode*. Although word frequency affects vowel acoustics, with vowels in more frequent words being more contracted (Munson and Solomon, 2004), a mix of high frequency, low frequency, and non-words were included to provide a consistent phonetic frame of reference.

TABLE I. Orthographic representation and International Phonetic Alphabet (IPA) transcription of target words. Left columns: /l/-final targets. Right columns: /d/-final targets. Non-words are underlined.

Coda /l/		Coda /d/	
Orthography	IPA	Orthography	IPA
heel	hiːl	*heed*	hiːd
hill	hɪl	*hid*	hɪd
hell	hel	*head*	hed
hal	hæl	*had*	hæd
hule	hʉːl	*who'd*	hʉːd
hurl	hɜːl	*herd*	hɜːd
harl	hɐːl	*hard*	hɐːd
hull	hɐl	*hud*	hɐd
hooll	hʊl	*hood*	hʊd
hall	hoːl	*horde*	hoːd
holl	hɔl	*hod*	hɔd
hail	hæɪl	*hade*	hæɪd
hile	hɑel	*hide*	hɑed
hoil	hɔɪl	*hoyd*	hɔɪd
howl	hæɔl	*howd*	hæɔd
hole	həʉl	*hode*	həʉd

Speakers read each word as it was presented orthographically on a computer monitor. Non-words were accompanied by a rhyming helper word, e.g., *hule—sounds like tool*. Recordings were monitored by a phonetically trained native speaker of AusE, and participants were asked to repeat erroneous items again with the correct pronunciation using the rhyming prompt—no items were modelled by the researcher.

Items were presented in random order in three blocks. The task also included practice words at the beginning of the session—none of which contained coda /d/ or /l/—and vowels produced in other contexts (hV, hVn, hVt). After each block of words, ten short sentences were elicited.

Participants were recorded between 2004 and 2009 in a sound treated recording studio at Macquarie University, Sydney. Speech data were captured using an AKG C535 EB microphone, Cooledit 2000 audio recording software via M-Audio delta66 soundcard to a Pentium 4 PC at 44.1 kHz sampling rate.

C. Phonetic analysis

32 (targets) × 3 (repetitions) × 29 (participants) – 1 = 2783 tokens were analysed; a repetition of *who'd* is missing for one speaker. Segment boundaries were automatically located using

the MAUS forced aligner (Kisler *et al.*, 2017; Schiel, 1999) with the AusE grapheme-to-phoneme converter, and manually corrected. The vowel onset was determined on the basis of voicing onset and sudden increase in amplitude (M0; Fig. 2). The vowel-/d/ boundary was determined on the basis of amplitude drop (M1; Fig. 2). The rime offset in /l/-final targets was not corrected (M1; Fig. 2). Because there is no discernible boundary between the vowel and the following /l/ in /hVl/ words, the entire /hVl/ rime was analysed instead of selecting an arbitrary boundary in the vowel-lateral transition (Fig. 2). Segmentation errors were corrected by a trained phonetician only when vowel onset or the vowel offset before coda /d/ was misplaced by more than 30 ms. To minimise potential imprecisions in formant measurements, the first and the last 30 ms of the vowel and the rime were discarded prior to extracting formant values (T0 and T1 in Fig. 2). A boundary threshold larger than the customary 20 ms was chosen because pre-trained force aligners have been shown to be less accurate than train/align models, but are more appropriate for a relatively small dataset like the present one (Fromont and Watson, 2016; González *et al.*, 2018).

Formant frequencies were estimated at every 10 ms throughout the analysis window from a 5 ms Gaussian window with 75% overlap and 25 ms formant analysis window with 55 dB dynamic range and a pre-emphasis filter increasing spectral slope above 2700 Hz by 6 dB/octave in Praat (Boersma and Weenink, 2013). To optimise formant settings for each speaker, four formants were tracked up to 4500 Hz ceiling for speakers who produced comparatively lower $F2$ and $F3$ or five formants were tracked up to a maximum frequency of 5000 Hz for speakers who produced a comparatively higher $F2$ or $F3$ trajectory. Formant trajectories were manually corrected by the first author using a MATLAB-based interface that superimposed formant estimates over a broad band spectrogram calculated over 5 ms windows with 40% overlap, allowing for corrections of estimates that did not align with the visible formants. After hand-correction, all formant values 1.5 times above or below the interquartile range for each formant in each vowel were rechecked.

Acoustic targets of monophthongs were located automatically in the corrected formant trajectories using the following criteria:

- $F1$ maximum
  - low vowels (/æ, ɐː, ɐ, ɔ/) before /d/ and /l/, as $F1$ maximum indicates the phonetically lowest point
- $F2$ minimum
  - high back vowels (/ʊ, oː/) before /d/, as $F2$ minimum indicates the phonetically backmost point
- $F2$ maximum
  - high front vowels (/iː, ɪ, e/) before /d/ and /l/, as $F2$ maximum indicates the phonetically frontmost point
  - /ʉː/ before /d/, as /ʉː/ is a fronted vowel, characterised by a high $F2$ in AusE
  - /ɜː/ before /l/, as $F2$ lowers considerably between an /ɜː/ target and an /l/ target
- Temporal midpoint
  - /ɜː/ before /d/, as the formant trajectories of mid-central /ɜː/ do not show considerable formant change in the pre-/d/ context
- 25% of the rime
  - high back vowels (/ʊ, ʉː, oː/) before /l/, as there is no considerable formant change in these rimes

Neither the first nor the second acoustic target were located for diphthongs, as several diphthong tokens did not exhibit two targets in the pre-lateral context.

DCTs were used to model the major dynamic properties of vowels in both types of rimes using EMUR (Harrington and Cassidy, 1994; Watson and Harrington, 1999; Winkelmann *et al.*, 2019). The first three DCT coefficients characterise formant changeover time: the zeroth coefficient (k_0) represents the mean of a formant trajectory multiplied by $\sqrt{2}$; the first coefficient (k_1) represents the direction and magnitude of the curve of the trajectory: a greater negative k_1

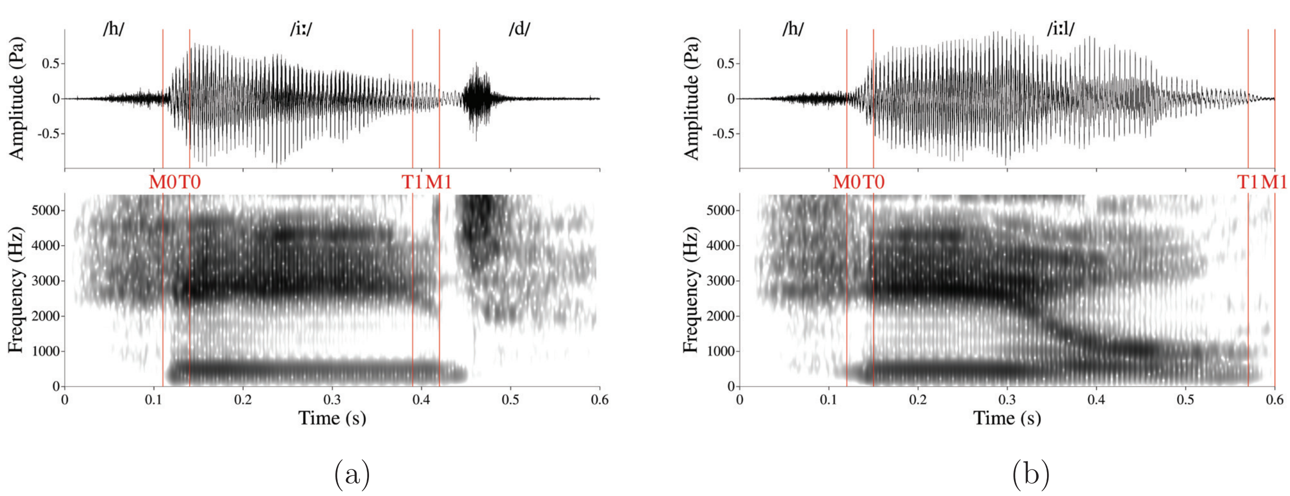


FIG. 2. (Color online) Acoustic landmarks defining the analysis window, exemplified by *heed* (a) and *heel* (b). *M*0: vowel onset determined by MAUS. *M*1: vowel offset (pre-/d/ context) and rime offset (pre-/l/ context) determined by MAUS. *T*0 marks the beginning and *T*1 marks the end of the analysis window.

https://doi.org/10.1121/10.0003499

corresponds to greater positive slope; the 2nd coefficient (k_2) represents the trajectory's curvature: a positive k_2 corresponds to an upward pointing curvature and a greater value corresponds to a narrow curvature (Harrington, 2010). Each token was represented parametrically by a total of 9 DCT coefficients (3 formants × 3 coefficients).

D. Statistical analysis

1. Effect of coda /l/ on monophthong targets

The effect of coda consonants on the acoustic targets of the monophthongs was examined using Linear Mixed-Effect models (LMMs) using the *LMER* function from the *LME4* package (Bates *et al.*, 2015), followed by least square means tests in the *EMMEANS* package (Lenth, 2019; Searle *et al.*, 1980) to evaluate the effect of /l/ on the mean target of each vowel adjusted for the means of other levels of factors in the LMM. We constructed three LMMs with the dependent variables $F1$, $F2$, and $F3$, and the interacting independent variables Vowel (sum-coded) and Coda (treatment coded, comparing /l/ to the baseline /d/); we used the factor Vowel rather than vowel features to test whether all vowels pattern consistently according to their place of articulation (front vs back, high vs low). The model included a random by-participant intercept and a by-participant random slope for the effect of coda to account for interspeaker variation. p-values were calculated with the *LMERTEST* package (Kuznetsova *et al.*, 2017) using Satterthwaite's degrees of freedom method. We constructed another three LMMs with the same structure, but without an interaction between Coda and Vowel to assess the effect of the Vowel-Coda interactions on model fit through model comparisons using a Chi-squared test. When adding Vowel-Coda interaction significantly improved model fit for $F1$, $F2$, and $F3$, least-square means analysis with Bonferroni correction was used to assess the effect of coda /l/ on the respective formant value of each vowel.

2. Spectral similarity

Spectral similarity across all diphthongs and monophthongs in the /d/- and /l/-context was tested by creating separate confusion matrices for pre-/d/ vowels and lateral-final rimes using random forest classification in the randomForest package (Liaw and Wiener, 2002). Random forest is a supervised classification algorithm that builds several decision trees and aggregates their result (Burger, 2018). Each decision tree splits the dataset (e.g., formant values of vowels) into subsets (e.g., back versus front vowels) based on descriptor values (e.g., high or low $F2$) (Burger, 2018). Building a random forest model consists of a training phase during which the algorithm learns the categories based on category labels (e.g., vowel labels) and descriptors (e.g., formant values, durational values) by building several binary decision trees (Burger, 2018). Then, in the testing phase, the remaining data is classified into the previously learnt categories based on descriptors only (Burger, 2018). Comparison of the original category labels and the category labels assigned by the random forest analysis provides a confusion matrix (Burger, 2018).

During the training phase, random forest classification builds several decision trees to learn the categories present in the data. Each tree is based on a bootstrap sample from the training data (customarily and in this paper 75% of the data) and random selection of descriptors. As training uses several bootstrap samples and different selection of descriptors, cross-validation is not required (Breiman, 2002). After a decision tree is built, the random forest classification makes a prediction, called out-of-bag prediction, about the data not in the bootstrap sample, based on the descriptors' values (Liaw and Wiener, 2002). After a pre-set number of trees has been built, out-of-bag predictions are aggregated: a low out-of-bag error rate indicates that the algorithm made successful predictions about the data left out in the iterations, and learnt the categories successfully, whereas a high out-of-bag error rate indicates that the algorithm could not make accurate predictions about the data left out from the iteration and was less successful in learning the categories (Liaw and Wiener, 2002).

Once the model is trained on a dataset, the second phase is the testing phase during which the model can be tested on the classification of novel data (customarily the remaining 25% of the original data), which are provided to the model without category labels. As a last step, the model's classification of the novel data is compared to the original category labels thus creating a confusion matrix between the original labels and the algorithm's labels, in which confusion rates indicate similarity between vowel categories.

To visualise the similarity between vowel categories and extract p-values, we ran a hierarchical cluster analysis on the confusion matrices output by the random forest analysis. Hierarchical cluster analysis takes the individual vowel categories as single-element clusters. At the first step, it merges two single-element clusters into a larger, binary-branching cluster. At each following step, it merges two clusters until it merges all the vowel categories into a single binary-branching cluster. Members within a cluster are maximally similar and the members of two separate clusters are maximally dissimilar; similarity was measured using Ward's method (Ward, 1963). To attest the robustness of clusters made of two or more vowel categories, we extracted the Approximately Unbiased p-value for each multi-element cluster by repeating the hierarchical cluster analysis on the same confusion matrices using multiscale bootstrap sampling in the *pvclust* package (Efron *et al.*, 1996; Suzuki and Shimodaira, 2006). Approximately unbiased p-value expresses the frequency with which a multi-element cluster appears in bootstrapping, and a multi-element cluster is considered to occur significantly frequently when it occurs in more than 95% of the resamples. The results of hierarchical cluster analysis are represented on a dendrogram: elements that are clustered together are similar to each other, and the lower the cluster is split from the other elements, the higher the spectral similarity between the members of the cluster. The location of nodes can be used for comparing between-cluster similarity across dendrograms.

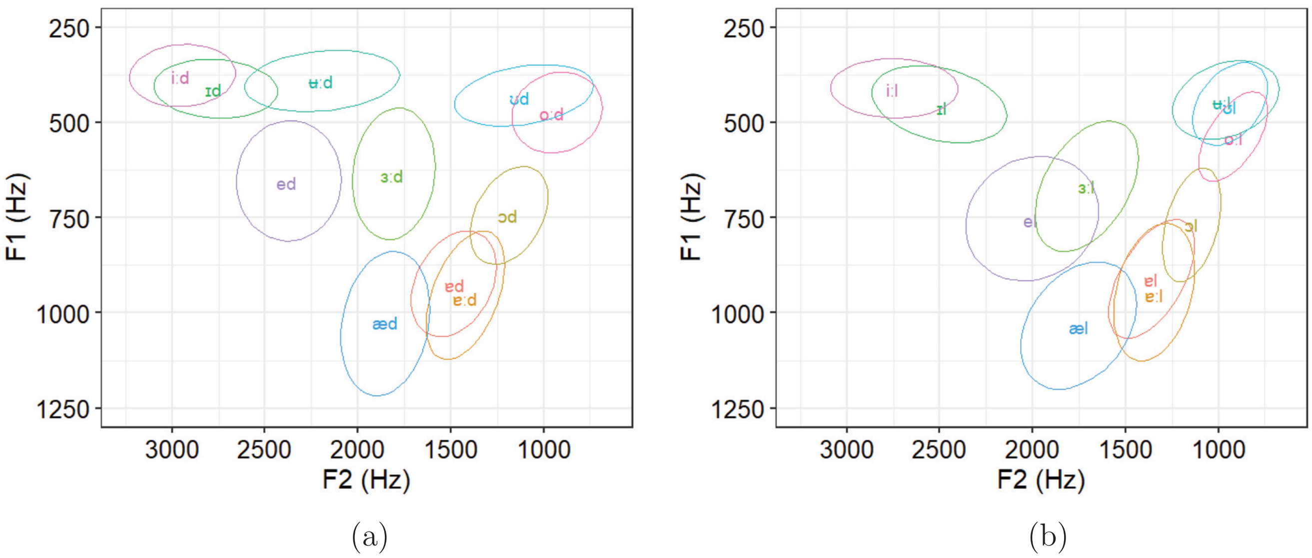


FIG. 3. (Color online) Acoustic monophthong targets produced before /d/ (a) and /l/ codas (b). IPA labels: mean $F1$ and $F2$ values (Hz). Ellipses: 95% confidence intervals.

To test spectral similarity in the /d/- and /l/-contexts, we first trained two random forest classification models to learn 16 vowel categories in the /d/-context and 16 vowel categories in the /l/-context based on the DCT coefficients, duration values, and vowel labels using 75% of the /d/-final and 75% of the /l/-final tokens. The remaining 25% of the tokens were used to test the classifier, by grouping unlabelled values based just on DCT coefficients and duration values. Separate confusion matrices were created by coda-condition. Last, the confusion matrices were fed into an agglomerative hierarchical cluster analysis using Ward's method (Ward, 1963) to measure between-vowel similarity based on the confusability rates of the vowels. All statistical analyses were carried out in R (R Core Team, 2018).

III. RESULTS

A. Effect of /l/ on the monophthong targets

We compared model fits between LMMs with and without Vowel-Coda interactions and found that models including the interactions fit the data significantly better for $F1$, $F2$, and $F3$ ($p < 0.001$ for model comparisons). Therefore, we report the main effect of /l/ from the models containing the interaction. Coda /l/ overall increases $F1$ ($\beta = 33.32, t_{28.01} = 11.41, p < 0.001$), decreases $F2$ ($\beta = -249.88, t_{28.01} = -28.93, p < 0.001$), and increases $F3$ ($\beta = 40.7, t_{28.01} = 4.77, p < 0.001$) (Figs. 3 and 4, Table II). Significant vowel-coda interactions are reported in Table III.

As the interactions significantly improved the model fit for all models, planned comparisons assessed the effect of coda /l/ on the $F1$, $F2$, and $F3$ of each vowel, using least-square means with Bonferroni correction (Table IV). Positive vowel-coda interactions (Table III) show that, compared to the overall effect, $F1$ increases more in the /l/-context for /e, ʉː, oː/. Similarly, the negative interactions show that $F1$ increases less in the pre-lateral context for /ɐː, ɐ/ than the overall effect (Table III). In line with the negative interactions, least-squares mean test shows no significant difference for the already low /ɐː, ɐ/ vowels (Table IV). In addition, least-square means test found no significant effect of coda /l/ on low /æ/ (Table IV). All other vowels show a significantly higher $F1$ in the /l/-context (Table IV).

Vowel-coda interactions in the initial LMM show that the $F2$ of /ɪ, e, ʉː/ is decreased before coda /l/ more than the overall effect, but $F2$ is lowered less than the overall effect for all other vowels (Table III). Least-square means test shows that even those vowels which showed a smaller effect for coda /l/ in the LMM had a significantly lower $F2$ in the /l/-context, except for back /oː/ (Table IV).

Vowel-coda interactions show that the $F3$ of /ɪ, ʉː, ɐ, ʊ, oː/ is increased before coda /l/ more than the overall effect, but

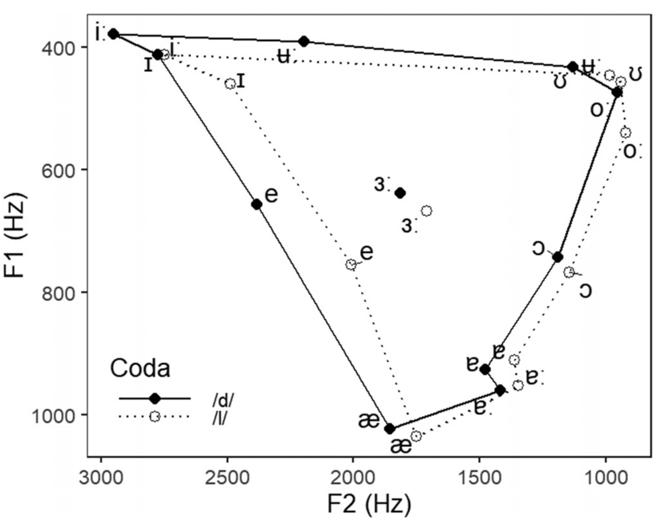


FIG. 4. Mean acoustic monophthong targets produced before /d/ and /l/ codas.

TABLE II. Mean formant values (Hz) and durations (ms) in hVd and hVl rimes.

Coda	Vowel	*F*1	*F*2	*F*3	Vowel duration
/d/	iː	379	2954	3329	298
	ɪ	413	2775	3255	177
	e	658	2382	3149	171
	æ	1023	1856	3005	209
	ʉː	391	2197	2684	295
	ɜː	638	1814	2886	307
	ɐː	961	1419	3034	329
	ɐ	927	1479	2995	158
	ʊ	433	1132	2882	175
	oː	475	953	3023	313
	ɔ	743	1191	2984	169
/l/	iː	413	2751	3204	424
	ɪ	460	2489	3075	365
	e	755	2011	3021	346
	æ	1036	1750	2987	395
	ʉː	446	983	3017	397
	ɜː	668	1711	2865	431
	ɐː	953	1347	3065	446
	ɐ	910	1360	3079	362
	ʊ	457	937	3131	375
	oː	540	920	3186	428
	ɔ	768	1146	3045	393

has a smaller than overall effect on /iː, e, ɜː/ (Table III). Least-square means test shows that *F*3 in the /l/-context was significantly lower for front vowels /iː, ɪ, e/, and significantly higher for /ʉː, ɐ, ʊ, oː/ (Table IV). Coda /l/ did not have a significant effect on the *F*3 of /æ, ɜː, ɐː, ɔ/. Therefore, least-mean square test does not show a consistent pattern on the effect of coda /l/ on *F*3.

TABLE III. Significant vowel-coda interactions in modelling the effect of coda /l/ on pre-lateral vowel targets compared to pre-/d/ vowel targets.

Parameter	Vowel	*β*	*df*	*t-value*	*p*-value
*F*1	e	64.2	7.93	8.1	<0.001
	ʉː	22.1	7.95	2.78	0.005
	ɐː	−40.9	7.93	−5.15	<0.001
	ɐ	−49.5	7.93	−6.25	<0.001
	oː	31.8	7.93	4.01	<0.001
*F*2	iː	47.2	1835.01	2.71	0.006
	ɪ	−37.1	1835.01	−2.13	0.034
	e	−121.7	1835.01	−6.98	<0.001
	ʉː	−964.1	1835.14	−55.12	<0.001
	ɜː	146.3	1835.01	8.39	<0.001
	ɐː	177.7	1835.01	10.19	<0.001
	ɐ	130.4	1835.01	7.48	<0.001
	ʊ	55.7	1835.01	3.19	0.001
	oː	216.6	1835.01	12.42	<0.001
	ɔ	204.6	1835.01	11.73	<0.001
*F*3	iː	−165.6	1835.01	−8.3	<0.001
	ɪ	220.6	1835.01	−11.05	<0.001
	e	−168.8	1835.01	−8.46	<0.001
	ʉː	291.5	1835.18	14.56	<0.001
	ɜː	−61.6	1835.01	−3.09	0.002
	ɐ	43.8	1835.01	2.20	0.028
	ʊ	207.4	1835.01	10.40	<0.001
	oː	122.9	1835.01	6.16	<0.001

TABLE IV. Effect of coda /l/ on *F*1, *F*2, and *F*3 values (Hz) at acoustic target compared to coda /d/. *β* shows the effect of coda /l/ compared to coda /d/ on the least-square mean of the vowel formant. Standard error (*SE*), *t-ratio*, and *p*-value calculated from least square means.

Parameter	Vowel	*β*	*SE*	*t-ratio*	*p*-value
*F*1	iː	33.8	8.45	3.999	0.0007
	ɪ	47.0	8.45	5.557	<0.0001
	e	97.5	8.45	11.539	<0.0001
	æ	12.8	8.45	1.518	1
	ʉː	55.4	8.48	6.538	<0.0001
	ɜː	29.3	8.45	3.471	0.0059
	ɐː	−7.6	8.45	−0.895	1.0
	ɐ	−16.2	8.45	−1.920	0.6064
	ʊ	24.0	8.45	2.835	0.0553
	oː	65.1	8.45	7.704	<0.0001
	ɔ	25.3	8.45	2.990	0.0343
*F*2	iː	−202.7	19.5	−10.412	<0.0001
	ɪ	−286.9	19.5	−14.742	<0.0001
	e	−371.6	19.5	−19.091	<0.0001
	æ	−105.5	19.5	−5.419	<0.0001
	ʉː	−1214.0	19.5	−62.207	<0.0001
	ɜː	−103.6	19.5	−5.321	<0.0001
	ɐː	−72.2	19.5	−3.709	0.0025
	ɐ	−119.4	19.5	−6.136	<0.0001
	ʊ	−194.2	19.5	−9.978	<0.0001
	oː	−33.3	19.5	−1.711	0.963
	ɔ	−45.3	19.5	−2.326	0.2239
*F*3	iː	−124.9	21.7	−5.755	<0.0001
	ɪ	−179.8	21.7	−8.285	<0.0001
	e	−128.0	21.7	−5.899	<0.0001
	æ	−18.5	21.7	−0.852	1
	ʉː	332.2	21.8	15.261	<0.0001
	ɜː	−20.9	21.7	0.962	1
	ɐː	30.9	21.7	1.422	1
	ɐ	84.6	21.7	3.895	0.0012
	ʊ	248.2	21.7	11.432	<0.0001
	oː	163.7	21.7	7.539	<0.0001
	ɔ	60.8	21.7	2.801	0.0574

The duration of all short vowels was 57% of the long-vowel duration in the /d/ condition, and the duration of all rimes containing short vowels was 88% of rimes containing long vowels in the /l/ condition.

B. Spectral similarity

Formant trajectories for all vowels were modelled using the first three DCT coefficients (see Tables VI and VII in the Appendix). Two random forest classification models were trained on DCT coefficients, duration values, and vowel labels using 75% of the tokens to learn 16 vowel categories in each coda condition. Out-of-bag error rate in the testing phase was 3.55% in the /d/-context and 24.07% in the /l/-context, indicating that DCT coefficients and duration values can classify vowels more accurately in the /d/- than in the /l/-context.

Twenty-five percent of the tokens were used to test the classification algorithms; the output of the random forest classification algorithm was compared to the original vowel labels, resulting in two confusion matrices (Figs. 5 and 6). In the /d/-context, seven vowels were classified with 100% accuracy (/ɪ, ʉː, ʊ, æɪ, ɔɪ, æɔ, əʉ/), whereas in the /l/-context only the rime /el/ was classified perfectly. In the /d/-context error rates were small: the least accurately classified vowels were central /ɜː/ and back /ɔ/, identified with, respectively, 83% and 85% accuracy.

The pre-lateral rime pairs /uːl-ʊl, əʉl-ɔl, æɔl-æl/, whose members were hypothesised to undergo acoustic contrast reduction, have a high confusion rate (Fig. 6): 26% of /ʉːl/ tokens were classified as /ʊl/ and 28% of /ʊl/ tokens were classified as /ʉːl/; 43% of /əʉl/ tokens were classified as /ɔl/ and 16% of /ɔl/ tokens were classified as /əʉl/; 30% of /æɔl/ tokens were classified as /æl/ and 30% of /æl/ tokens were classified as /æɔl/. In contrast, all of the /ʉː, ʊ, æɔ, əʉ/ tokens were identified correctly in the /d/-context, /æ/ was confused with /ɐ/ (9%), not with /æɔ/, and /ɔ/ was misidentified as /ʊ/ (12%) and not as /əʉ/. Members of the pre-lateral pair /iːl-ɪl/ were also hypothesised to undergo spectral contrast reduction and the confusion rate between /iːl/ and /ɪl/ is higher in the /l/-context (19% of /iːl/ tokens misidentified as /ɪl/ while 5% of /ɪl/ tokens misidentified as /iːl/, without any confusion in the other direction) than in the /d/-context (5% of /iː/ tokens identified as /ɪ/). Despite the notable confusion between /iː/ and /ɪ/ in the /l/ context, it is smaller than for the other three vowel pairs that are confusable in this context.

We used hierarchical cluster analysis to test whether the patterns of confusion correspond to statistically significant

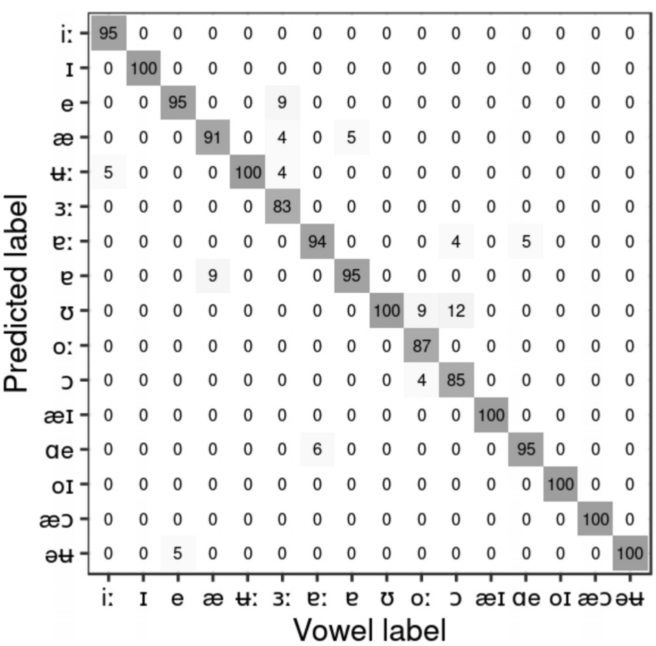


FIG. 5. Confusion matrix of vowels produced before /d/ codas, based on DCT coefficients ($k0$, $k1$, $k2$) of formants ($F1$, $F2$, $F3$) and mean vowel duration. Columns show the percentage of tokens classified for each vowel target. Rows show the percentage of tokens classified by the random forest classification algorithm as a certain vowel.

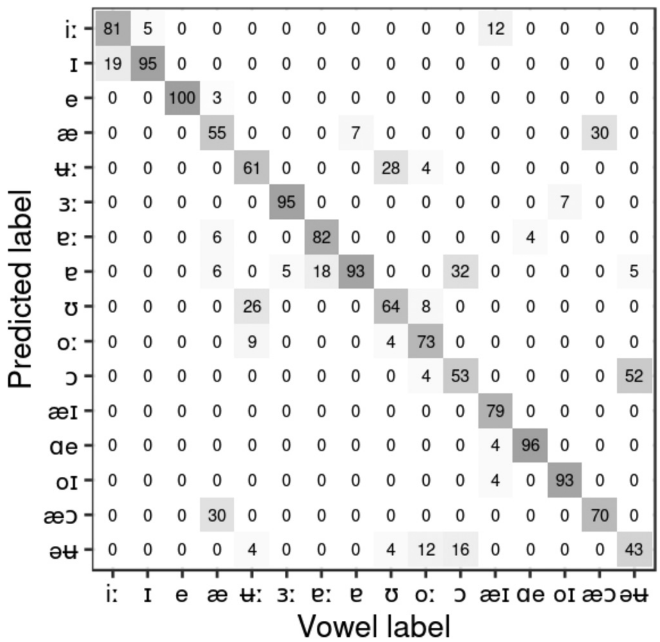


FIG. 6. Confusion matrix of vowels produced before /l/ codas, based on DCT coefficients ($k0$, $k1$, $k2$) of formants ($F1$, $F2$, $F3$) and mean vowel duration. Columns show the percentage of tokens classified for each vowel target. Rows show the percentage of tokens classified by the random forest classification algorithm as a certain vowel.

contrast reduction between AusE vowels. In the /d/-context, the only cluster that appears with significant frequency, namely in 100% of the bootstrap samples, was the cluster consisting of all diphthongs and monophthong vowels except the three back monophthongs /oː, ʊ, ɔ/ (Fig. 7). In the /d/-context no vowel pairs are confused with significant frequency; that is no two such vowels are found which are maximally similar to each other and maximally different from the rest (Fig. 7). In the /l/-context, the cluster of /iːl, ɪl, oːl, oɪl, ɜːl, el, æɪl, ɑel/ occurs with significant frequency. In addition, the pairs /iːl-ɪl, ʉːl-ʊl, əʉl-ɔl, æɔl-æl/ occur significantly frequently in a cluster, indicating that the members of these pairs are maximally similar to each other (Fig. 7).

Vertical location of the nodes of the dendrograms (Fig. 7) indicates similarity between the clusters: the lower a node is located, the more similar the members of the cluster are. The close vertical alignment of the nodes in Fig. 7 in the /d/ condition shows that members of the vowel dyads are only slightly more similar to each other than to other dyads. For instance, /ʊ/ and /ɔ/ merge into a cluster at approximately 1.4, and the /ʊ-ɔ/ cluster merges with /oː/ at approximately 1.5, indicating that /ʊ/ and /ɔ/ are only a little more similar to each other than the /ʊ-ɔ/ cluster is to /oː/ (Fig. 7). In contrast, in the /l/ condition, members within the vowel pairs /ʉːl-ʊl, əʉl-ɔl, æɔl-æl/ are maximally similar to each other, as the nodes of their respective dyads branch at 0.5 (/əʉl-ɔl/) or below (/ʉːl-ʊl, æɔl-æl/) (Fig. 7).

Both random forest analysis and hierarchical cluster analysis indicate that spectral contrast is reduced between the members of the pairs /iːl-ɪl, ʉːl-ʊl, əʉl-ɔl, æɔl-æl/. In the random forest analysis, the members of these pairs are

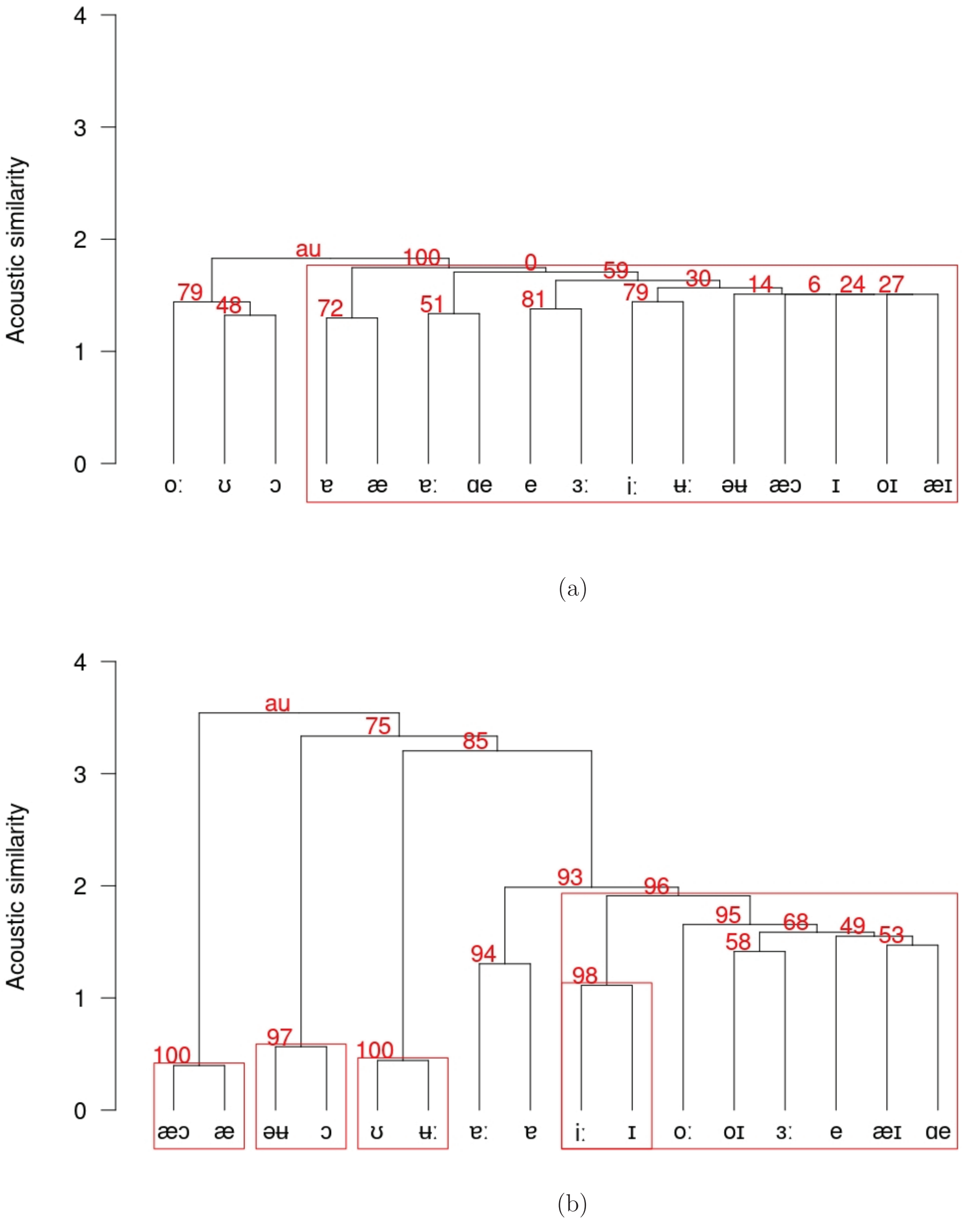


FIG. 7. (Color online) (a) Acoustic vowel similarity before /d/ codas, based on vowel confusion; (b) acoustic rime similarity in /l/-final rimes, based on rime confusion. Lower branching signals higher confusion rates. AU, Approximately unbiased *p*-value indicates the frequency with which a cluster appears in bootstrapping. Red boxes highlight clusters appearing with significant frequency.

systematically confused. In the hierarchical cluster analysis, these pairs form significantly frequently recurring dyads that are maximally similar to each other in the pre-lateral vowel space.

/iːl/ and /ɪl/ show a lower confusion rate in the random forest analysis compared to the other three key vowel pairs, and they are also merged later in hierarchical cluster analysis. For the pairs /ʉːl-ʊl, əʉl-ɔl, æɔl-æl/, random forest provides more details than hierarchical cluster analysis. Random forest analysis shows that /ʉːl/ is primarily confused with /ʊl/ and to a lesser extent with /oːl/ (/ʉːl/ and /ʊl/ are confused in almost 30% of the tokens for both rimes, and /ʊl/ and /oːl/ are confused in 4% of the tokens for both rimes). The high confusion rate between /ʉːl/ and /ʊl/ leads to these vowels forming a dyad in hierarchical cluster analysis, while the smaller confusion rate between /ʉːl, ʊl/, and /oːl/ is not captured by hierarchical cluster analysis. Similarly, random forest misidentifies /ɔl/ as /ɐl/ (32%) and as /əʉl/ (16%) and misidentifies /æl/ as /ɐːl, ɐl/ (6%, –6%) and as /æɔl/ (30%). However, in the hierarchical cluster analysis /ɔl/ clusters with /əʉl/, not /ɐl/ due to 52% of /əʉl/ tokens being misidentified as /ɔl/, while /æl/ clusters with /æɔl/, not /ɐːl, ɐl/ due to 30% of /æɔl/ tokens being misidentified as /æl/.

IV. SUMMARY OF RESULTS

(1) Effect of coda /l/ on monophthong targets compared to coda /d/:
   (a) All vowels have a higher $F1$, except for /æ, ɐː, ɐ, ʊ/, indicating phonetic lowering before coda /l/.
   (b) All vowels have a lower $F2$, except for /oː/ and /ɔ/, indicating phonetic backing before coda /l/.
   (c) Front vowels /iː, ɪ, e/ have lower $F3$ before coda /l/, while central and back /ʉː, ɐ, ʊ, oː/ have higher $F3$.

(2) Spectral contrast reduction:

(a) Increased out-of-bag error rate in random forest analysis indicates that a higher percentage of vowels were misidentified in the /l/-context than in the /d/-context.
(b) Random forest analysis indicates that confusion of pre-/l/ vowels is pairwise and systematic; such patterns were rarely observed in the pre-d context.
(c) Hierarchical cluster analysis shows that the members of the lateral-final pairs /iːl-ɪl, ʉːl-ʊl, əʉl-ɔl, æɔl-æl/ are maximally similar to each other; no such pairings were found among the pre-/d/ vowels.

V. DISCUSSION

A. Acoustic patterns

1. Lowering of monophthongs

Hypothesis (1) predicted that front vowels would have a higher *F*1, that is, they would be phonetically lowered in pre-lateral position compared to pre-/d/ position. Hypothesis (1) largely holds, as we found increased *F*1 for all front vowels (/iː, ɪ, e/) except for front /æ/. In addition, most back vowels were also found to be significantly lowered in pre-lateral contexts. The biggest lowering effect can be observed in /e/, whose target distribution shifts toward /æ/, similar to shifts observed in Melbourne/Victoria dialects (Cox and Palethorpe, 2004; Loakes *et al.*, 2010c). However, random forest and hierarchical cluster analysis did not classify /e/ as similar or confusable with /æ/ in the /l/-context, most probably due to the lack of overlap between pre-lateral /e/ and /æ/.

The only front vowel that did not lower before laterals was /æ/, which can potentially be explained by its already high *F*1 in the /d/ condition. The low vowels /ɐː/ and /ɐ/ did not lower either, similar to the observation of Bernard (1985) and Palethorpe and Cox (2003). The lack of phonetic lowering in /æ, ɐː, ɐ/ indicates that /æ/ might pattern with the phonologically low vowels due to its high *F*1. This pattern appears again as pre-/d/ /æ/ and /ɐ/ are classified as similar (Fig. 7).

2. Backing of monophthongs

Hypothesis (2) predicted that front vowels would have a lower *F*2, that is, they would be phonetically backed in pre-lateral position compared to pre-/d/ position. Hypothesis (2) holds, as we found decreased *F*2 for all front vowels before coda /l/, compared to coda /d/. In addition, back and low vowels were also phonetically backed except for /oː/ and /ɔ/.

The greatest backing effect was observed for /ʉː/, whose target *F*2 is on average 1214 Hz lower before coda /l/ than before coda /d/. As a result, /ʉː/ overlaps acoustically with /ʊ/ in the /l/-context, unlike in the pre-/d/ context, where it acoustically approaches /ɪ/ (Fig. 3). The backing influence of the lateral on /ʉː/ is further corroborated in the analysis of spectral similarity: in the /l/-context /ʉː/ shows similarity to /ʊ/ and to a lesser extent to long back /oː/. In contrast, in the /d/-context /ʉː/ shows some similarity to front /iː/ and central /ɜː/. The fact that /ʉː/ shows similarity to /iː/ and not to /ɪ/, even though the latter is acoustically closer to /ʉː/ in the *F*1–*F*2 vowel space, is due to the fact that the presented analysis of spectral similarity considers vowel length when classifying vowels. Therefore, in the /d/-context, long vowels are clustered with long vowels, but in the /l/-context long-short vowel pairs cluster together due to the reduction of the duration contrast.

In addition, we found that /e/ partially overlaps acoustically with /ɜː/ in the pre-lateral environment due to the lowering of its *F*2. However, we did not find spectral contrast reduction between /e/ and /ɜː/.

3. Acoustic contrast reduction

Hypothesis (3) predicted that acoustic contrast would be reduced between the members of the pairs /iːl-ɪl, ʉːl-ʊl, əʉl-ɔl, æɔl-æl/. Analysis of spectral similarity shows that acoustic vowel contrast is reduced between the members of these pairs, as the members of each pair are maximally similar to each other.

The pairwise similarity of long-short vowels in the /l/ context was absent in the /d/ context because the model for testing similarity included duration as a distinguishing cue. In the /d/-context, vowel similarity within members of any cluster and between members of separate clusters is comparable (Fig. 7).

Increased spectral similarity in the /l/-context compared to the /d/ condition could be due to the fact that formant trajectories were measured in the rime, and thus all include /l/. However, if the overlap in the coda was the main cause of the increased confusion rates, all rimes would be confused to the same extent. That is, the dendrogram would be similar to that of the /d/-context, as it would show comparable similarity within members of clusters as between members of different clusters and would not show the pairwise similarity of key vowel pairs. Therefore, the dendrogram in the /l/-context indicates that the increased confusion rates are due to contrast reduction in the vocalic part of the rime.

/iː/ and /ɪ/ are spectrally more similar to each other than to any other vowel in the /l/-context; however, the extent of spectral similarity is smaller between the members of the pair /iːl-ɪl/ than the members of the pairs /ʉːl-ʊl, əʉl-ɔl, æɔl-æl/. These tentative results are best explained by the fact that both vowels are backed and lowered to a similar extent, with pre-lateral /iː/ remaining more peripheral than pre-lateral /ɪ/ (Figs. 3 and 4). In addition, pre-lateral /iː/ and /ɪ/ might be differentiated by the presence of onglide in /iː/ (Cox *et al.*, 2014). The high front target is followed by a steep *F*2 transition to /l/ (Fig. 8).

Increased spectral similarity between pre-lateral /ʉː/ and /ʊ/ is attributed to the *F*2 drop in pre-lateral /ʉː/ throughout the vowel, which makes high central /ʉː/ similar to high back /ʊ/ in the pre-lateral context (Figs. 4 and 8). Not only is the vowel target backed [see Sec. V A 2, Figs. 3(a) and 4], but the entire *F*2 trajectory is low across the rime in the /l/-context (Fig. 8).

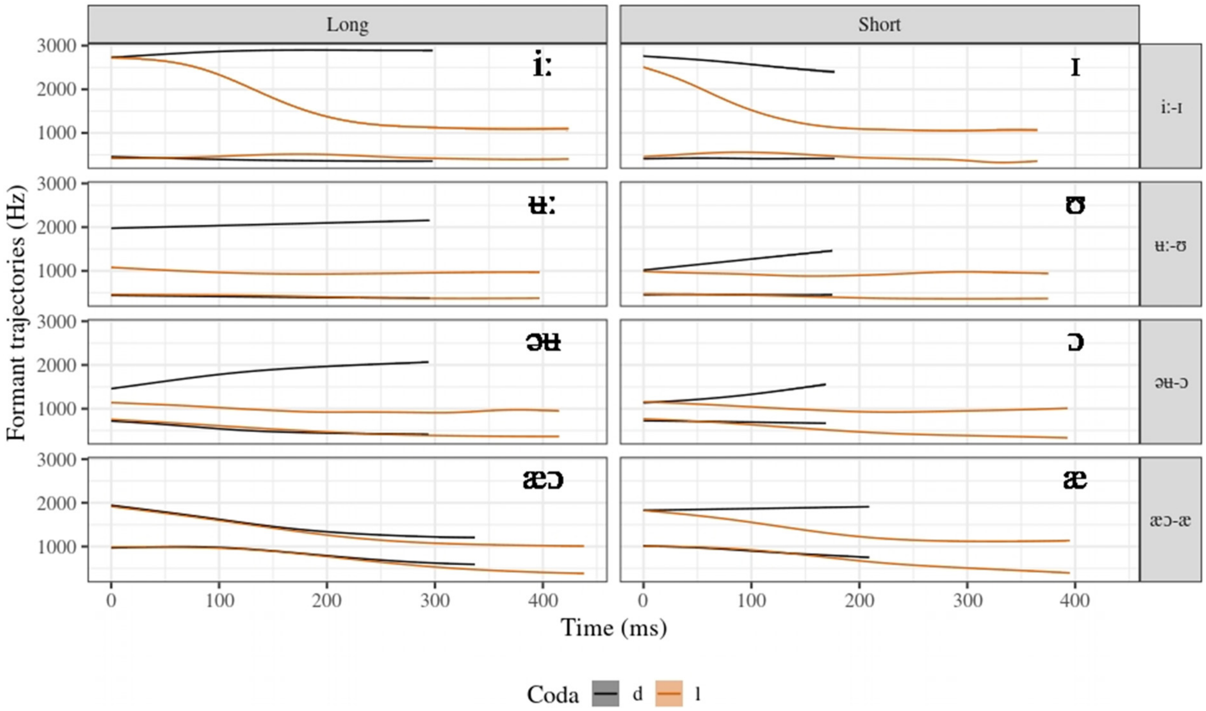


FIG. 8. (Color online) Mean $F1$ and $F2$ trajectories by coda context (black: /d/, red: /l/) and vowel pair. First row: /iː-ɪ/. Second row: /ʉː-ʊ/. Third row: /əʉ-ɔ/. Fourth row: /æɔ-æ/. Left: long vowel. Right: short vowel.

Increased spectral similarity between /əʉ/ and /ɔ/ is best explained by the diphthong's lowering and backing of the first target and the loss of the high central second target, shown by the overall lower $F2$ trajectory in the /l/-context (Fig. 8). As the high central second target of the pre-lateral diphthong is backed, it becomes similar to mid-back /ɔ/ (Fig. 8). In contrast, $F2$ of /əʉ/ in the /d/-context shows a higher first target followed by a steep rise as it transitions from the schwa target to the [ʉː] target.

Increased spectral similarity between pre-lateral /æɔ-æ/ is best explained by the fact that the $F2$ trajectory of /æl/ becomes similar to that of /æɔl/ (Fig. 8). /æɔ/ has a falling $F2$ both in the /d/- and in the /l/-context, as is expected in both conditions as the diphthong in the /d/-context and the rime in the /l/-context contains a transition from a high $F2$ to a low $F2$. In contrast, /æ/ has a rising $F2$ in the /d/-context due to the vowel-alveolar transition (Delattre *et al.*, 1955), whereas /æ/ has a falling $F2$ in the /l/-context due to the vowel-/l/ transition, making the $F2$ trajectory more similar to /æɔl/ (Fig. 8).

The vowel pairs /iː-ɪ, ʉː-ʊ, əʉ-ɔ, æɔ-æ/ also contrast in terms of length. In the /d/-context, duration of the short key vowels is 59% of the long key vowels, in line with Cox (2006), and mean duration of key rimes with short vowels is 79% of the duration of key rimes containing long vowels (Table V). In contrast, the key /l/-final rimes containing short vowels are 91% of key /l/-final rimes containing long vowels (Table V). Reduced duration contrast in the /l/-context further increases similarity between key long-short vowel pairs, whereas the larger duration contrast in the /d/-context results in vowels being classified according to length (Fig. 7). However, duration contrast reduction between the /d/- and the /l/-context cannot be assessed without separating the vowel from the following liquid for which we have found no reliable method.

The acoustic targets and the durations of pre-/d/ vowels in the current study are consistent with standard descriptions of AusE (Cox, 1999, 2006; Cox and Palethorpe, 2001). In addition, the pairing of /ʉː/ with /iː/ in the cluster analysis of the /d/-condition is in line with the fronting of the AusE /ʉː/ (Cox, 1999; Cox and Palethorpe, 2001; Elvin *et al.*, 2016; Harrington *et al.*, 1997). Our results confirm the increased acoustic similarity between /iː-ɪ, ʉː-ʊ, əʉ-ɔ, æɔ-æ/ in the pre-lateral context noted by Palethorpe and Cox (2003).

TABLE V. Duration contrast reduction from pre-/d/ long and short vowels to /l/-final rimes containing long and short vowels.

Context	Vowel pair	Long (ms)	Short (ms)	Short:Long
Pre-/d/ vowels	/iː-ɪ/	298	177	0.59
	/ʉː-ʊ/	295	175	0.59
	/əʉ-ɔ/	294	169	0.57
	/æɔ-æ/	337	209	0.62
	Mean	306	183	0.59
/d/-final rimes	/iː-ɪ/	397	317	0.80
	/ʉː-ʊ/	398	316	0.80
	/æɔ-æ/	429	348	0.81
	/əʉ-ɔ/	415	320	0.77
	Mean	409	325	0.79
/l/-final rimes	/iː-ɪ/	424	365	0.86
	/ʉː-ʊ/	396	375	0.95
	/æɔ-æ/	438	395	0.90
	/əʉ-ɔ/	415	393	0.95
	Mean	418	382	0.91

B. Articulatory explanations

The phonetic backing and lowering of pre-lateral vowels can be attributed to the coarticulatory influence of the dorsal gesture of /l/ on the preceding vowel, as has been reported for American English (Gick *et al.*, 2002; Gick and Wilson, 2006; Giles and Moll, 1975; Sproat and Fujimura, 1993). In American English, tongue dorsum lowering and retraction typically precedes coronal articulation in coda laterals and may overlap with the vowel (Giles and Moll, 1975; Proctor *et al.*, 2019; Sproat and Fujimura, 1993). The overall increase in $F1$ and overall decrease in $F2$ observed in AusE pre-lateral vowels is consistent with a pattern of production in which the lowered and retracted tongue dorsum gesture of coda /l/ coarticulates with the vowel gesture (Fant, 1960). In particular, the phonetic backing of /ʉː/ observed here is consistent with the articulatory backing of this vowel observed in previous work for AusE (Lin *et al.*, 2012). The backed tongue position in the production of pre-lateral /ʉː/ might make it articulatorily similar to /ʊ/. Similarly to AusE, in Standard Southern British English and West Yorkshire English, the lateral-final rimes in *fool* and *full* show acoustic and articulatory contrast reduction compared to the pre-obstruent vowels in *food* and *foot* (Gorman and Kirkham, 2020). Contrast reduction occurs due to $F2$ lowering in both *fool* and *full* and to tongue dorsum backing in *fool*, despite the tongue dorsum fronting in *full* (Gorman and Kirkham, 2020).

The reduction in acoustic contrast between /æɔ-æ/ before laterals in the AusE data is also consistent with the articulatory characterization of the dorsal gesture associated with American English laterals: a magnetic resonance imaging (MRI) study of [ɫ] and /ɔ/ reported articulatory similarities between the dorsal gestures of [ɫ] and /ɔ/ (Gick *et al.*, 2002). As a result, the monophthong /æ/ followed by an /ɔ/-like /l/ can be spectrally similar to the diphthong /æɔ/, whereas the second target of the diphthong /æɔ/ might be encroached upon by the following /ɔ/-like /l/.

Articulatory similarity between /ɔ/ and /l/ can potentially also play a role in the loss of the second target of /əʉː/, as the backed [ʉ] can be similar to /ɔ/ and therefore to /l/, leading to the loss of contrast between the second target of the diphthong and /l/. This account is consistent with the articulatory backing of the second target of /əʉ/ in the pre-/l/ context (Lin *et al.*, 2012).

When coda /l/ is preceded by a high front vowel, the vowel and /l/ place competing demands on the tongue dorsum: the vowel target requires a raised and fronted tongue dorsum whereas the /l/ target requires it to be lowered and backed (Gick and Wilson, 2006). These competing demands result in a long transition between the two segments during which the tongue passes through a schwa-like posture (Gick and Wilson, 2006). Our acoustic data from AusE are consistent with these articulatory accounts of American English, as /iː/ and /ɪ/ exhibited a relatively front target followed by a long steep $F2$ fall to reach the /l/ target.

Although the observed $F1$-raising and $F2$-lowering can be attributed to tongue raising and backing (Fant, 1960), this well-established relationship between tongue lowering and $F1$ and between tongue backing and $F2$ might break down in the /l/-context (Strycharczuk and Scobbie, 2017). For example, in Standard Southern British English, $F2$ difference between /uː/ and /ʊ/ is reduced, similarly to AusE; however, articulatory distinctions are maintained (Strycharczuk and Scobbie, 2017). Therefore, an articulatory study is needed to address vowel-lateral coarticulation and articulatory contrast reduction in AusE lateral-final rimes. Such research would reveal whether there is articulatory contrast reduction in the tongue body gesture between the members of the lateral-final pairs /iːl-ɪl, ʉːl-ʊl, əʉl-ɔl, æɔl-æl/ in AusE; or whether spectral contrast reduction is better attributed to changes in lip-rounding; or to different timing relations between the /l/ and /d/-context in the coordination of the vowel and the coda gesture.

C. Implications for sound change: Pre-lateral vowel mergers?

A vowel merger is defined as the loss of contrast between two or more categories due to the loss of phonetic differentiation either across the board or in a particular phonological environment (Maguire *et al.*, 2013). In the Harrington *et al.* (2018) interactive phonetic model of sound change, the prerequisite of sound change is that typical realisations of two phonemes are acoustically distinct, but their highly coarticulated realisations become acoustically similar to each other. As listeners and speakers interact, atypical speaker realisations are incorporated into the listener's phoneme representation, shifting its boundary closer to the second phoneme until the categories overlap, potentially leading to a merger (Harrington *et al.*, 2018).

Acoustic contrast reduction within the pairs /ʉː-ʊ, əʉ-ɔ, æɔ-æ/ in pre-lateral environments is consistent with the interactive phonetic model of sound change and with a contextual vowel merger conditioned by coda /l/. Vowel-lateral coarticulation creates atypical realisations for these vowels, shifting their boundaries closer to each other and leading to overlap. This is best exemplified by the vowel /ʉː/: /ʉː/ moves into the vowel plane of /ʊ/ (Fig. 3), making pre-lateral /ʉː/ a potential candidate for a vowel merger with pre-lateral /ʊ/ in the New South Wales dialect of AusE. In addition, the perceptual confusion between /ʉː-ʊ, əʉ-ɔ, æɔ-æ/ provide further support for a potential perceptual merger, as listeners are not always able to distinguish pre-lateral vowels on the basis of spectral and durational cues (Szalay *et al.*, 2018).

While our analysis of spectral similarity indicates that contrast is reduced even considering dynamic $F1$, $F2$, $F3$, and duration information, our methods cannot show whether the phonemes are differentiated: both Random Forest and Hierarchical Cluster Analysis classified the tokens into pre-defined 16 vowel categories. Increased similarity between categories is consistent both with a merger and with reduced acoustic contrast. To explore whether the phonemes undergo a conditional merger in the pre-lateral environment, an apparent time or a sociolinguistic study is needed to better understand the implications for the actuation of sound change in key pre-lateral vowels of AusE.

VI. CONCLUSION

In AusE, $F1$ is increased and $F2$ is decreased in the acoustic target of pre-lateral vowels compared to coda /d/, indicating phonetic lowering and retraction. In addition, spectral and durational contrast is reduced within the pairs /iːl-ɪl, ʉːl-ʊl, əʉl-ɔl/, and /æɔl-æl/ (e.g., *feel-fill, fool-full, role-roll, howl-Hal*). Spectral contrast reduction is potentially the result of the coarticulatory effect of the dorsal gesture of /l/ reported in other varieties of English. The observed spectral contrast reduction may reflect necessary conditions for conditional vowel mergers in the pre-lateral environment.

ACKNOWLEDGMENTS

This project was approved by the Macquarie University Human Research Ethics Committee under HE26SEP2008-R06061L&P. This research was supported in part by iMQRTP 2015144, ARC DE150100318, and MQSIS 9201501719 grants. We thank Peter Humburg for his help in the statistical analysis, and the members of the Phonetics Lab, the Centre for Language Sciences and the ARC Centre of Excellence in Cognition and its Disorders at Macquarie University for their comments, feedback, and support.

APPENDIX: DCT COEFFICIENTS OF /D/- AND /L/-FINAL RIMES

TABLE VI. DCT coefficients (k_0, k_1, k_2) of formants ($F1$, $F2$, $F3$) in pre-obstruent vowels.

	$F1$			$F2$			$F3$		
Vowel	k_0	k_1	k_2	k_0	k_1	k_2	k_0	k_1	k_2
iː	563	41	10	4016	−59	−47	4634	−39	−20
ɪ	593	−6	−3	3786	81	−25	4538	43	6
e	915	31	−8	3274	30	−8	4440	−9	3
æ	1330	88	−30	2629	−41	26	4287	−45	19
ʉː	580	25	−1	2887	−59	3	3717	−16	32
ɜː	889	21	−13	2574	−41	11	4096	−28	15
ɐː	1265	42	−18	2035	−89	68	4302	−27	−11
ɐ	1238	55	−18	2193	−104	27	4260	−26	2
ʊ	649	−4	−4	1635	−166	53	4065	44	−14
oː	730	−15	9	1374	−113	88	4280	34	−55
ɔ	1002	16	−7	1723	−104	44	4201	27	−19
æɪ	812	206	47	3579	−262	−65	4414	−54	−18
ɑe	1251	95	−79	2398	−438	99	4224	−18	53
oɪ	745	74	−41	2545	−867	−9	4155	−25	101
æɔ	1257	155	−78	2192	341	15	4199	−56	31
əʉ	780	136	21	2507	−243	−30	3743	52	28

TABLE VII. DCT coefficients (k_0, k_1, k_2) of formants ($F1$, $F2$, $F3$) in /l/-final rimes.

	$F1$			$F2$			$F3$		
Vowel	k_0	k_1	k_2	k_0	k_1	k_2	k_0	k_1	k_2
iː	642	12	−75	2527	915	134	4501	−154	182
ɪ	677	73	−62	2104	624	267	4523	−236	32
e	851	203	−32	1907	423	155	4622	−246	2
æ	1087	295	−70	1980	366	36	4552	−276	22
ʉː	590	48	−16	1370	43	55	4430	−186	10
ɜː	835	130	−73	1925	332	14	4367	−308	80
ɐː	1060	233	−76	1740	138	−25	4568	−226	37
ɐ	897	252	6	1663	150	47	4683	−173	−51
ʊ	587	63	−10	1316	5	52	4535	−126	−12
oː	697	77	−31	1308	−4	32	4655	−163	19
ɔ	778	194	5	1433	86	60	4634	−193	−24
æɪ	826	191	−1	2393	698	−64	4387	−144	133
ɑe	1075	221	−45	2124	147	−339	4304	−172	93
oɪ	744	94	−30	2017	299	−289	4270	−244	141
æɔ	1099	286	−81	1960	410	72	4451	−288	94
əʉ	748	177	18	1398	86	64	4580	−201	3

[1]The symbols used are those recommended by Cox and Palethorpe (2007) for AusE.

Altendorf, U., and Watt, D. (**2008**). "The dialects in the South of England: Phonology," in *Varieties of English*, edited by B. Kortmann and C. Upton (Mouton de Gruyter, Berlin, New York), pp. 194–222.

Arnold, L. (**2015**). "Multiple mergers: Production and perception of three pre-/l/ mergers in Youngstown, Ohio," Univ. Penn. Work. Papers Ling. **21**(2), 1–10.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Softw. **67**(1), 107820.

Bernard, J. R. (**1970**). "Toward the acoustic specification of Australian English," STUF Lang. Typol. Univ. **23**(1-6), 113–128.

Bernard, J. (**1985**). "Some local effects of post-vocalic (l)," in *The Cultivated Australian: Festschrift in Honour of Arthur Delbridge*, edited by J. E. Clark (Buske, Hamburg), pp. 319–332.

Blevins, J. (**2006**). "A theoretical synopsis of evolutionary phonology," Theor. Ling. **32**, 117–165.

Boersma, P., and Weenink, D. (**2013**). "Praat: Doing phonetics by computer (version 5.3.42) [computer program]," http://www.praat.org (Last viewed 11 February 2021).

Bradley, D. (**2004**). "Regional characteristics of Australian English: Phonology," in *Varieties of English: The Pacific and Australasia*, edited by K. Burridge and B. Kortmann (Mouton de Gruyter, Berlin), pp. 645–655.

Breiman, L. (**2002**). "Manual on setting up, using, and understanding random forests v3.1," Statistics Department University of California Berkeley, https://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf (Last viewed 11 February 2021).

Burger, S. V. (**2018**). *Introduction to Machine Learning with R: Rigorous Mathematical Analysis* (O'Reilly Media, Inc., Sebastopol, CA).

Butcher, A. (**2006**). "Formant frequencies of /hVd/ vowels in the speech of South Australian females," in *Proceedings of the 11th Australian International Conference on Speech Science and Technology*, December 6–8, Canberra, Australia, pp. 449–453.

Cox, F. (**1999**). "Vowel change in Australian English," Phonetica **56**, 1–27.

Cox, F. (**2006**). "The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers," Aust. J. Ling. **26**(2), 147–179.

Cox, F., and Fletcher, J. (**2017**). *Australian English Pronunciation and Transcription* (Cambridge University Press, Cambridge).

Cox, F., and Palethorpe, S. (**2001**). "The changing face of australian english vowels," in *English in Australia*, edited by D. Blair and P. Collins (John Benjamins Publishing Company, Amsterdam, Philadelphia), pp. 17–44.

Cox, F., and Palethorpe, S. (**2004**). "The border effect: Vowel differences across the NSW-Victorian border," in *Proceedings of the Conference of Australian Linguistics Society*, July 13–15, Sydney, Australia.
Cox, F., and Palethorpe, S. (**2007**). "Australian English," J. Int. Phon. Assoc. **37**(3), 341–350.
Cox, F., Palethorpe, S., and Bentink, S. (**2014**). "Phonetic archaeology and 50 years of change to Australian English /iː/," Aust. J. Ling. **34**(1), 50–75.
Delattre, P. C., Liberman, A. M., and Cooper, F. S. (**1955**). "Acoustic loci and transitional cues for consonants," J. Acoust. Soc. Am. **27**(4), 769–773.
Efron, B., Halloran, E., and Holmes, S. (**1996**). "Bootstrap confidence levels for phylogenetic trees," Proc. Natl. Acad. Sci. USA **93**(23), 13429–13434.
Elvin, J., Williams, D., and Escudero, P. (**2016**). "Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English," J. Acoust. Soc. Am. **140**(1), 576–581.
Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton, The Hague).
Fromont, R., and Watson, K. (**2016**). "Factors influencing automatic segmental alignment of sociophonetic corpora," Corpora **11**(3), 401–431.
Garrett, A., and Johnson, K. (**2013**). "Phonetic bias in sound change," in *Origins of Sound Change*, edited by A. C. L. Yu (Oxford University Press, Oxford, UK), pp. 51–97.
Gick, B., Kang, M. A., and Whalen, D. H. (**2002**). "MRI evidence for commonality in the post-oral articulations of English vowels and liquids," J. Phon. **30**(3), 357–371.
Gick, B., and Wilson, I. (**2006**). "Excrescent schwa and vowel-laxing: Cross-linguistic responses to conflicting articulatory targets," in *Papers in Laboratory Phonology*, edited by L. M. Goldstein, D. H. Whalen, and C. T. Best (Mouton de Gruyter, Berlin), pp. 635–659.
Giles, S., and Moll, K. (**1975**). "Cinefluorographic study of selected allophones of English /l," Phonetica **31**(3-4), 206–227.
González, S., Grama, J., and Travis, C. (**2018**). "Comparing the accuracy of forced-aligners for sociolinguistic research," in *CoEDL Fest*, February 5–8, Melbourne, Australia.
Gorman, E., and Kirkham, S. (**2020**). "Dynamic acoustic-articulatory relations in back vowel fronting: Examining the effects of coda consonants in two dialects of British English," J. Acoust. Soc. Am. **148**(2), 724–733.
Harrington, J. (**2010**). *Phonetic Analysis of Speech Corpora* (Wiley-Blackwell, Chichester, UK).
Harrington, J., and Cassidy, S. (**1994**). "Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English," Lang. Speech **37**(4), 357–373.
Harrington, J., Cox, F., and Evans, Z. (**1997**). "An acoustic phonetic study of broad, general, and cultivated Australian English vowels," Aust. J. Ling. **17**, 155–184.
Harrington, J., Kleber, F., Reubold, U., Schiel, F., and Stevens, M. (**2018**). "Linking cognitive and social aspects of sound change using agent-based modeling," Top. Cogn. Sci. **10**(4), 707–728.
Harris, J. (**1994**). *English Sound Structure* (Blackwell, Oxford, UK).
Hyman, L. M. (**2013**). "Enlarging the scope of phonologization," in *Origins of Sound Change: Approaches to Phonologization*, edited by A. C. L. Yu (Oxford University Press, Oxford, UK), pp. 3–28.
Kisler, T., Reichel, U., and Schiel, F. (**2017**). "Multilingual processing of speech via web services," Comput. Speech Language **45**, 326–347.
Kleber, F., Harrington, J., and Reubold, U. (**2012**). "The relationship between the perception and production of coarticulation during a sound change in progress," Lang. Speech **55**(3), 383–405.
Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (**2017**). "lmerTest package: Tests in linear mixed effects models," J. Stat. Softw. **82**(13), 30702.
Labov, W., Ash, S., and Boberg, C. (**2008**). *The Atlas of North American English, Phonetics, Phonology and Sound Change* (Gruyter Mouton, Berlin).
Lenth, R. (**2019**). "Emmeans: Estimated marginal means, aka least-squares means, r package (version 1.3.4) [computer program]," https://CRAN.R-project.org/package=emmeans (Last viewed 11 February 2021)..
Lewis, E., and Loakes, D. (**2012**). "The /ɐlC-ɔlC/ sound change in Australian English: Preliminary resɔlts," in *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, December 3–6, Sydney, Australia, pp. 73–76.
Liaw, A., and Wiener, M. (**2002**). "Classification and regression by randomforest," R News **2**(3), 18–22.
Lin, S., Palethorpe, S., and Cox, F. (**2012**). "An ultrasound exploration of Australian English /CVl/ words," in *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, December 3–6, Sydney, Australia, pp. 105–108.
Loakes, D., Clothier, J., Hajek, J., and Fletcher, J. (**2014**). "An investigation of the /el/-/æl/ merger in Australian English: A pilot study on production and perception in South-West Victoria," Aust. J. Ling. **34**(4), 436–452.
Loakes, D., Graetzer, N., Hajek, J., and Fletcher, J. (**2012**). "Vowel perception in Victoria: Variability, confusability and listener expectation," in *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, December 3–6, Sydney, Australia.
Loakes, D., Hajek, J., and Fletcher, J. (**2010a**). "Issues in the perception of the /el/-/æl/ contrast in Melbourne: Perception, production and lexical frequency effects," in *Proceedings of 13th Australasian International Conference on Speech Science and Technology*, December 14–16, Melbourne, Australia.
Loakes, D., Hajek, J., and Fletcher, J. (**2010b**). "(Mis)perceiving /el/-/æl/ in Melbourne English: A micro-analysis of sound perception and change," in *Proceedings of 13th Australasian International Conference on Speech Science and Technology*, December 14–16, Melbourne, Australia.
Loakes, D., Hajek, J., and Fletcher, J. (**2010c**). "The /el/-/æl/ sound change in Australian English: A preliminary perception experiment," in *Selected Papers from the 2009 Conference of the Australian Linguistic Society*, edited by Y. Treis and R. De Busser (Australian Linguistics Society, Melbourne, Australia).
Loakes, D., Hajek, J., and Fletcher, J. (**2011**). "/æl/-/el/ transposition in Australian English: Hypercorrection or a competing sound change?," in *Proceedings of the 17th International Congress of Phonetic Sciences*, August 17–21, Hong Kong.
Maguire, W., Clark, L., and Watson, K. (**2013**). "Introduction: What are mergers and can they be reversed?," English Lang. Ling. **17**(2), 229–239.
Munson, B., and Solomon, N. P. (**2004**). "The effect of phonological neighborhood density on vowel articulation," J. Speech Lang. Hear. Res. **47**, 1048.
Oasa, H. (**1989**). "Phonology of current Adelaide English," in *Australian English: The Language of a New Society*, edited by P. Collins and D. Blair (University of Queensland Press, St Lucia, Queensland), pp. 271–287.
Ohala, J. J. (**1989**). "Sound change is drawn from a Pool of synchronic variation," in *Language Change: Contributions to the Study of Its Causes*, edited by L. E. Breivik and E. H. Jahr (Mouton de Gruyter, Berlin), pp. 173–198.
Ohala, J. J. (**1993**). "The phonetics of sound change," in *Historical Linguistics: Problems and Perspectives*, edited by C. Jones (Longman, London), Chap. 9, pp. 237–278.
Palethorpe, S., and Cox, F. (**2003**). "Vowel modification in pre-lateral environments," in *International Seminar on Speech Production*, March 31–April 3, Sydney, Australia.
Proctor, M., Walker, R., Smith, C., Szalay, T., Narayanan, S., and Goldstein, L. (**2019**). "Articulatory characterization of English liquid-final rimes," J. Phon. **77**, 100921.
R Core Team (**2018**). *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).
Schiel, F. (**1999**). "Automatic Phonetic Transcription of Non-Prompted Speech," in *Proceedings 14th International Congress of Phonetic Sciences*, August 1–7, San Francisco, CA, pp. 607–610.
Searle, S. R., Speed, F. M., and Milliken, G. A. (**1980**). "Population marginal means in the linear model: An alternative to least squares means," Am. Stat. **34**(4), 216–221.
Sproat, R., and Fujimura, O. (**1993**). "Allophonic variation in English /l/ and its implications for phonetic implementation," J. Phon. **21**(3), 291–311.
Strycharczuk, P., and Scobbie, J. M. (**2017**). "Fronting of Southern British English high-back vowels in articulation and acoustics," J. Acoust. Soc. Am. **142**(1), 322–331.
Suzuki, R., and Shimodaira, H. (**2006**). "Pvclust: An R package for assessing the uncertainty in hierarchical clustering," Bioinformatics **22**(12), 1540–1542.
Szalay, T., Benders, T., Cox, F., and Proctor, M. (**2018**). "Production and perception of length contrast in lateral-final rimes," in *Proceedings of the 17th Australasian International Conference on Speech Science and Technology*, December 4–7, Sydney, Australia, pp. 127–132.
Thomas, B., and Hay, J. (**2005**). "A pleasant malady: The Ellen/Allan merger in New Zealand English," Te Reo **48**, 69–93.

Turton, D. (**2014**). “Variation in English /l/: Synchronic reflections of the life cycle of phonological processes,” Ph.D. thesis, University of Manchester, Manchester, UK.

Wade, L. (**2017**). “The role of duration in the perception of vowel merger,” J. Assoc. Lab. Phonol. **8**(1), 30.

Ward, J. H. (**1963**). “Hierarchical grouping to optimize an objective function,” J. Am. Stat. Assoc. **58**(301), 236–244.

Watson, C. I., and Harrington, J. (**1999**). “Acoustic evidence for dynamic formant trajectories in Australian English vowels,” J. Acoust. Soc. Am. **106**(1), 458–468.

Winkelmann, R., Jaensch, K., Cassidy, S., and Harrington, J. (**2019**). “emuR: Main Package of the EMU Speech Database Management System, r package version 2.0.3,” https://CRAN.R-project.org/package=emuR (Last viewed 11 February 2021).