Reconsidering lateral vocalisation: Evidence from perception and production of Australian English /l/

Szalay, T.; Benders, T.; Cox, F.; Proctor, M.

DOI
10.1121/10.0014249

Publication date
2022

Document Version
Final published version

Published in
Journal of the Acoustical Society of America

Citation for published version (APA):
Szalay, T., Benders, T., Cox, F., & Proctor, M. (2022). Reconsidering lateral vocalisation: Evidence from perception and production of Australian English /l/. *Journal of the Acoustical Society of America*, *152*(4), 2106-2116. https://doi.org/10.1121/10.0014249

JASA ARTICLE

Reconsidering lateral vocalisation: Evidence from perception and production of Australian English /l/[a)]

Tünde Szalay,[1,b)] Titia Benders,[2,b),c)] Felicity Cox,[3] and Michael Proctor[3]
[1]Discipline of Communication Sciences, University of Sydney, Sydney, New South Wales 2006, Australia
[2]Department of Linguistics, The University of Amsterdam, Amsterdam 1012, The Netherlands
[3]Department of Linguistics, Macquarie University, Sydney, New South Wales 2109, Australia

ABSTRACT:
Lateral vocalisation is assumed to arise from changes in coronal articulation but is typically characterised perceptually without linking the vocalised percept to a coronal articulation. Therefore, we examined how listeners' perception of coda /l/ as vocalised relates to coronal closure. Perceptual stimuli were acquired by recording laterals produced by six speakers of Australian English using electromagnetic articulography (EMA). Tongue tip closure was monitored for each lateral in the EMA data. Increased incidence of incomplete coronal closure was found in coda /l/ relative to onset /l/. Having verified that the dataset included /l/ tokens produced with incomplete coronal closure—a primary articulatory cue of vocalised /l/—we conducted a perception study in which four highly experienced auditors rated each coda /l/ token from vocalised (3) to non-vocalised (0). An ordinal mixed model showed that increased tongue tip (TT) aperture and delay correlated with vocalised percept, but auditors ratings were characterised by a lack of inter-rater reliability. While the correlation between increased TT aperture, delay, and vocalised percept shows that there is some reliability in auditory classification, variation between auditors suggests that listeners may be sensitive to different sets of cues associated with lateral vocalisation that are not yet entirely understood. https://doi.org/10.1121/10.0014249

(Received 20 March 2022; revised 18 August 2022; accepted 8 September 2022; published online 11 October 2022)
[Editor: Richard A. Wright] Pages: 2106–2116

I. INTRODUCTION

Lateral vocalisation is a well-known and widespread phenomenon that has been observed in many languages, as well as in historical language descriptions (Recasens, 1996). Underlying much of this large and diverse body of work is the assumption that lateral vocalisation is a relatively unified phenomenon arising from common, cross-linguistic phonetic mechanisms associated with changes in coronal articulation. This assumption of uniformity is informed and reinforced by the common description of lateral vocalisation in broad segmental terms using a small set of phonetic symbols, based on perceptual accounts (Recasens, 1996). For example, vocalised /l/ in English has been described as a standard allophone (Rubach, 2006; Straka, 1968); a variant associated with ethnolects (Edwards, 2008; Hancock, 1974), geographical regions (Ash, 1982; Bauer and Warren, 2008; Horvath and Horvath, 2002; Wells, 1982), children's speech (Johnson and Britain, 2007; Lin and Demuth, 2015), and non-standard language varieties (Hualde, 2005); a mechanism involved in sound change (Colantoni and Steele, 2005; Gil, 1990; Horvath and Horvath, 2001; Lin *et al.*, 2014; Lynch, 2008; Recasens, 1996); and an indexical marker of social identity (Scobbie and Wrench, 2003; Turton, 2017). In Australian English (AusE), vocalised /l/ has been observed as an allophone associated with young, male speakers and specific phonetic contexts, e.g., it is more frequent before /k/ (e.g., *milk*) than before /t/ (e.g., *tilt*) (Borowsky, 2001; Horvath and Horvath, 2002). In many studies (e.g., Ash, 1982; Borowsky, 2001; Hancock, 1974; Johnson and Britain, 2007; Wells, 1982), including all studies on AusE, lateral vocalisation has been studied perceptually or impressionistically, so no direct information is available about the underlying articulation. Therefore, examining the correspondence between coronal articulation and a vocalised percept is required to further our understanding of /l/-vocalisation. Articulatory studies have also revealed more variation and complexity in lateral production than has previously been assumed (Strycharczuk *et al.*, 2020; Strycharczuk and Scobbie, 2020; Ying *et al.*, 2021), and it is not clear how and how consistently these complex articulatory patterns correlate with a vocalised percept.

A. Patterns of vocalised /l/ in production and perception

The English lateral approximant is a multi-gestural segment canonically articulated with a central coronal closure and tongue dorsum retraction (Giles and Moll, 1975; Ladefoged and Maddieson, 1996). Formation of lateral

[a)]This paper is part of a special issue on Reconsidering Classic Ideas in Speech Communication.
[b)]Also at: Department of Linguistics, Macquarie University, Sydney, New South Wales 2109, Australia.
[c)]Electronic mail: tuende.szalay@sydney.edu.au

channels may result from the simultaneous coronal fronting and dorsum retraction (Ladefoged and Maddieson, 1996) or may be partly actively controlled (Ying *et al.*, 2021). In the few English varieties for which articulatory data are available, differences between clear onset /l/, dark coda /l/, and vocalised coda /l/ have been shown to arise from complex patterns of production rather than just changes in coronal articulation (Giles and Moll, 1975; Scobbie and Pouplier, 2010; Strycharczuk *et al.*, 2020; Strycharczuk and Scobbie, 2020). Clear /l/ is characterised by a relatively advanced tongue dorsum, a raised tongue tip (TT), and the TT gesture temporally preceding the tongue dorsum gesture, whereas dark /l/ is characterised by a more retracted tongue dorsum, less raised TT, and the tongue dorsum gesture temporally preceding the TT gesture (Sproat and Fujimura, 1993; Turton, 2017).

Lateral vocalisation arises from the loss of contact between the TT and the alveolar ridge, delay in TT raising, increased dorsal retraction, and loss of lateral channel(s) (Giles and Moll, 1975; Scobbie and Pouplier, 2010; Strycharczuk *et al.*, 2020; Strycharczuk and Scobbie, 2020). The articulation of vocalised /l/ is primarily characterised by spatial reduction of the coronal gesture, resulting in increased aperture between the tongue blade and the palate, and a lack of TT closure in the central alveolar region (Browman and Goldstein, 1995; Giles and Moll, 1975; Hardcastle and Barry, 1989; Scobbie and Pouplier, 2010; Scobbie *et al.*, 2007; Strycharczuk and Scobbie, 2020; Wrench and Scobbie, 2003). When contact is not achieved, the point of maximum TT raising or minimum TT aperture has been considered as the gestural target of /l/ (Giles and Moll, 1975; Strycharczuk and Scobbie, 2020). Lateral vocalisation has also been examined in the temporal domain, where the TT gesture is delayed, sometimes occurring after the acoustic offset of /l/ (Browman and Goldstein, 1995; Strycharczuk and Scobbie, 2020). In Standard Southern British English, TT delay appears to follow from spatial reduction, as delayed coronal gestures are always produced with spatial reduction, indicating that /l/-vocalisation may be primarily spatial rather than a temporal phenomenon (Strycharczuk and Scobbie, 2020).

Despite the association between clear /l/ and onset position, and dark or vocalised /l/ and the coda position, clear, dark, and vocalised /l/ are not categorical positional allophones, as a continuum has been observed between the configurations associated with each of these lateral variants (Sproat and Fujimura, 1993; Turton, 2017). Changes from clear to dark /l/ have been shown to be gradient and conditioned by duration and morphological complexity (Lee-Kim *et al.*, 2013; Sproat and Fujimura, 1993). Gradient and categorical /l/-darkening may coexist in the same dialect (Turton, 2017).

Vocalised /l/ forms a continuum with dark /l/, as articulatory characteristics of vocalised /l/ may also be present in dark /l/, e.g., both dark and vocalised /l/ are characterised by the loss of lateral channels in New Zealand English (Strycharczuk *et al.*, 2020; Strycharczuk and Scobbie, 2020; Turton, 2017). Vocalised /l/ shows gradience from "less vocalised" to "more vocalised," as the spatial reduction and temporal delay of TT gestures in lateral production are context-dependent (Lin *et al.*, 2014; Strycharczuk and Scobbie, 2020). For example, speakers of Southern British English exhibited complete closure in word-medial /l/, a relatively small aperture in word-final pre-vocalic /l/, and a large aperture in word-final pre-consonantal /l/ (Strycharczuk and Scobbie, 2020). American English speakers produced pre-consonantal /l/ without alveolar closure; however, aperture minima were larger before velars than labials and larger before labials than alveolars (Lin *et al.*, 2014). The increase in minimum aperture and the increase in maximal delay are speaker-specific: when closure is not reached, the size and the temporal location of minimum aperture vary between speakers (Strycharczuk and Scobbie, 2020). In contrast to Southern British and American English, New Zealand English coda /l/ varied categorically between vocalised and dark based on the size of the alveolar aperture (Strycharczuk *et al.*, 2020). The size of the posterior dorsal restriction shows less positional and interspeaker variation, and the dorsal gesture has not been shown to undergo reduction correlated with coronal reduction (Giles and Moll, 1975; Lin *et al.*, 2014).

Vocalised /l/ has been described as a back voicoid that may or may not be rounded (Hall-Lew and Fix, 2012; Hardcastle and Barry, 1989; Wells, 1982). The back vowel-like percept has been hypothesised to arise from the incomplete or delayed coronal closure for two reasons (e.g., Hardcastle and Barry, 1989; Tollfree, 1999). First, it has been hypothesised that the acoustic cues to coronal closure are lost when closure is not achieved, or they may be masked by the following segment or the absence of sound when minimum aperture is delayed beyond the acoustic offset of /l/ (Strycharczuk and Scobbie, 2020; Tollfree, 1999). Second, dorsum retraction may create a back-vowel-like percept due to its articulatory similarity to back vowels, as back vowels, particularly American English /ɔ/, are articulated with a tongue dorsum retraction similar to that of /l/ (Gick *et al.*, 2002; Hardcastle and Barry, 1989). Although listeners can be consistent in their perception of canonically vocalised /l/ (Hall-Lew and Fix, 2012), it is not clear how, and how consistently, these less canonical and more diverse articulatory patterns correspond to percepts of vocalised /l/ (Hall-Lew and Fix, 2012; Hardcastle and Barry, 1989). In addition, exactly what is meant by vocalised /l/ is often not defined at all in the literature describing how laterals are perceived.

Auditory ratings of coda /l/ as vocalised or non-vocalised by phoneticians and phonologists have been used in several sociolinguistic studies, but these differ in the measurement methods and rating scales used as well as in the number of raters employed (e.g., Ash, 1982; Hall-Lew and Fix, 2012; Horvath and Horvath, 2001; Stuart-Smith *et al.*, 2006). Ash (1982) used a five-step scale from vocalised to non-vocalised and linked each step to different articulatory characteristics. For instance, step 0 was defined as clear /l/ characterised by "an apico-alveolar contact," step 2 was

https://doi.org/10.1121/10.0014249

defined as vocalised /l/ with "apparent raising of some part of the tongue," and step 3 was "heavily vocalised with minimal or no raising of the tongue" (Ash, 1982). Although an articulatory definition was provided for each step, articulatory measurements were not taken (Ash, 1982). Other studies used a four-step (Hall-Lew and Fix, 2012) or binary coding (e.g., Horvath and Horvath, 2001; Stuart-Smith *et al.*, 2006) without referencing to articulatory characteristics. The number of coders ranged from one (Ash, 1982; Horvath and Horvath, 2001) to two (Stuart-Smith *et al.*, 2006) to 53 (Hall-Lew and Fix, 2012).

In Hall-Lew and Fix (2012), the reliability of auditory coding was tested using a perceptual survey in which a total of 53 listeners rated a total of 100 tokens of /l/ from vocalised to non-vocalised. Most of the listeners were native listeners of various dialects of English (46 of 53, North American, British, Irish, Australian). All listeners were researchers in linguistics, mostly from sociophonetics, phonetics, and non-phonetic sociolinguistics. Results showed low variation between listeners, and individual tokens received relatively consistent ratings (Hall-Lew and Fix, 2012). Listeners were aware of how their own regional accent may affect their perceptual ratings, but no difference was found between listeners based on social or linguistic background (Hall-Lew and Fix, 2012). While Hall-Lew and Fix (2012) showed the reliability of auditory ratings of /l/-vocalisation, they did not address how the variation in listeners' rating corresponds to variation in /l/ production.

B. AusE /l/-vocalisation

AusE is known for having /l/-vocalisation conditioned by phonetic and sociophonetic factors (Borowsky, 2001; Borowsky and Horvath, 1997; Horvath and Horvath, 1997, 2001, 2002; Wells, 1982). However, AusE /l/-vocalisation has only been studied through impressionistic-auditory ratings: a single researcher encoded tokens of /l/ as non-vocalised or vocalised, when calculating the likelihood of /l/-vocalisation in different phonetic and sociophonetic environments (Borowsky, 2001; Borowsky and Horvath, 1997; Horvath and Horvath, 1997, 2001, 2002).

According to Borowsky (2001) and Horvath and Horvath (2001), AusE pre-pausal and pre-consonantal /l/ are more likely to be vocalised compared to pre-vocalic /l/. Place of articulation of adjacent segments conditions the likelihood of /l/-vocalisation: /l/-vocalisation is least likely before an alveolar, more likely before a bilabial, and most likely before a velar consonant (Borowsky, 2001). Borowsky (2001) proposes that articulatory similarity between the coronal gesture of /l/ and /t/ inhibits /l/-vocalisation, because achieving coronal closure in /t/ increases the likelihood of achieving coronal closure in pre-/t/ laterals. In contrast, articulatory similarity between the posterior gesture of /l/ and /k/ facilitates /l/-vocalisation due to the overall tongue retraction in the /lk/ sequence decreasing the likelihood of achieving a front coronal closure (Borowsky, 2001). Similarly, a preceding back vowel facilitates /l/-vocalisation compared to a front vowel due to the coarticulatory influence of the back vowel decreasing the likelihood of alveolar closure in /l/ (Borowsky, 2001). In contrast, an articulatory study of British English indicated that preceding back vowels inhibit /l/-vocalisation compared to front vowels (Hardcastle and Barry, 1989). This difference may be attributed to the different methodologies as well as to regional accent differences. Vowel length also affects the likelihood of /l/-vocalisation as a preceding long monophthong or diphthong facilitates /l/-vocalisation compared to a preceding short monophthong. This might be explained by the syllable structures associated with /l/ final rimes, which vary according to preceding vowel length: /l/ tends to be syllabic after long vowels in AusE (Borowsky, 2001).

The dataset used to examine phonetic factors by Borowsky (2001) was also used to show that sociophonetic and dialectal factors affect the likelihood of /l/-vocalisation (Horvath and Horvath, 1997, 2001, 2002). Speakers below 30 years were found to vocalise more than older speakers, female speakers more than male speakers, and working class speakers more than middle class speakers (Horvath and Horvath, 1997, 2001, 2002). /l/-vocalisation was more common in the South Australian data collected in Mount Gambier and Adelaide compared to data collected in other states, indicating region-specific dialectal differences in AusE (Horvath and Horvath, 1997, 2001, 2002).

C. Aims and hypothesis

Although it has been commonly assumed that loss of TT contact results in the perception of /l/ as vocalised (Gick, 1999; Hall-Lew and Fix, 2012; Hardcastle and Barry, 1989; Strycharczuk and Scobbie, 2020), the relationship between listeners' perception and speakers' production has not yet been systematically studied. In particular, it is not known to what extent a vocalised percept correlates with the spatial and temporal aspects of coronal articulation associated with laterals produced in different contexts. Therefore, our aim was to examine how listeners' perception of /l/ as vocalised relates to incomplete and delayed coronal closure observed in /l/ production in AusE. We hypothesised that the likelihood of a vocalised percept would be increased by (1) increased aperture of the coronal constriction and (2) delayed achievement of maximal TT height.

II. METHODS

To examine correlations between /l/ production and /l/ perception, we conducted two experiments. Production of /l/ by six speakers of AusE was recorded using electromagnetic articulography (EMA) to be used as stimuli in the perception task. Speakers were selected to minimise interspeaker variation. To maximise intraspeaker variation, laterals were elicited across a range of phonological environments, targeting phonetic contexts that are reported to affect the likelihood of /l/-vocalisation. Having verified with articulatory measurements that the laterals in this experimental corpus are produced with varying degrees of incomplete TT closure, we carried out a perception study, in which four phonetically trained listeners rated every coda /l/ token for vocalisation.

A. Articulatory data collection

1. Speakers

Articulatory data were recorded from six female native speakers of AusE (mean age = 23.4, range = 20–27) to be used as stimuli. Speakers were born and raised in New South Wales (NSW). All but one participant had two NSW-born parents; W2 had one NSW-born and one Victoria-born parent. Speakers received course credit and/or $40/h for participation. None of the participants reported any current or past reading, hearing, or speaking disorders.

2. Material

Laterals were elicited with 39 three-word phrases containing word-final or word-initial /l/ in the second word. Coda laterals were elicited with 33 phrases, across three vowel contexts (front, back, low), four consonantal contexts (glottal, labial, alveolar, dorsal), and three syllable types (Table I). To create three syllable types, we manipulated vowel length and coda complexity: target words contained a simple coda with a short vowel, a simple coda with a long vowel, or a complex coda with a short vowel. Onset laterals were elicited with six phrases with /lVp/ words containing the same high, low, or back vowels, for articulatory comparison. To provide a consistent phonetic frame of reference, non-words were used when necessary (Table I). No phonotactically illicit words were elicited.

3. Articulatory procedure

Participants were instructed to read the phrases aloud while seated approximately 150 cm from a computer screen. They were introduced to the task and the experimental materials with a short practice block. Each phrase was presented orthographically, and presentation was timed automatically. Phrases were divided into two blocks: the first block contained the pre-vocalic onset and pre-glottal targets, the second the pre-alveolar, pre-labial, and pre-velar targets. Targets were randomised within blocks, and the order of the blocks was counterbalanced between participants. Blocks were repeated eight times, eliciting a total of 312 phrases per participant.

Articulatory data were acquired with EMA using an NDI (Waterloo, Canada) Wave system. Lingual articulation was tracked with TT, tongue body, tongue dorsum, and left and right lateral sensors attached to the tongue. Sensors were attached to the upper and lower lips to track lip aperture and lip rounding. One sensor was attached to the gumline below the lower incisor to measure jaw movement. Reference sensors were attached to the nasion and to the left- and right mastoid to track head movement. Sensors were sampled at a rate of 100 Hz. The occlusal plane was located with a bite trial, and the midline of the palate was traced with a palate probe.

Audio was acquired using two microphones located 150 cm from the lips and offset by 15°. A Røde (Sydney, Australia) NTG-1 was connected through a Focusrite (High Wycombe, UK) OctoPre MkII preamplifier to the NDI Wave system. The Røde NTG-1 recorded synchronised acoustic data simultaneously with the spatial data from the sensor coils but also recorded the background noise generated by the NDI Wave system. A second microphone (Røde NT1-A) was connected through a separate Focusrite OctoPre MkII preamplifier to the computer presenting the experimental stimuli, capturing the utterance as a series of WAV files sampled at 44 100 Hz using SpeechRecorder (Draxler and Jansch, 2017). The second microphone recorded unsynchronised acoustic data with less background noise as it was not connected to the NDI Wave system.

4. Articulatory analysis

A total of 39 (target) × 8 (repetitions) × 6 (speakers) = 1872 tokens were elicited. Tokens were excluded if they were misread (77 tokens), if the audio file was corrupted (58 tokens), or if the sensors were tracked incorrectly (27 tokens). In total, 1710 (283 onset, 1427 coda) tokens were included in the experimental corpus.

Sensor position data were corrected for head movement with reference to the fixed reference sensors and rotated into a common coordinate system defined around an origin located on the midsagittal occlusal plane, immediately behind the upper incisors. Sensor traces were low-pass filtered with a 6th-order elliptical low-pass filter (10 Hz cutoff) and conditioned using a discrete cosine transform (DCT)-based penalized least squares discretised smoothing spline (Garcia, 2010).

TABLE I. Target words containing /l/ preceded by the first or followed by the last word of the carrier phrase.

		Word-final				Cluster		
Vowel context	Word-initial	Glottal	Bilabial	Alveolar	Velar	Bilabial	Alveolar	Velar
Front	*far* ***leap***	***peel*** *harp*	***peel*** *parp*	***peel*** *tarp*	***peel*** *karp*			
	far ***lip***	***pill*** *harp*	***pill*** *parp*	***pill*** *tarp*	***pill*** *karp*	***pilp*** *harp*	***pilt*** *harp*	***pilk*** *harp*
Back	*fee* ***lorp***	***Paul*** *heap*	***Paul*** *peep*	***Paul*** *teep*	***Paul*** *keep*			
	fee ***lop***	***pol*** *heap*	***pol*** *peep*	***pol*** *teep*	***pol*** *keep*	***polp*** *heap*	***polt*** *heap*	***polk*** *heap*
Low	*fee* ***larp***	***parl*** *heap*	***parl*** *peep*	***parl*** *teep*	***parl*** *keep*			
	fee ***lup***	***puhl*** *heap*	***puhl*** *peep*	***puhl*** *teep*	***puhl*** *keep*	***pulp*** *heap*	***pult*** *heap*	***pulk*** *heap*

https://doi.org/10.1121/10.0014249

The location of the palate was refined from the complete lingual trajectories of three lingual sensors (tongue dorsum, body, and tip). Because the tongue comes into contact with the roof of the mouth during obstruent production and when resting against the palate, a convex hull defined over this set of points describes the upper limit of lingual excursion, onto which the probe-defined palate trace was mapped.

As the beginning and the end of /l/ could not be identified reliably in the articulatory data due to the lack of discernible boundary between /l/ and the adjacent segments, the analysis window was defined from the acoustic /l/ onset to the end of the vowel for onset /l/ and from the acoustic vowel onset to the end of /l/ for coda /l/ (T0 and T1 in Fig. 1). Acoustic landmarks were identified using the forced-aligner MAUS (Schiel, 1999). All landmarks were manually checked and hand-corrected where necessary based on changes in voicing and amplitude. Articulatory trajectories between the acoustic landmarks were extracted.

TT aperture at each point in time was calculated as the Euclidean distance of the TT sensor to the closest point on the palate, with 0 corresponding to full closure and increasing with incomplete closure. As TT aperture yielded a negative distance in some tokens (i.e., the TT was measured to be above the palate) due to the palate location being approximated, a speaker-specific constant was added to the TT aperture. The coronal gestural target for each lateral was located automatically at the TT aperture minima in the first 30% of the analysis window for onset /l/ or at the local minimum in the last 40% of the analysis window for coda /l/. The time-normalised location of the TT aperture minima in the analysis window was also extracted as a proxy for TT delay. [For more details on post-processing articulatory data, see Szalay (2020), Chap. 6.]

To verify that the corpus included tokens of vocalised coda /l/, i.e., tokens of /l/ produced with incomplete coronal closure, a primary articulatory cue of vocalised /l/, we compared TT aperture in onset and coda laterals. The effect of position on TT aperture was examined with a linear mixed-effect model (LM) using the *lmer* function from the *lme4* package in R (Bates *et al.*, 2015; R Core Team, 2021). We constructed an LM using the dependent variable TT aperture, with the independent variable position (treatment coded, comparing coda /l/ to the baseline onset) and a random by-participant intercept and a by-participant random slope for the effect of position to account for interspeaker variation. The model used the Gaussian family as the residuals followed a normal distribution. p-values were calculated with the *lmerTest* package (Kuznetsova *et al.*, 2017) using Satterthwaite's degrees of freedom method (Giesbrecht and Burns, 1985; Hrong-Tai Fai and Cornelius, 1996). Our model shows that TT aperture is significantly increased for laterals produced in coda position ($\beta = 3.17$, $t_{1.04} = 3.04$, $p = 0.0289$) compared to onset laterals (Fig. 2).

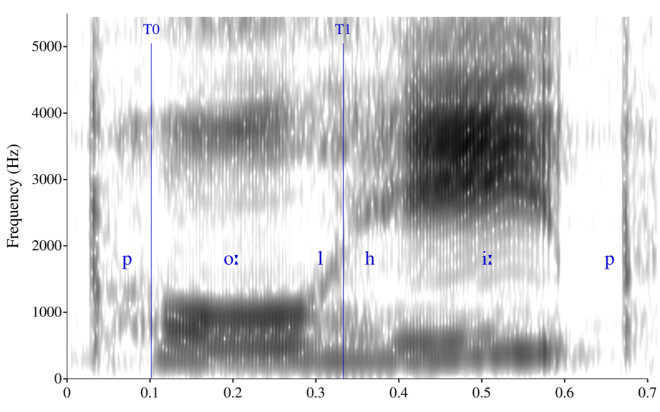

FIG. 1. (Color online) Identifying analysis window in coda laterals, exemplified by *Paul heap* (W4, third repetition). T0, start of the analysis window marked by vowel onset; T1, end of the analysis window, marked by /l/ offset. Phonemic symbols are from hand-corrected annotation.

B. Auditors

In the perception task, four experienced phoneticians listened to each of the coda recordings in the experimental corpus and rated each coda lateral token for vocalisation. Auditors were native speakers of AusE who were born in Australia or migrated to Australia before the age of 1. They were members of the Department of Linguistics at Macquarie University with varying levels of experience in phonetic research. Auditor 1 was a postgraduate research student of phonetics. Auditors 2 and 4 hold a Ph.D. in phonetics and have taught AusE phonetics and phonology at a university level. Auditor 3 has also taught AusE phonetics and phonology at a university level and holds an International Phonetic Association (IPA) Certificate of Proficiency in the Phonetics of English. None of the auditors reported any hearing, reading, or speaking disorders. Only auditor 4 (F.C.) was familiar with the experiment design. Auditors received gift vouchers for their time.

C. Material

The experimental corpus included the 33 types of coda lateral targets produced two to eight times by speakers W2, W3, W4, W5, W7, and W10. The number of repetitions of laterals in each context varied, because of data exclusion

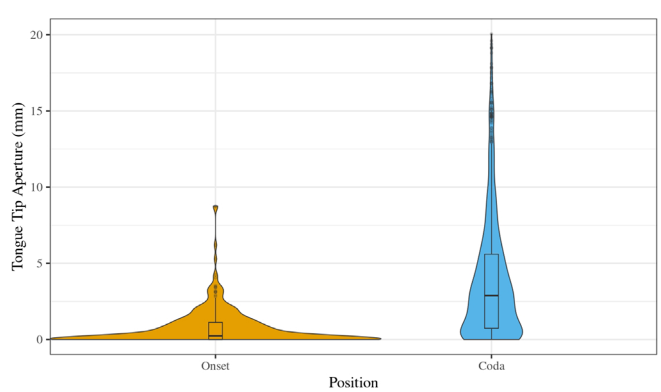

FIG. 2. (Color online) TT aperture (mm) by position. x axis, position. Orange, onset. Blue, coda. Greater TT aperture indicates increased degree of incomplete closure.

(see Sec. IV). Onset laterals were not included in the perception experiment, as onset laterals are not expected to be vocalised. The experimental corpus consisted of 1427 unique tokens; in addition, approximately 10% of the tokens were repeated to measure intra-rater reliability, yielding 1572 ratings per auditor. Repeated tokens were chosen randomly prior to data collection and were the same for all auditors.

Audio recordings of speaker W8, an excluded EMA participant, were used for familiarisation with the task. Data produced by W8 were excluded from the articulatory analysis due to technical difficulties in the articulatory data collection; however, the audio recordings matched the remaining stimuli.

Audio recordings were taken from the second, Røde NT1-A microphone that recorded unsynchronised acoustic data with less background noise. Recordings were amplitude-normalised and truncated after the first word of the three-word phrase, and 0.3 s of silence was added to the beginning, so that auditors were only presented with the last two words of the phrase.

D. Procedure

Auditors rated coda /l/ for vocalisation on a scale ranging from 0 to 3. Auditors were instructed to select 0 for tokens they perceived as non-vocalised and select 1–3 for vocalised tokens, with 3 representing "maximally vocalised." Exemplars were not provided, as the ratings were intended to represent the expert auditors' perception. Prior to the task, listeners were informed that the audio they were about to hear had been recorded during an EMA experiment and contained some amount of background noise produced by the EMA machine and potentially "unusual" articulations caused by the speakers having sensors on their tongues. Listeners were introduced to the task with a short practice session, listening to audio recordings of ten words and rating them.

Auditors were seated in front of a computer monitor located at eye height at a distance of 50 cm and wore Sennheiser (Wedemark, Germany) 380 Pro headphones with the volume adjusted to a comfortable listening level. Auditors were instructed to respond as accurately as possible. To begin each trial, a fixation cross was displayed in the centre of the screen. After 500 ms, the response options appeared, and simultaneously the target phrase started playing. Auditors entered their rating using a button box. Auditors heard each phrase once only and could not change their response. Audio was presented, and ratings and response time were recorded using Expyriment (Krause and Lindemann, 2014).

As rating all tokens took more than 3 h, tokens were divided into three blocks to create shorter tasks to be completed over three separate days, within a 4-day time frame. Each block contained audio from one vowel context, including the repeated tokens from the vowel context. Each block contained audio from all consonant contexts, syllable types, and speakers. Blocks were organised by vowel, because visual inspection of the TT aperture data indicated that grouping by vowel context yielded blocks in which tokens showed a variety of TT aperture. Alternative groupings, such as by speaker, consonant context, or syllable type, yielded blocks with less variation; random grouping in which tokens vary with respect to speaker, vowel and consonant context, and syllable type were reported to be too difficult by pilot participants. The order of the blocks was randomised between auditors, and items were randomised within blocks. After the end of the experiment, auditors were asked about their experience and difficulties with the task in an exit survey.

E. Statistical analysis

Auditors' ratings from 0 to 3 were analysed as ordinal data. Intra-rater and inter-rater agreement were tested using Krippendorff's alpha (Krippendorff, 2011). Krippendorff's alpha calculates agreement between two or more datasets on ordinal data, such that disagreements are weighted differently, e.g., rating the same token as 2 and 3 or as 0 and 1 provides a higher agreement score compared to rating the same token maximally differently with 0 and 3 (Krippendorff, 2011). Krippendorff's alpha ranges from –1 (inverse agreement) to 1 (complete agreement), with 0 indicating no agreement (Krippendorff, 2011). All measurements of intra- and inter-rater reliability were calculated using the library *irr* in R (Gamer *et al.*, 2019; R Core Team, 2021).

Correlation between auditors' rating and TT aperture and TT delay was examined using ordered regression mixed models (ORMMs) using the *clmm* function from the *ordinal* package in R (Christensen, 2019; R Core Team, 2021). We constructed one ORMM with the dependent variable rating, and the independent variables TT aperture and TT delay (continuous variables). The model included a random by-speaker and by-auditor intercept to account for interspeaker and interlistener variation. *p*-values were calculated with the *lmerTest* package (Kuznetsova *et al.*, 2017) using Satterthwaite's degrees of freedom method (Giesbrecht and Burns, 1985; Hrong-Tai Fai and Cornelius, 1996). Only one rating per token was included in the model, yielding 1427 responses per auditor; second ratings of tokens introduced for measuring intra-rater reliability were excluded from the analysis.

III. RESULTS

Krippendorff's alpha indicates weak to moderate intra-rater reliability for three auditors (Table II and Fig. 3). Auditor 3 showed no intra-rater reliability due to the lack of variance in their responses (Fig. 3). Auditor 1 showed slight and auditors 2 and 4 showed moderate agreement on ratings. Krippendorff's alpha indicated no agreement between the four auditors (Table II and Fig. 4).

ORMM indicates that TT aperture shows a significant positive correlation with an increase in vocalisation ratings

TABLE II. Inter- and intra-rater reliability of ratings using Krippendorff's alpha.

Listeners	Intra-rater reliability	Inter-rater reliability
Auditor 1	0.288	0.005
Auditor 2	0.480	
Auditor 3	−0.003	
Auditor 4	0.487	

($\beta = 0.19$, $z_{0.05} = 4.12$, $p < 0.0001$), indicating that tokens with larger aperture are more likely to be perceived as vocalised (Fig. 5). ORMM indicates that TT delay shows a significant positive correlation with an increase in vocalisation ratings ($\beta = 1.6$, $z_{0.35} = 4.53$, $p < 0.0001$), indicating that tokens with larger delay are more likely to be perceived as vocalised (Fig. 6). There was no significant interaction between TT aperture and TT delay ($\beta = -0.1$, $z_{0.05} = -1.89$, $p = 0.059$). That is, there is no evidence for tokens produced with both an increased aperture and a delayed aperture minimum to be more likely to be perceived as vocalised than tokens produced with either. Random effects for speaker showed no variability in the ratings according to the speakers (variance = 0, standard deviation = 0). Random effects for auditor showed variability in the ratings according to the auditors (variance = 4.247, standard deviation = 2.061).

IV. DISCUSSION

We hypothesised that the likelihood of a vocalised percept would be increased (1) by incomplete closure and (2) by delayed achievement of maximal TT height. Our results are broadly consistent with our hypotheses, as increased TT aperture and delay increased vocalised ratings. However, listeners showed low to moderate intra-rater reliability and no inter-rater reliability. Low intra-rater reliability and the lack of inter-rater reliability contrasts with the results of Hall-Lew and Fix (2012), who found that phoneticians can reliably identify vocalised /l/ using auditory-impressionistic methods. These different findings may be due to differences in the materials and methods used: Hall-Lew and Fix (2012) selected tokens that the authors rated unequivocally for /l/-vocalisation on a scale ranging from 1 ("definitely consonantal") to 4 ("definitely vocalised") based on acoustic and auditory observations. In contrast, in the current study, the presence of vocalisation in the experimental materials presented to listeners was determined using articulatory measures of coronal constriction aperture. As a result, our study contained lateral exemplars that were vocalised according to a primary articulatory metric, even though tokens in our stimuli showed smaller aperture compared to vocalised tokens in General American English (Proctor *et al.*, 2019). However, it may be the case that these laterals were not produced with other accompanying phonetic properties or with a sufficiently large TT aperture that give rise to the perception of being canonically vocalised, which may have caused the large number of ratings as non-vocalised.

Low inter-rater reliability might be attributed to individual listeners' reliance on different articulatory-acoustic cues for vocalisation. One potential articulatory property that may lead to the acoustic cue for a canonically vocalised percept may be the lack of lateral channel(s) (Strycharczuk *et al.*, 2020). In New Zealand English, an English dialect closely related to AusE, the loss of lateral channel formation is a characteristic of dark /l/, while the loss of the lateral channel combined with the loss of coronal closure is the characteristic of vocalised /l/. That is, in New Zealand English, loss of lateral channel(s) in dark /l/ precedes incomplete coronal closure in /l/-vocalisation in the course of a sound change (Strycharczuk *et al.*, 2020). Tokens without

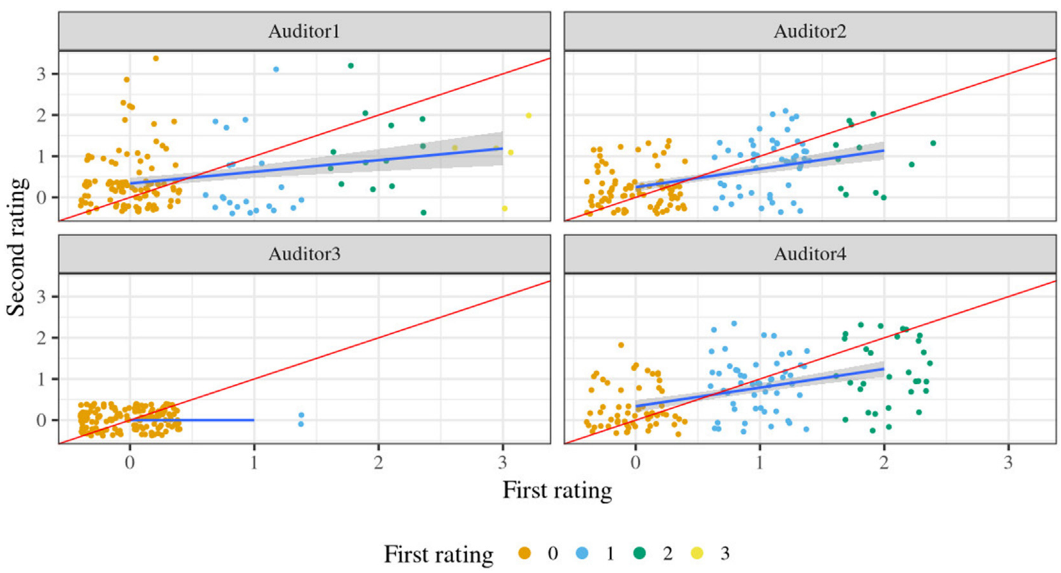

FIG. 3. (Color online) Intra-rater reliability for each auditor. *x* axis, first rating; *y* axis, second rating. Blue regression line, observed correlation between first and second rating. Red line, one-to-one correlation between first and second rating, plotted as a reference. Colour, vocalisation rating: orange, 0; blue, 1; green, 2; yellow, 3. Jitter has been added to aide visualization by avoiding overlapping datapoints.

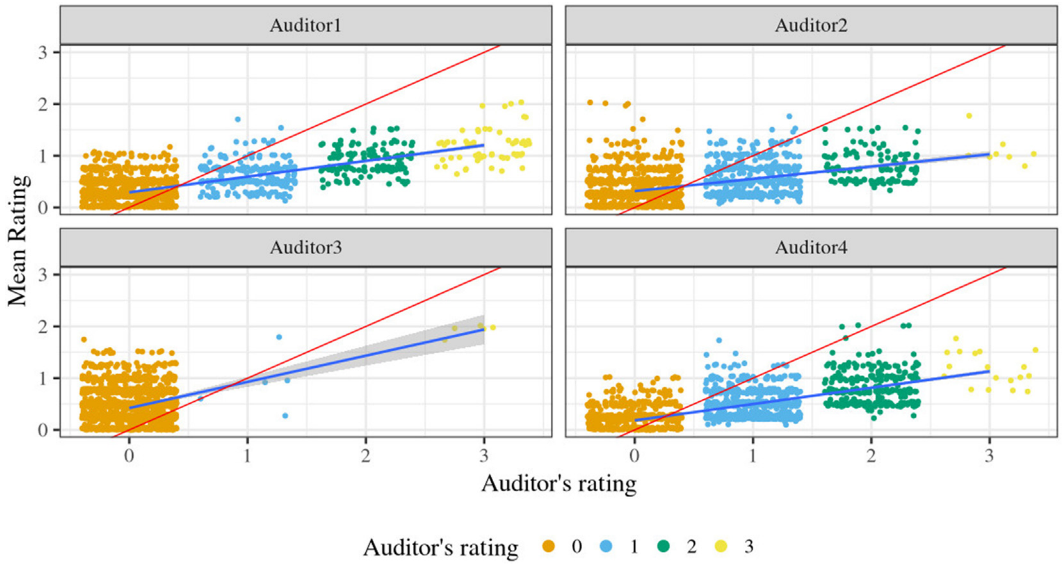

FIG. 4. (Color online) Inter-rater reliability for each auditor. *x* axis, auditor's rating. *y* axis, token's mean rating across four auditors. Blue regression line, observed correlation between auditor's rating and token's mean rating. Red line, one-to-one correlation between listener's rating and token's mean rating, plotted as a reference. Colour, vocalisation rating: orange, 0; blue, 1; green, 2; yellow, 3. Jitter has been added to aide visualization by avoiding overlapping datapoints.

lateral channel and with coronal closure might be rated as vocalised by listeners relying on the lack of lateral channel(s) to identify vocalised /l/, while those relying on incomplete closure would rate the same token as non-vocalised. In AusE, no positional differences have been found in lateral channel formation between onset and coda position, potentially because the loss of the lateral channel is associated with vocalised /l/, and the only study to look at onset-coda differences in lateral channel formation excluded vocalised tokens from the dataset (Ying *et al.*, 2021).

Labialisation of /l/ may also lead to the acoustic cue for a canonically vocalised percept, as incomplete TT closure and tongue dorsum backing creates the configuration of back vowels or /w/, which are prototypically rounded both

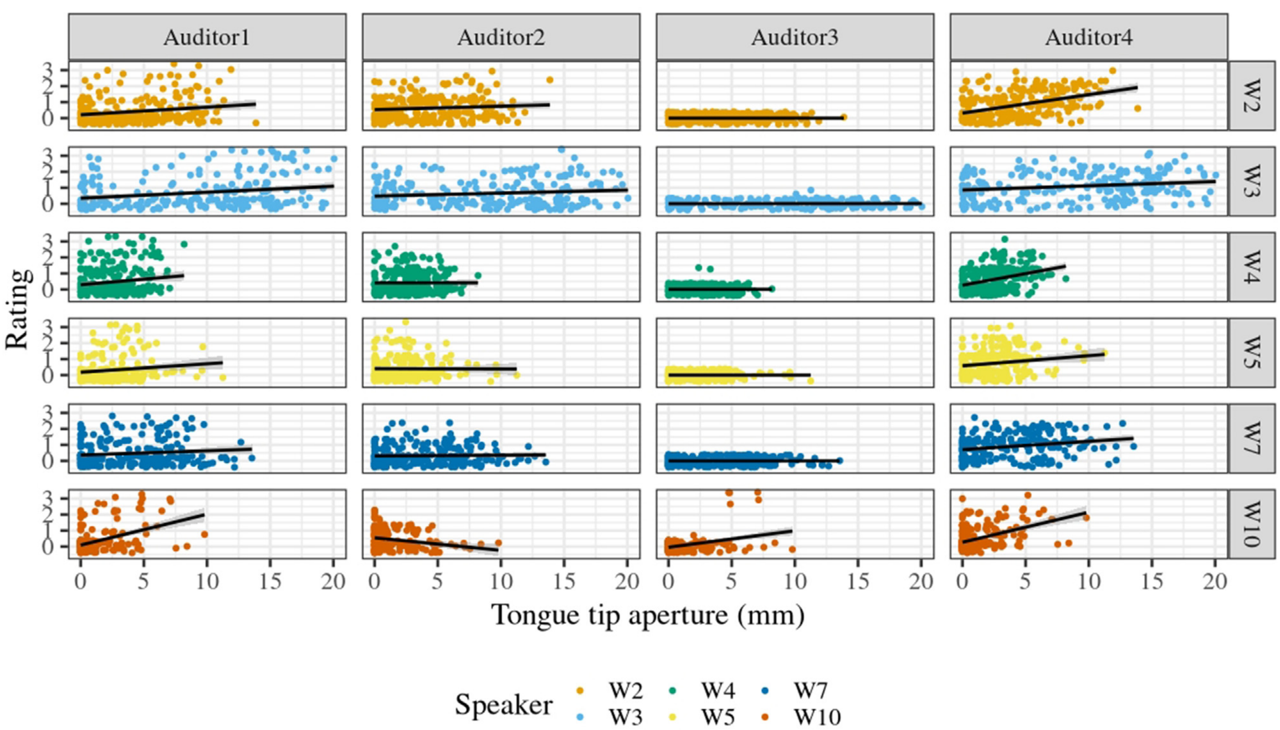

FIG. 5. (Color online) Correlation between TT aperture and auditors' rating. Columns, auditors. Rows, speakers. *y* axis, auditor's rating (larger rating indicates vocalised percept); *x* axis, TT aperture (larger TT aperture indicates incomplete closure). Regression line, observed correlation between auditor's rating and TT aperture. Colour, speaker. Jitter has been added to aide visualization by avoiding overlapping datapoints.

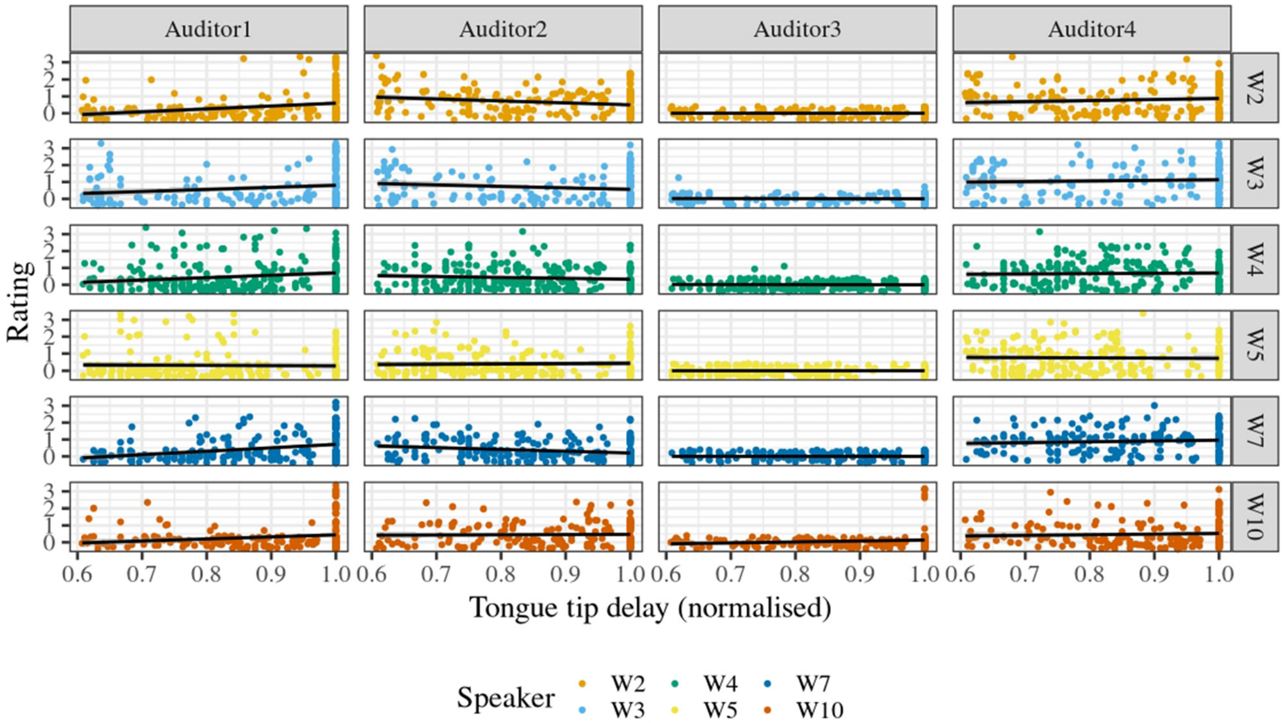

FIG. 6. (Color online) Correlation between TT delay and auditor's rating. Columns, auditors; rows, speakers. *y* axis, auditor's rating (larger rating indicates vocalised percept); *x* axis, TT delay. Regression line, observed correlation between auditor's rating and TT delay. Colour, speaker. jitter has been added to aid visualization by avoiding overlapping datapoints.

in English and cross-linguistically (Gick *et al.*, 2002; Recasens, 1996). A high back rounded vowel or /w/ was used to transcribe /l/-vocalisation by children acquiring dark coda /l/ (Lin and Demuth, 2015). In fact, when children acquiring AusE /l/ do not produce the adult-like velar and alveolar gestures, they may produce coda /l/ with some combination of labial, velar, and/or alveolar constrictions (Lin and Demuth, 2015). Of the possible combinations of the three gestures, non-adult like coda /l/ is most often produced with velar and labial constrictions, which is not perceived as an /l/ but transcribed as a high back rounded vowel or /w/ by trained phoneticians (Lin and Demuth, 2015). In contrast, coda /l/ produced with labial, velar, and coronal constriction is perceived and transcribed as an /l/, although not adult-like (Lin and Demuth, 2015).

Listeners might also be sensitive to contextual variation in /l/-vocalisation. Coda /l/ was collected in a variety of vowel and consonantal contexts, and a more detailed articulatory analysis of the experimental corpus shows that coda /l/ does not form a homogeneous group; instead, it contains tokens with a wide range of TT aperture that varies according to the place of articulation of the adjacent segments [Szalay (2020), Chap. 6]. In particular, a following velar consonant, a preceding back vowel, or a coda cluster facilitates incomplete TT closure, whereas a following alveolar consonant or a preceding long vowel inhibits incomplete TT closure [Szalay (2020), Chap. 6]. This contextual variation in TT aperture could have given rise to inter- and intra-rater variation; however, data from more listeners are required to explore the effect of context on the perception of /l/-vocalisation.

Listeners might also be sensitive to coarticulatory cues to /l/ carried by the preceding vowel to a different extent. Pre-/l/ vowel allophones differ systematically from pre-obstruent vowels (Szalay *et al.*, 2021); thus, /l/-influenced vowels might provide cues to /l/-identity for some listeners. This account would predict that a token containing an /l/-influenced vowel and no coronal closure may be perceived as non-vocalised by listeners sensitive to the coarticulatory vowel cues, whereas listeners sensitive to coronal aperture may rate the same token as vocalised. To further explore what articulatory and acoustic cues contribute to a vocalised /l/ percept, better understanding of articulatory and acoustic differences between canonical and non-canonical /l/ is required.

While the articulation of coda /l/ showed variation according to its phonetic context [Szalay (2020), Chap. 6], our speaker sample was homogeneous with respect to the sociolinguistic factors that govern /l/-vocalisation in AusE. /l/-vocalisation in AusE is associated with young, female, and/or working class speakers (Horvath and Horvath, 2001), and our speakers were matched for age (young) and gender (female); social class could not be established. Therefore, listeners might have had only a few sociolinguistic cues for building speaker-specific expectations regarding /l/-vocalisation. If auditors relied on subconscious biases or conscious expectations based on the speaker's background despite having only a few cues, then cues to age and gender might have increased the number of "vocalised" ratings, while cues to level of education might have decreased them. In contrast, speakers in Hall-Lew and Fix (2012) represented African

American, Asian American, and European American speakers, and the results show "a clear correlation between perceived ethnicity of the speaker and the average vocalisation rating of the speaker." Speaker ethnicity can be cued by several factors, such as vowel identity (Hall-Lew, 2009; Thomas and Reaser, 2004); therefore, listeners could have relied on their overall perception of the speaker's ethnicity and/or other aspects of speaker identity to evaluate each exemplar of /l/.

The main limitation of our study is the small number of listeners; however, increasing the number of listeners is difficult, partly due to the number of restrictions placed on listener selection and mainly due to the difficulty of the task. Despite the strict selection criteria, all four of the (highly experienced) listeners reported that they found the task long, repetitive, and difficult. Therefore, expanding the task to include greater numbers of more diverse auditors would be challenging. Despite these difficulties, future studies with more listeners are required to identify listeners' patterns in the perception of vocalised /l/.

Despite the poor intra-rater and the lack of inter-rater reliability, listeners were universally sensitive to the articulatory factors of TT aperture and delay, as increase in aperture and increase in delay both increased the likelihood of a vocalised percept. This result is consistent with the assumption that key mechanisms of articulation result in robust acoustic cues to common properties of non-canonical laterals.

V. CONCLUSION

Analysis of listeners' perception indicates that vocalised percept corresponds to incomplete closure and delay of the coronal gesture. The correlation between spatial reduction, temporal delay, and vocalised percept shows that there is some reliability in auditory classification of laterals and is consistent with the assumption that articulatory and auditory /l/-vocalisation might correspond to the same underlying phenomenon.

However, this study has revealed that, even among listeners experienced in phonetic description of laterals, there is considerable inconsistency in the perception and characterisation of vocalised and non-vocalised /l/. Therefore, the observed variation between auditors may complicate the interpretation of insights on /l/-vocalisation and highlights the importance of having multiple auditors when using auditory-impressionistic ratings.

ACKNOWLEDGMENTS

This project was approved by the Macquarie University Human Research Ethics Committee under HE26SEP2008-R06061L&P. This research was supported in part by Australian Research Council Grant Nos. ARC 746 DE150100318 and ARC FT180100462 and by Macquarie University Grant Nos. iMQRTP 2015144 and MQSIS 9201501719. We thank Louise Ratko and Ioanna Anastasopoulou for their help in data collection 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.

Ash, S. (**1982**). "The vocalization of /l/ in Philadelphia," Ph.D. thesis, University of Pennsylvania, Philadelphia, PA.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Softw. **67**(1), 1–48.

Bauer, L., and Warren, P. (**2008**). "New Zealand English: Phonology," in *A Handbook of Varieties of English*, edited by B. Kortmann and E. W. Schneider (Mouton De Gruyter, Berlin), pp. 580–602.

Borowsky, T. (**2001**). "The vocalisation of dark l in Australian English," in *English in Australia*, edited by D. Blair and P. Collins (John Benjamins, Amsterdam), pp. 69–87.

Borowsky, T., and Horvath, B. (**1997**). "L-vocalisation in Australian English," in *Variation, Change and Phonological Theory*, edited by F. Hinskens, R. van Hout, and W. L. Wetzels (John Benjamins, Amsterdam), pp. 101–123.

Browman, C. P., and Goldstein, L. M. (**1995**). "Gestural syllable position effects in American English," in *Producing Speech: Contemporary Issues*, edited by F. Bell-Berti and J. L. Raphael (AIP, New York).

Christensen, R. H. B. (**2019**). "Ordinal-regression models for ordinal data," R package version 2019.12-10, https://CRAN.R-project.org/package=ordinal (Last viewed September 28, 2022).

Colantoni, L., and Steele, J. (**2005**). "Liquid asymmetries in French and Spanish," Toronto Work. Papers Linguist. **24**, 1–14.

Draxler, C., and Jansch, K. (**2017**). "Speechrecorder (version 3.10.0) [computer program]," https://www.bas.uni-muenchen.de/Bas/software/speechrecorder/ (Last viewed September 28, 2022).

Edwards, W. F. (**2008**). "African American vernacular English: Phonology," in *A Handbook of Varieties of English*, edited by B. Kortmann and E. W. Schneider (Mouton De Gruyter, Berlin), pp. 383–392.

Gamer, M., Lemon, J., Fellows, I., and Singh, P. (**2019**). "*irr: Various Coefficients of Interrater Reliability and Agreement*," R package version 0.84.1, https://cran.r-project.org/web/packages/irr/irr.pdf (Last viewed September 28, 2022).

Garcia, D. (**2010**). "Robust smoothing of gridded data in one and higher dimensions with missing values," Comput. Stat. Data Anal. **54**(4), 1167–1178.

Gick, B. (**1999**). "A gesture-based accounts of intrusive consonants in English," Phonology **16**(1), 29–54.

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.

Giesbrecht, F. G., and Burns, J. C. (**1985**). "Two-stage analysis based on a mixed model: Large-sample asymptotic theory and small-sample simulation results," Biometrics **41**, 477–486.

Gil, F. M. (**1990**). "Topics in Spanish historical phonology: A non-linear approach," Ph.D. thesis, University of Southern California, Los Angeles, CA.

Giles, S., and Moll, K. (**1975**). "Cinefluorographic study of selected allophones of English /l/," Phonetica **31**(3), 206–227.

Hall-Lew, L. (**2009**). "Ethnicity and phonetic variation in a San Francisco neighborhood," Ph.D. thesis, Stanford University, Stanford, CA.

Hall-Lew, L., and Fix, S. (**2012**). "Perceptual coding reliability of (L)-vocalization in casual speech data," Lingua **122**(7), 794–809.

Hancock, I. F. (**1974**). "English in Liberia," Am. Speech **49**(3), 224–229.

Hardcastle, W., and Barry, W. (**1989**). "Articulatory and perceptual factors in /l/ vocalisations in English," J. Int. Phon. Assoc. **15**, 3–17.

Horvath, B. M., and Horvath, R. J. (**1997**). "The geolinguistics of a sound change in progress: /l/ vocalization in Australia," Penn Work. Pap. Linguist. **4**(1), 109–124.

Horvath, B. M., and Horvath, R. J. (**2001**). "A multilocality study of a sound change in progress: The case of /l/ vocalization in New Zealand and Australian English," Lang. Var. Change **13**, 37–57.

Horvath, B. M., and Horvath, R. J. (**2002**). "The geolinguistics of /l/ vocalisation in Australia and New Zealand," J. Sociolinguist. **6**(3), 319–346.

Hrong-Tai Fai, A., and Cornelius, P. L. (**1996**). "Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments," J. Stat. Comput. Simul. **54**(4), 363–378.

https://doi.org/10.1121/10.0014249

Hualde, J. I. (**2005**). *The Sounds of Spanish with Audio CD* (Cambridge University, New York).

Johnson, W., and Britain, D. (**2007**). "L-vocalisation as a natural phenomenon: Explorations in sociophonology," Lang. Sci. **29**, 294–325.

Krause, F., and Lindemann, O. (**2014**). "Expyriment: A Python library for cognitive and neuroscientific experiments," Behav. Res. **46**(2), 416–428.

Krippendorff, K. (**2011**). "Computing Krippendorff's alpha-reliability," Departmental Papers, Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (**2017**). "lmerTest package: Tests in linear mixed effects models," J. Stat. Softw. **82**(13), 1–26.

Ladefoged, P., and Maddieson, I. (**1996**). *The Sounds of the World's Languages* (Blackwell, Oxford, UK).

Lee-Kim, S.-I., Davidson, L., and Hwang, S. (**2013**). "Morphological effects on the darkness of English intervocalic /l/," Lab. Phonol. **4**(2), 475–511.

Lin, S., Beddor, P., and Coetzee, A. (**2014**). "Gestural reduction, lexical frequency, and sound change: A study of post-vocalic /l/," Lab. Phonol. **5**(1), 9–36.

Lin, S., and Demuth, K. (**2015**). "Children's acquisition of English onset and coda /l/: Articulatory evidence," J. Speech Lang. Hear. Res. **58**(1), 13–27.

Lynch, J. (**2008**). "Liquid vocalization and loss in Central Vanuatu," Ocean. Linguist. **47**(2), 294–315.

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 (**2021**). *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).

Recasens, D. (**1996**). "An articulatory-perceptual account of vocalization and elision of dark /l/ in the Romance languages," Lang. Speech **39**(1), 63–89.

Rubach, J. (**2006**). "Polish: Phonology," in *Encyclopedia of Language and Linguistics*, 2nd ed., edited by E. K. Brown and A. Anderson (Elsevier, Boston), pp. 676–679.

Schiel, F. (**1999**). "Automatic phonetic transcription of non-prompted speech," in *Proceedings of the 14th International Congress of Phonetic Sciences*, edited by J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey, August 1–7, San Francisco, CA, pp. 607–610.

Scobbie, J. M., and Pouplier, M. (**2010**). "Syllable structure and external sandhi: An EPG study of vocalisation and retraction in word-final English /l/," J. Phon. **38**, 240–259.

Scobbie, J. M., Pouplier, M., and Wrench, A. A. (**2007**). "Conditioning factors in external sandhi: An EPG study of English /l/ vocalisation," in *Proceedings of the 16th International Congress of Phonetic Sciences*, August 6–10, Saarbrücken, Germany, pp. 441–444.

Scobbie, J. M., and Wrench, A. A. (**2003**). "An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English," in *Proceedings of the 15th International Congress of Phonetic Sciences*, August 3–9, Barcelona, Spain.

Sproat, R., and Fujimura, O. (**1993**). "Allophonic variation in English /l/ and its implications for phonetic implementation," J. Phon. **21**(3), 291–311.

Straka, G. (**1968**). "Contributions à la description et à l'histoire des consonnes L" ("Contributions to the description and history of 'L' consonants"), Trav. Linguist. Lit. Univ. Strasbourg **6**, 267–326.

Strycharczuk, P., Derrick, D., and Shaw, J. (**2020**). "Locating delateralization in the pathway of sound changes affecting coda /l/," Lab. Phonol. **11**(1), 21.

Strycharczuk, P., and Scobbie, J. M. (**2020**). "Gestural delay and gestural reduction. Articulatory variation in /l/-vocalisation in Southern British English," in *The Corpus Phonology of English: Multifocal Analyses of Variation*, edited by A. Prezwozny, C. Viollain, and S. Navarro (Edinburgh University, Edinburgh, UK).

Stuart-Smith, J., Timmins, C., and Tweedie, F. (**2006**). "Conservation and innovation in a traditional dialect: L-vocalization in Glaswegian," English World-Wide **27**(1), 71–87.

Szalay, T. (**2020**). "Production and perception of lateral-final rimes in Australian English," Ph.D. thesis, Macquarie University, Sydney, Australia.

Szalay, T., Benders, T., Cox, F., Palethorpe, S., and Proctor, M. (**2021**). "Spectral contrast reduction in Australian English /l/-final rimes," J. Acoust. Soc. Am. **149**(2), 1183–1197.

Thomas, E. R., and Reaser, J. (**2004**). "Delimiting perceptual cues used for the ethnic labeling of African American and European American voices," J. Sociolinguist. **8**(1), 54–87.

Tollfree, L. (**1999**). "South East London English: Discrete versus continuous modelling of consonantal reduction," in *Urban Voices: Accent Studies in the British Isles*, edited by P. Foulkes and G. J. Docherty (Arnold, London), pp. 163–184.

Turton, D. (**2017**). "Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English," Lab. Phonol. **8**(1), 13.

Wells, J. C. (**1982**). *Accents of English* (Cambridge University, Cambridge, UK).

Wrench, A. A., and Scobbie, J. M. (**2003**). "Categorising vocalisation of English /l/ using EPG, EMA and ultrasound," in *Proceedings of the Sixth International Seminar on Speech Production*, December 8–10, Sydney, Australia, pp. 314–319.

Ying, J., Shaw, J. A., Carignan, C., Proctor, M., Derrick, D., and Best, C. T. (**2021**). "Evidence for active control of tongue lateralization in Australian English /l/," J. Phon. **86**, 101039.