Articulatory-acoustic mappings in Australian English /l/

Jia Ying[1], Christopher Carignan[2], Michael Proctor[3], Jason A. Shaw[4], Donald Derrick[5], Catherine T. Best[1]

[1]MARCS Institute, Western Sydney University, Penrith NSW 2751, Australia

[2]Division of Psychology and Language Sciences, University College London, London WC1E 6BT, UK

[3]Department of Linguistics, Macquarie University, North Ryde NSW 2109, Australia

[4]Department of Linguistics, Yale University, New Haven CT 06511, USA

[5]New Zealand Institute of Brain and Behaviour, University of Canterbury, Christchurch 8041, New Zealand

Abstract

The relationship between the articulatory and acoustic properties of lateral approximants is still imperfectly understood. Research on articulatory-acoustic relations in /l/ production has been primarily focused on mid-sagittal articulation, such as tongue tip raising and tongue body retraction (Sproat & Fujimura, 1993; Browman & Goldstein, 1995). The relative lack of data on para-sagittal articulation represents a critical gap in our understanding of /l/ production. The present study investigated the acoustic properties associated with para-sagittal lateralization in Australian English /l/. Production of /l/ in two different vowel environments (/æ/ and /ɪ/) by six speakers of Australian English was tracked using three-dimensional electromagnetic articulography, with synchronized audio recordings. Degree of lateralization was tracked over time, by comparing parasagittal tongue height to mid-sagittal height at the tongue blade. Analysis of the relationship between formant frequencies and degree of tongue lateralization revealed a positive correlation between F3 values and tongue lateralization on the dominant side. This finding indicates that acoustic characterization can be directly related to articulatory data. Tongue lateralization is a strong predictor of F3 frequency.

Keywords: /l/ production; articulation and acoustics; tongue lateralization; electromagnetic articulography

I. INTRODUCTION

The lateral approximant /l/ has several variants in most varieties of English (Huffman, 1997). As well as syllabic realizations (e.g. [l̩] in ‘middle’), two main positional allophones of /l/ have been identified in American English (Sproat & Fujimura, 1993; Browman & Goldstein, 1995; Lin, 2011) and some varieties of British English (Cruttenden, 2008; Turton, 2017): syllable-onset, and syllable-coda variants. Onset /l/ is typically more palatalized, and coda /l/ is typically more pharyngealized or velarized (Chomsky & Halle, 1968; Ladefoged, 2001). Articulatorily, both variants of /l/ have been shown to involve two coordinated gestures in the mid-sagittal plane: a tongue tip gesture and a tongue dorsum gesture (Sproat & Fujimura, 1993; Browman & Goldstein, 1995). The two lateral allophones differ both in the relative magnitude of movement in the tongue tip and dorsum gestures, and the relative timing between the two gestures. In addition to syllable position effects, coarticulation with a neighboring vowel can also affect tongue position. A recent electromagnetic articulography (EMA) study on English /l/ production demonstrated that an adjacent vowel effects /l/ production in the mid-sagittal plane more than in the para-sagittal plane (Ying et al., submitted).

Many studies have examined the articulatory (Sproat & Fujimura, 1993; West, 1999; Wrench & Scobbie, 2003) and acoustic (Giles & Moll, 1975; Epsy-Wilson, 1992; Huffman, 1997) characteristics of English /l/. Only a few studies, however, have addressed the relationship between the acoustics and the articulation of /l/ (e.g., Ying, Shaw, Kroos, & Best, 2012), and none that we are aware of compare the acoustics of /l/ to the dynamics of lateral channel formation. To address this deficit, this study investigates how the acoustic and articulatory properties during /l/ production are correlated, particularly with respect to the formation of the lateral channel(s) during /l/ production, to improve our current understanding of acoustic-articulatory relations in /l/.

A. The articulatory characteristics of /l/

/l/s are typically produced with an alveolo-palatal lingual contact along the mid-sagittal plane, with air flowing along one or both sides of the tongue blade (Ladefoged & Maddieson, 1996). /l/ production involves a raising and fronting tongue tip gesture, and a lowering and retracting tongue body gesture. Ladefoged and Maddieson (1996) propose that mid-sagittal articulation alone could give rise to the lateral channel(s) that characterize /l/: when the tongue tip stretches forward at the same time that the tongue dorsum retracts, the two antagonistic gestures elongate the tongue and cause the tongue blade at either side or both sides to pull away from the side teeth, creating a lateral channel for airflow along the sides of the tongue.

Sproat and Fujimura (1993) also proposed that American English /l/s involve two gestures, but described it them as a consonantal apical gesture (tongue tip) and a vocalic dorsal gesture (tongue dorsum). The consonantal tongue tip raising and fronting gesture is relatively stronger in onset /l/ than in coda /l/. On the other hand, the production of coda /l/ is more dominated by the ‘vocalic’ dorsal gesture (Sproat & Fujimura, 1993). This is consistent with evidence that coda /l/ is realized with a delayed tongue tip gesture. In a cinefluorographic study conducted by Giles and Moll (1975), the tongue body was lower and more retracted in syllable-coda /l/ than in syllable-onset /l/, and tongue tip raising and fronting was less pronounced as the tongue body pulled further back. Onset /l/s, conversely, were produced with a more anterior tongue dorsum position and undershoot of apical alveolo-palatal contact.

Those studies examined lingual articulation of /l/ in the mid-sagittal plane only. However, lateral production cannot be characterized solely in terms of mid-sagittal tongue movements. A few authors (Stone et al., 1991, 1992; Stone & Lundberg, 1996; Alwan et al., 1997; Narayanan et al., 1997) have speculated that active para-sagittal tongue blade gestures are involved in lateral channel formation. One recent ultrasound study has demonstrated that the production of /l/ also includes lateralization of the tongue blade, at least in syllable-onset /l/s (Lin, Beddor & Coetzee, 2014). Narayanan, Byrd and Kaun (1999) have also suggested that /l/ production goes beyond the mid-sagittal constriction targets. They argue that the para-sagittal tongue blade, the tongue tip, and the tongue body movements are all active and independent movements but work together to produce an /l/. The complex tongue movements during the production of /l/ take advantage of the tongue as a “boneless, jointless structure that can elevate, depress, widen, narrow, extend, retract, and move differentially, both laterally-to-medially and left-to-right” (Stone, Faber and Cordaro 1991).

B. The acoustic characteristics of /l/

The acoustic properties of /l/ vary depending on vowel contexts, syllable positions and speakers. These variations make /l/ more difficult than other consonants to characterize acoustically. Lower formants are more typically more prominent than the higher formants, which are reduced in intensity due to the influence of lateral anti-formants (Narayanan et al., 1999). The first formant (F1) frequency in /l/ is low. The second formant (F2) frequency varies depending on the location of the tongue body according to syllable position and vowel context (Ladefoged and Maddieson, 1996). Table I shows a summary of F1 and F2 frequencies in English /l/ across syllable position from two classic datasets (Lehiste, 1964; Bladon & Al-Bamerni, 1976).

Lehiste (1964) examined the acoustic of /l/s in different syllable positions (onset vs. coda) and vowel contexts (high vowel context /i, u/ vs. low vowel context /æ, ɔ:/) in American English. She measured mean F1, F2 and F3 frequencies in /l/. She found that F3 is high in /l/s. She suggested that the spectral discontinuities might result from anti-formants and identified the high F3 as F4. This claim is in line with Fant’s (1960) study. Only F1 and F2 frequencies are shown in Table I. In another study, Bladon and Al-Bamerni (1976) investigated allophonic variations in the quality of Received Pronunciation (PR) /l/. Onset /l/s, coda /l/s and syllabic /l/s under a wide range of vowel contexts were examined. Measurements were made of a few spectral properties including the first four formants (F1, F2, F3 and F4). They found that the most determinant /l/ quality is F2 frequency. Only F1 and F2 frequencies were reported in their study (Bladon & Al- Bamerni, 1976).

F3 is associated with the cavity in front of the closure location (Fant, 1960), and is also affected by the back cavity (Steven, 1998; Carter, 2002, p. 84). In the following discussion, we discuss what is already known about how the location of the tongue tip and tongue body constrictions along the mid-sagittal vocal tract affect the F1 and F2.

TABLE I: Summary of the first two formant frequencies and/or range in American English and British English /l/s across different syllable positions.

Data	Syllable position	Formants		Data	Syllable position	Formants	
		F1	F2			F1	F2
Lehiste, 1964	onset	295 Hz	950 Hz	Bladon & Al-Bamerni, 1976	onset	300~425 Hz	1100~1600 Hz
	coda	455 Hz	795 Hz		coda	350~550 Hz	700~1000 Hz

C. Articulatory-acoustic relations in /l/ production

We focus our discussion of articulatory-acoustic relations for /l/ on Fant’s tube theory of speech acoustics (1960) and a modification of the tube theory (Stevens, 1998). Fant (1960) modelled the production of /l/ as a tube with lateral side channel(s) in which F1 is approximately a Helmholtz resonance with acoustic mass due to the lateral airflow. A low F2 value is caused by the influence of tongue body retraction on the pharyngeal cavity. F3 is roughly a resonance of the oral cavity anterior to that constriction. The production of /l/ with alveolo-palatal tongue tip contact forms two sub-cavities, one behind the constriction by the tongue blade, the other under the raised tongue tip. The cavity under the tongue tip is acoustically coupled with the back cavity. This adds poles and zeros in the transfer function. According to Fant (1960), the sub-cavity formed by the tongue blade contributes to the additional zero, while the entire cavity contributes to the additional pole. This additional pole-zero pair can cause formant shifting.

Regarding the effect of the lateral channel(s) on the acoustics, Stevens (1998) proposed a modification of Fant's tube theory. He proposed that the frequency of F3 is affected by the lateral channel(s). When the two lateral channels are asymmetrical, additional zeros are added to the vocal tract. The location of the zeros varies across speakers with the length of the lateral channel. Stevens (1998) claimed that the lateral channel formation should modify the spectrum in the frequency range of 2500 ~ 4000 Hz, i.e., typically of F3.

Narayanan, Alwan and Haker (1995, 1997) confirmed the presence of asymmetrical lateral channels in their magnetic resonance imaging (MRI) and electropalatographic (EPG) data in American English /l/, and proposed some acoustic consequences of the articulatory configurations revealed by these data. Most of the spectral energy of /l/ is typically concentrated below 5 kHz. Helmholtz resonances between a large back cavity and the lossy oral configuration result in a low frequency wide bandwidth F1 in the range 250 to 500 Hz. An F2 in the range 1250 to 1450 Hz is the characteristic resonance of a back cavity with a length between 12 and 14 cm (Narayanan et al. 1997: 1074).

The studies above provide rich information on the relations between articulatory-acoustic characteristics of English /l/ based on consideration of articulation in the mid-sagittal plane. By definition, /l/s are apical consonants with mid-sagittal closure behind the incisor. This mid-sagittal tongue tip closure causes the lateral channel to form along the side(s) of the tongue blade (Narayanan et al., 1999), which curl downwards. Formation of the lateral channel has acoustic consequences. According to Stevens (1998), it contributes a pole-zero-pole cluster in the frequencies below 5 kHz. In addition to the acoustic effects of the lateral channel, the formation of a sublingual cavity (the area under the raised tongue tip) can also affect the acoustics. The sublingual cavity size predicts F3 frequency. It lowers the frequency associated with the front cavity as the volume of the front cavity increases (Narayanan & Alwan, 1996). Stevens (1998) also suggests that the sublingual cavity can set up an additional pole-zero pair around 2000 ~ 5000 Hz in the spectrum.

Recasens and Espinosa (2005, 2012) summarize the observed articulatory-acoustic relationships in /l/ production: 1) Higher F1 frequency is associated with lower jaw position, lower tongue dorsum height, and wider cross-sectional lateral constriction, whereas a lower F1 frequency is associated with higher jaw position, higher tongue dorsum height, and narrower cross-sectional lateral constriction; 2) Higher F2 frequency is associated with higher and more fronted tongue dorsum, greater dorso-palatal contact, shorter back cavity, and narrower back constriction, whereas lower F2 frequency is associated with lower and more backed tongue dorsum, less dorso-palatal contact, longer back cavity and wider back constriction: 3) Higher F3 frequency is more associated with coda allophones than onsets; for clear laterals and laterals produced in low vowel contexts, F3 frequency is lower. Though these findings are based on /l/ production in Spanish, a few cross-language studies have shown that the articulatory and acoustic characteristics of dental [l]s in various languages appear to be very similar (Narayanan, Alwan, & Haker, 1997).

Most of the studies reviewed above have focused on the relation of the mid-sagittal gestural movements (tongue tip raising and tongue dorsum retraction) and their acoustic consequences. Thus far, there have been no studies on the acoustic characteristics associated with para-sagittal tongue blade movements during lateral channel formation, and their relationship to mid-sagittal articulatory movements. Investigating those articulatory-acoustic relationships was the goal of the study reported here.

D. The present study

The aim of the present study was to determine the relationships between the articulatory and acoustic properties of lateral channel formation, specifically in Australian English /l/. Using 3D electromagnetic articulography (EMA), we examined tongue movement at key points on the midline (mid-sagittal) and sides (para-sagittal) of the tongue to track the formation of the lateral channel during /l/ production. Combined with simultaneous high-quality audio recordings, these data provide new detail on the acoustic consequences of para-sagittal kinematics during production of /l/.

Lack of data on lateral channel formation represents an empirical gap which could potentially adjudicate between two different theoretical interpretations of the available kinematic data for /l/ production and their expected acoustic consequences. Extrapolating from past work on articulatory and acoustic properties of /l/ production (Stone et al, 1991, 1992; Stone & Lundberg, 1996; Alwan et al., 1997; Narayanan et al., 1997; Proctor, 2009), four predictions (P1, P2, P3 and P4) were formulated for the present study:

- Prediction 1: Tongue lateralization (i.e., para-sagittal tongue blade movement) will be inversely correlated with F1 frequency. Traditionally, F1 frequency is associated with tongue body height (Recasens & Espinosa, 2006). However, according to Lin, Beddor & Coetzee (2014), the tongue tip (TT) constriction value can indirectly reflect tongue blade constriction, i.e., tongue lateralization (Sproat & Fujimura, 1993; Proctor, 2011; Lin, Beddor & Coetzee). A moderate correlation between tongue lateralization and F1 frequency is expected, since F1 frequency has been found to be correlated with both TT and tongue blade constriction.

- Prediction 2: F3 frequency and tongue lateralization will be correlated. According to Fant (1960), F3 is associated with the cavity in front of the closure location. For this reason, a strong correlation between tongue lateralization and F3 frequency is expected.
- Prediction 3: F2 frequency is associated with tongue dorsum (TD) position; therefore, tongue lateralization is expected to be less strongly related to F2 frequency than it is to F1 and F3 frequencies.
- Prediction 4: As the effects of syllable position on tongue lateralization are minimal (Ying et al., submitted), the effects on both F1 and F3 frequencies will also be minimal.

II. METHODS

The present study was designed to investigate articulatory-acoustic relationships in lateral channel formation during /l/ production in Australian English (AusE). The approach was informed by our previous findings on the same articulatory data, in which we had determined the relationships between known variations in the timing of mid-sagittal gesture movements and the para-sagittal dynamics involved in /l/ formation (Ying et al., submitted).

A. Participants

Six monolingual AusE speakers (three females and three males; mean age 22.2 years, range = 19-35 years) participated (Ying et al., submitted). None of the participants were characterized as having atypical speech, and none had pervasive syllable-final lateral vocalization; that is, they articulated their final /l/s as lateral approximants rather than as high back vowels. All were living in Sydney at the time of data collection. They were paid for their participation and were naïve to the purpose of the experiment. Written consent was obtained from all participants and the study was conducted with approval from the ethics committee of Western Sydney University.

B. Experimental Material

Laterals were elicited word-medially in disyllabic words of the form /ˈCVb.lət/ or /ˈ(C)Vl.bət/, allowing comparison of both syllable-onset and syllable-coda /l/s. In both syllable-onset and syllable-coda positions, laterals were preceded by a stressed front vowel, either /æ/ or /ɪ/. The vowels /æ/-/ɪ/ were chosen because of the different constraints that they place on the shape of the tongue preceding /l/. During the production of /l/, the tongue tip (TT) is raised, the tongue middle (TM) is lowered, and the tongue dorsum (TD) is retracted gradually (Sproat & Fujimura, 1993; Goldstein, 1995; 1996; Campbell & Gick, 2003). According to Stone and Lundberg's (1996) 3D ultrasound study, the vowel /æ/ requires the opposite para-sagittal configuration as /l/; that is, /æ/ usually has a medial groove tongue shape, such that the sides of the tongue are curved up (instead of curved down as in laterals) to form a spoon-shaped, concave configuration in the coronal plane, a shape which conflicts with lateral channel formation. In contrast, a high front vowel such as /ɪ/ does not conflict to the same degree with the tongue shape required for /l/. The TT and TM are raised for /ɪ/ production, with the tongue showing a convexity in tongue blade in the coronal plane and a concavity at the tongue back. We found previously that vowel contexts affect /l/ production mid-sagittally, but not para-sagittally (Ying et al., submitted). Target words were read aloud in the carrier phrase 'keep __ here', with adjacent /p/ and /h/ chosen to minimize lingual coarticulation effects. /b/ preceding or following the /l/ is used in both target forms for this purpose as well. (see Table II). The stimuli were presented in 10-word blocks, with ten repetitions for each target word, randomized across blocks for each participant. Each recording session took approximately 25 minutes for a participant to complete.

TABLE II: Target words used to elicit Australian English laterals. Note that AusE is a non-rhotic variety of English, thus final ‘-ert’ is pronounced [ət].

preceding vowel	æ		ɪ	
mid-sagittal tongue shape	concave		convex	
Representations	orthographic	IPA	orthographic	IPA
CVC.lVC	tablet cablet	[ˈtæb.lət] [ˈkæb.lət]	tiblet kiblet	[ˈtɪb.lət] [ˈkɪb.lət]
(C)Vl.CVC	Talbot Calbert Albert	[ˈtæl.bət] [ˈkæl.bət] [ˈæl.bət]	Tilbert Kilbert Ilbert	[ˈtɪl.bət] [ˈkɪl.bət] [ˈɪl.bət]

C. Procedure

Experiments were conducted at the MARCS Institute Speech Production Laboratory at Western Sydney University. Articulographic data were acquired at a rate of 100 Hz using an NDI Wave electromagnetic articulography (EMA) system (Northern Digital Inc., Canada). Synchronized companion speech audio was recorded at a 22,050 Hz sampling rate using a Schoeps Colette Series Supercardioid microphone and EURORACK UB802 preamplifier. Tongue, lip and jaw movements were tracked using three EMA sensors affixed mid-sagittally to the tongue tip (TT; ~5 mm behind the apex), tongue middle (TM; ~20 mm behind the TT sensor) and tongue dorsum (TD; between 20 and 35 mm behind the TM sensor, depending on the participant’s tolerance), and another two sensors affixed para-sagitally to the sides of the tongue blade (on the top surface ~5 mm from the edges of the tongue and ~15 mm from both the TT sensor and the TM sensor). The TD sensor was located 45 to 60 mm posterior to the TT sensor, depending on each speakers’ comfort level. Figure 1 provides a schematic of the tongue sensor placement: the ‘Southern Cross’ configuration we designed for our prior analyses (Ying et al., submitted) to allow measurement of para-sagittal dynamics.

Sensors were also attached to the lower jaw on the gum line between the two central incisors; to the upper lip and lower lips along the vermillion border in the mid-sagittal plane; to the left mastoid (LM) and right mastoid (RM); and to the nasion (NA). The LM, RM and NA sensors were used for correction of head motion for post-collection data processing. Three sensors are required to account for the translation and rotation of the head using x, y, and z coordinates. The occlusal (bite) plane was determined by having speakers clench a semi-circular protractor between their upper and lower teeth. Two sensors were attached to the corners of the protractor and the third sensor was attached to the center of the circular portion of the protractor to define a rigid occlusal plane.

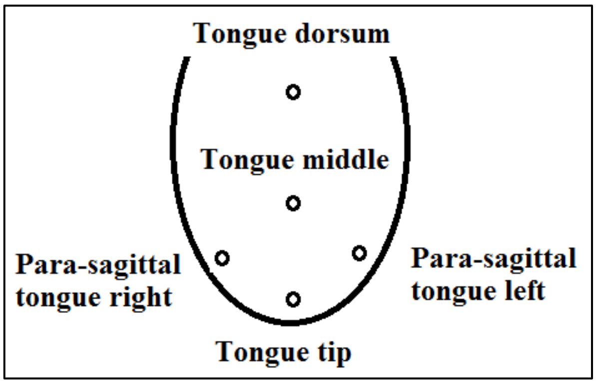


Figure 1: Tongue sensor positions in the ‘Southern Cross’ configuration designed for articulographic investigation of mid-sagittal and para-sagittal movements in speech (view of tongue from above; Ying et al., submitted).

Participants were familiarized with the target words before recording. Elicitation sentences were presented on a computer monitor placed approximately 120 cm in front of the participant, and participants were instructed to read the sentences at a comfortable speaking rate.

D. Data processing and measurements

Articulographic data were corrected for head movement and rotated into a common coordinate system: x = anterior-posterior; y = left-right; z = up-down. Sensor displacement was expressed with respect to an origin located on the speaker’s occlusal plane, along the midline and immediately behind the upper incisors. Kinematic data from the lingual sensors were filtered and smoothed using a robust DCT-based penalized least squares algorithm (Garcia, 2010). Smoothing splines (gss R package; Gu 2002) were applied to time-varying measurements derived from the kinematic data to observe general trends across tokens.

EMA data were first visualized using MVIEW, a MATLAB-based program developed by Mark Tiede at Haskins Laboratories (Tiede, 2005). MVIEW displays the positional signal of the sensors, time-aligned with the acoustic speech signal. Visualization of the data revealed that /l/ production primarily involved horizontal (x) motion of the TD sensor, and vertical (z) motion of the TM, TT and two para-sagittal tongue blade sensors, tongue blade left (TL) and tongue blade right (TR).

A set of temporal landmarks was identified in the acoustic signal to define a window in which the articulatory data could be measured. These landmarks were identified by visual inspection of acoustic waveforms and spectra in PRAAT (Boersma & Weenink, 2015), and articulatory analysis was conducted in MVIEW. The local maximum in TTz typically occurred within the period of acoustic evidence for the following /b/ closure in /ˈ(C)Vl.bət/ (syllable-coda /l/) words. In /ˈCVb.lət/ (syllable-onset /l/) words, the local maximum in TTz instead occurred before the period of acoustic evidence of the unstressed vowel /ə/. Based on these observations, Vl.b segment sequences for coda-/l/ words and Vb.l segment sequences for onset-/l/ words were both demarcated in PRAAT using the acoustic onset of stressed pre-lateral vowel and the acoustic onset of unstressed post-lateral vowel. This segmentation protocol ensured that the TTz gesture extremum (highest position, or peak) associated with /l/ production would occur within the segmentation boundaries for both coda and onset /l/ tokens. Figure 2 shows two examples of acoustic landmarks in vowel-/l/ sequences produced by female speaker F03.

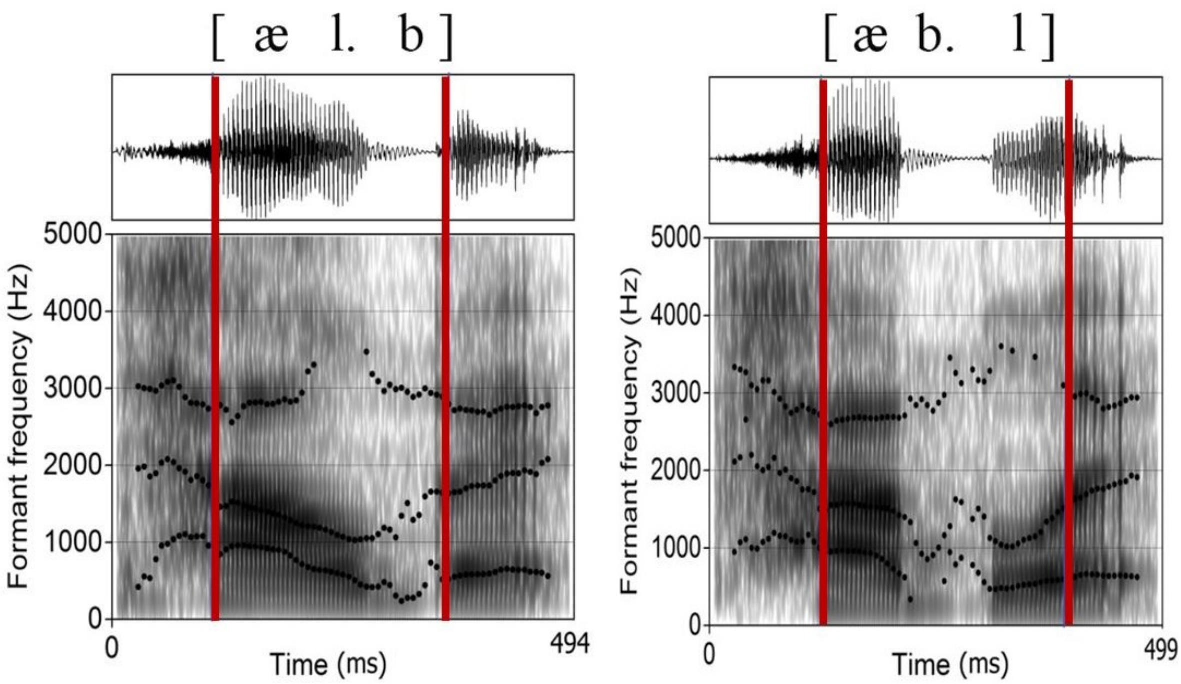


Figure 2: (Color online) Identification of Analysis Window**.** Acoustic waveform and spectrum of [æl.b] (left) and [æb.l] (right) produced by female speaker F03. Vertical red lines

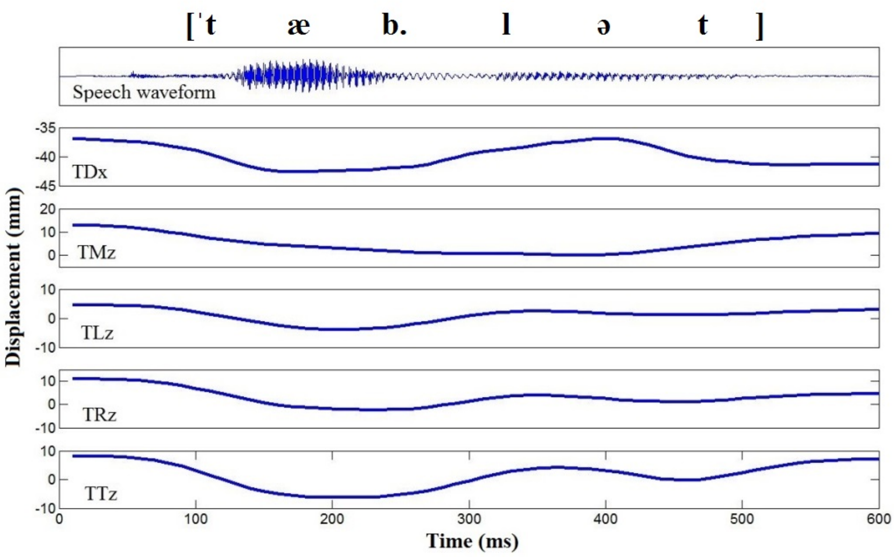
indicate the endpoint landmarks of the analysis window. Left landmark: acoustic onset of stressed pre-lateral vowel. Right landmark: acoustic onset of unstressed post-lateral vowel.

1. *Acoustic measurements*

Noise reduction was applied to all files using Audacity (Version 2.0.6). The parameters were set at 30 dB for noise reduction (dB), 1 dB for sensitivity (dB), 200 Hz for frequency smoothing (Hz), and 0.15 seconds for decay time (seconds). After reducing the noise, a three-formant estimation (F1, F2, and F3 frequency) was conducted in PRAAT using the Burg linear predictive coding (LPC) method. The F3 ceiling varies across speakers (3600, 3400 and 3900 Hz for the females, but 3200 Hz for all three males) so it was optimized independently for each speaker using a method similar to Escudero et al. (2009). F1, F2 and F3 measurements were using different ceiling heights in Praat. The ceiling heights were automatically determined for each speaker. The speaker-dependent measurements were imported to R, where the variance for each formant was calculated at each step, and the three formant variances were summed together. The summed variances were plotted, and the F3 ceiling that resulted in the lowest variance was logged (i.e., the F3 frequency ceiling that resulted in the most consistent measurements of F1, F2, and F3 frequency for the given speaker's data). The speaker-optimized ceilings were verified manually in PRAAT by comparing the resulting formant tracks against a broadband spectrogram. The formant tracks throughout the analysis window were the same identical time window as shown in Figure 2. The formant tracks were generated based on the speaker-optimized F3 ceiling frequencies. Z-score formant normalization was then carried out for each speaker.

2. *Articulatory measurements*

Temporal alignment of TD, TM, TT, TL and TR trajectories is illustrated in Figure 3, for the words 'talbot' and 'tablet'; differences in TT height and TD retraction are compared across syllable positions. In both syllable positions, the para-sagittal left and right sides of the tongue blade (TL and TR) are raised in concert with TT during /l/ production, with at least one side of the tongue blade (either TL or TR) slightly higher or lower than TT (for speaker F03). In ‘tablet’ TL is about 1.35 mm lower than TT at 350 ms. In ‘talbot’, TT is about 1.5 mm higher than TL. TDx retraction and TMz lowering can be observed in both tokens. The maximum TD retraction is about -42.6 mm at 170 ms and the maximum TM lowering is about 0.7 mm at 310 ms in ‘tablet’. This negative coordinate is relative to the occlusal plane, i.e., distance behind the occlusal plane origin. In ‘talbot’, the maximum TD retraction is about -41.8 mm at 160 ms and the maximum TM lowering is about -3.3 mm at 240 ms. Then TT, TL and TR rise gradually for the /l/ and lower slightly for the following /ə/. Thus, in syllable-coda position (tal.bot), TT is lower, TM is lower and TD is slightly less retracted than in syllable-onset position (tab.let), in which TT is raised, TM is lowered and TD is slightly more retracted.



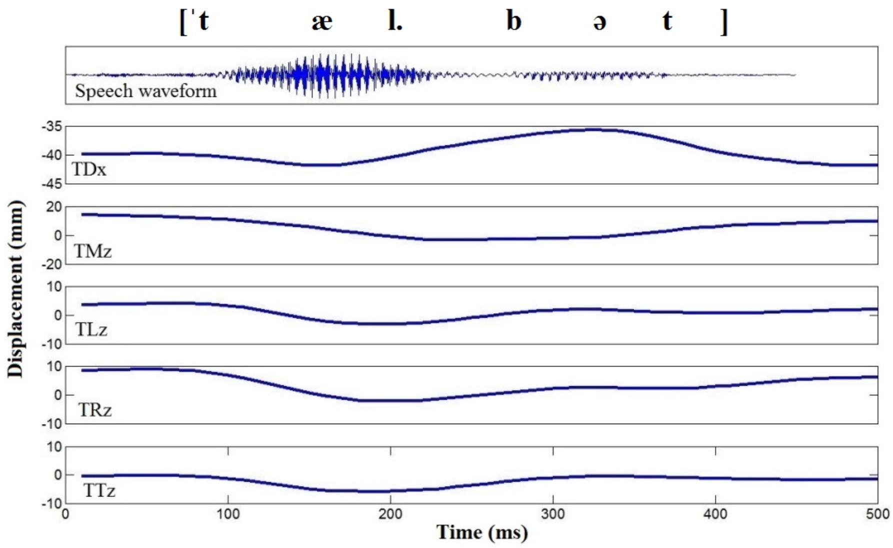


Figure 3: (Color online) /l/ articulation in onset (top panel – as shown in Table 3a) and coda (bottom panel – as shown in Table 3b) positions. Top-to-bottom in each panel: acoustic waveform; tongue dorsum movement in the anterior-posterior (x) dimension (TDx); then tongue middle (TMz), para-sagittal tongue left (TLz), para-sagittal tongue right (TRz) and tongue tip (TTz) movements in the vertical (z) dimension. Example utterances ‘tablet’ (top panel - as shown in Table 3a) and ‘talbot’ (bottom panel - as shown in Table 3b) produced by female speaker F03. x-axes: time (ms); y-axes: displacement (mm).

Partly due to previous limitations in sensing technology, and partly because of a prior focus on mid-sagittal articulation, there has been comparatively little prior research on para-sagittal dynamics of tongue movements in speech. It was for that reason that we developed the Southern Cross sensor configuration, which includes a para-sagittal sensor on each side of the tongue blade (Ying et al., submitted).

We developed novel analyses designed to address para-sagittal dynamics in /l/ production and inform our understanding of lateral channel formation, and at the same time to allow us to assess how variations in the mid-sagittal plane influence the timing of lateral channel formation. In order to estimate lateralization of the tongue blade in the coronal plane, a mid-sagittal tongue blade sensor is required. One methodological innovation in this study was that we mathematically estimated the tongue blade sensor in the mid-sagittal plane from relationships among the para-sagittal sensors and the mid-sagittal TT and TM sensors, instead of simply using the TT sensor data. The issue with using the TT sensor data alone is that the TT sensor is typically higher than the para-sagittal sensors in /l/ production, particularly when the anterior portion of the tongue rises to form a constriction for /l/. Also, the TT sensor and the para-sagittal sensors are not situated in the same sagittal or coronal planes. Estimating a mid-sagittal tongue blade sensor from the relationships among the para- and mid-sagittal sensors solves these two issues.

The steps we used to estimate the mid-sagittal tongue blade sensor are: Firstly, a second-order polynomial was fitted to the x (horizontal anterior-posterior) and z (vertical) dimensions of the three mid-sagittal sensors (TT, TM, TD), thus estimating the mid-sagittal curve of the tongue (concave, flat, or convex). Secondly, the average position of the para-sagittal tongue sensors (TR and TL) in the x dimension was used to locate an intersection point along the mid-sagittal polynomial. The x (front-back), y (side-to-side), and z (height) dimension values of this intersection point served as the estimated location of the mid-sagittal tongue blade sensor (virtual TB). This virtual sensor represents the intersection of the mid-sagittal (TT-TM-TD) and coronal planes at a location that falls midway between the two para-sagittal sensors (TR-TL). The virtual TB sensor location in the horizontal (x) plane was determined by the position of the dominant-side para-sagittal sensor for each token to ensure that all the relevant measurements used in our new para-sagittal lateralization index were located in the same coronal plane. The dominant side of tongue lateralization is the one with the lower para-sagittal sensor, as this is where the side branch characteristic of lateral anti-formants forms. We observed in our data that in each token of /l/ only one lateral channel was formed, on either left or right side of the tongue blade. While there may be a smaller side channel that admits airflow on the non-dominant side, or other speakers for whom the lateral airflow is equally bifurcated, these possibilities are not addressed by the current study.

To measure tongue lateralization in the coronal plane, we created an index referred to as ΔHeight, which captures the degree to which the dominant side of the tongue blade (the lower of the two para-sagittal sensors) differed in height from the mid-sagittal virtual TB sensor at any given point in time. During /l/ production, the speakers in this study typically tilted (i.e. roll around the x-axis) one side of the tongue blade more than the other, either left or right. Formula (1) below was used to calculate the token-by-token difference between the mid-sagittal virtual TB sensor and the dominant para-sagittal sensors in the vertical (z) dimension (1) at each time sample in the analysis. The time course of lateralization during /l/ production was indexed as the temporal relationship between the mid-sagittal virtual TB sensor and the dominant para-sagittal sensor positions over time.

$$\Delta \text{Height} = \text{virtual } TB_z - \text{dominant para-sagittal}_z \quad (1)$$

$$\text{where dominant para-sagittal} = TL_z \text{ or } TR_z$$

E. Dataset normalization process

The z-score normalization method was used to standardize the articulatory and the acoustic data, so that we can compare the dataset across speakers. The y-axis in all figures of the Results section show the normalized values for tongue lateralization, F1, F2 and F3 frequency.

III. RESULTS

A. The articulatory dynamics of tongue lateralization

Smoothing spline analysis of variance (SSANOVA) was applied to the temporal trajectories of the tongue lateralization from the acoustic onset of the stressed pre-lateral vowel to the acoustic onset of unstressed post-lateral vowel (300 ms), represented by ΔHeight. Figure 4 shows the time course of tongue lateralization (in the coronal plane) on the dominant side for /l/ tokens produced after stressed /æ/ versus /ɪ/. The y-axis shows the magnitude of tongue lateralization. In this measure, a value of zero indicates a flat tongue along the coronal plane between the para-sagittal sensor and the estimated tongue blade sensor, a positive value indicates that the dominant tongue side is lower than the midline of the tongue, and a negative value indicates that the dominant side of the tongue is higher than the midline of the tongue. The x-axis shows a time window of 800 ms, which covers the entire V-/l/ interval. The peak of tongue lateralization occurs at about 200 ms. The /l/s adjacent to /æ/ appear to achieve slightly greater magnitude of peak tongue lateralization compared to /l/s adjacent to /ɪ/. In addition, the lateralization peak appears to occur slightly earlier in /æ/ context than in /ɪ/ context. However, neither of these differences is significant, as there is considerable overlap between the peaks of the turquoise line for /ɪ/ and the rose-colored line for /æ/.

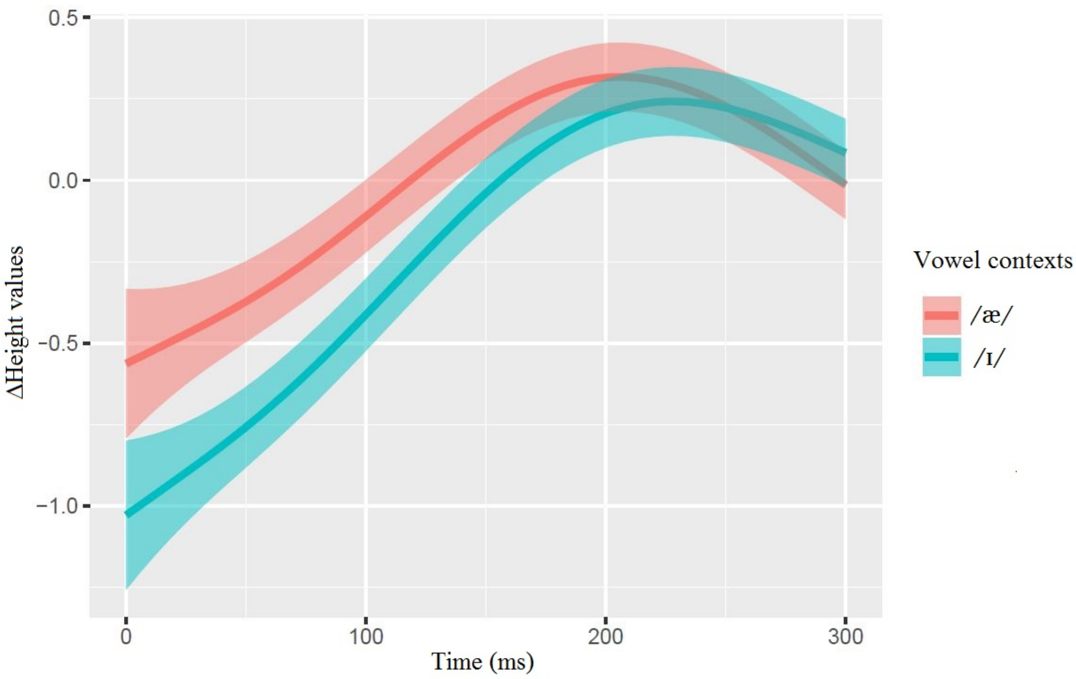


Figure 4: (Color online) Tongue lateralization (in the coronal plane) of /l/ tokens in the contexts of preceding stressed /æ/ versus /ɪ/. The y-axis shows ΔHeight (normalized tongue lateralization); higher values indicate greater lateralization. The x-axis shows a time window of 300 ms from the onset of the stressed vowel (V_1) and incorporates /l/ production in all cases. The relevant portion of the trajectories is the peaks between 180 ~ 250 ms, which indicate the timing and magnitude of maximal tongue lateralization as indexed by ΔHeight.

Figure 5 shows the temporal dynamics of tongue lateralization (ΔHeight) of the dominant side (by-token analysis) in the coronal plane over the same 300 ms interval from the onset of V_1. The ΔHeight is defined as the point in time at which the difference in height between the mid-sagittal blade and the para-sagittal blade is greatest. It differs across vowel by syllable position contexts. Higher ΔHeight values indicate greater magnitude of tongue lateralization, i.e., more lowering of dominant side of the tongue blade. The ΔHeight value of zero refers to a flat tongue shape. At the beginning of the temporal window (left side of Figure 4.5), i.e., during V_1, ΔHeight is higher, i.e., the dominant side of the tongue blade is already lower, for /æ/ preceding the onset /l/ (green line) than for the other three vowel by syllable position contexts. The coda /l/ following /ɪ/ (turquoise line) reaches its ΔHeight peak at roughly 250 ms, whereas the onset /l/ following /æ/ (green line), coda /l/ following /æ/ (rose-colored line), and onset /l/ following /ɪ/ (purple line) reach their ΔHeight peak at about 200 ms.

The maximum degree of lateralization obtained also varies across vowel by syllable position contexts. The coda /l/ following /æ/ (rose-colored line) and the onset /l/ following /ɪ/ (purple line) have a lower ΔHeight peak than the other two contexts, and their trajectories are almost overlapped. The onset /l/ following /æ/ (green line) has the highest ΔHeight peak among all four contexts. That of the coda /l/ following /ɪ/ (turquoise line) is lower than the onset /l/ following /æ/ (green line), but higher than the coda /l/ following /æ/ (rose-colored line) and the onset /l/ following /ɪ/ (purple line).

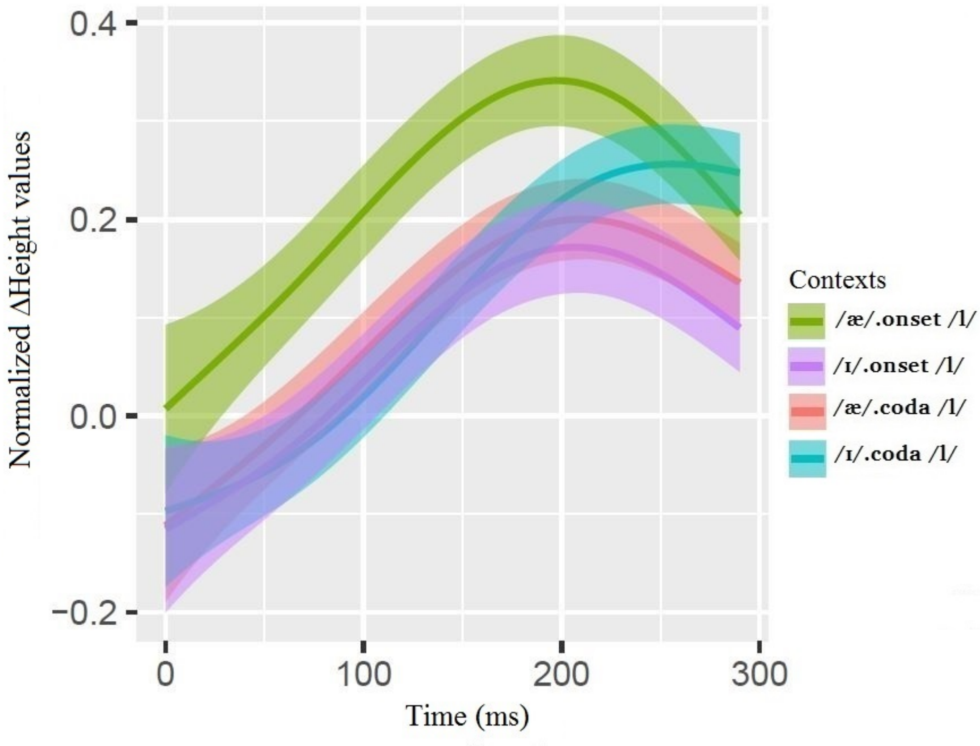


Figure 5: (Color online) Temporal dynamics of tongue lateralization on the dominant side of the tongue blade in the coronal plane. The x-axis shows the time window of 300 ms from the onset of V_1. The y-axis shows ΔHeight (normalized tongue lateralization). A value of zero indicates a flat tongue shape along the coronal plane between the para-sagittal sensor and the mid-sagittal tongue blade virtual sensor. A positive value indicates the dominant side of the tongue blade is lower than the mid-sagittal tongue blade virtual sensor.

B. The acoustic measurements of tongue lateralization

1. F1 frequency

Figure 6 shows the temporal dynamics of F1 frequency over the 300 ms interval from the onset of V_1. According to the ΔHeight results, the acoustic peak of lateralization for /l/ occurs between 200 to 300 ms. The F1 peaks in the /l/ are higher for coda /æ/ and /ɪ/ (rose-colored

424 line and turquoise line) than the rising but non-peaked values for onset /l/s following /æ/ and

425 /ɪ/ (purple line and green line). There is a steady drop following the peak for the coda /l/s.

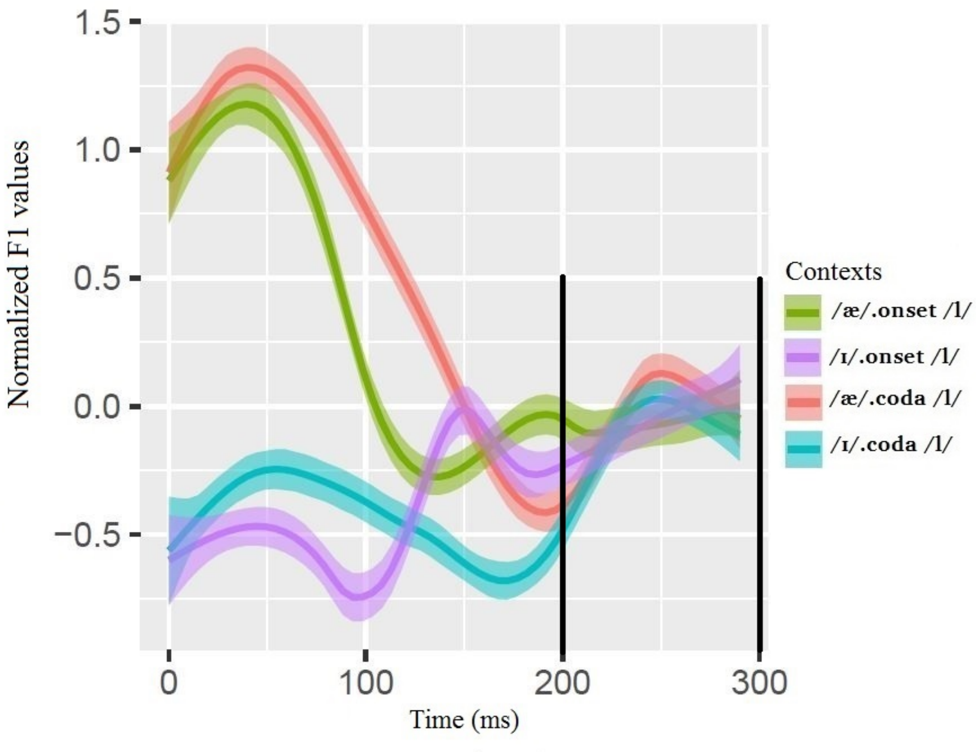


Figure 6: (Color online) Temporal dynamics of F1 frequency. The x-axis shows the time window of 300 ms from the onset of V_1. The y-axis shows normalized F1 values. Black lines on the x-axis indicate the target time window.

426 ***2. F2 frequency***

427 Figure 7 shows the temporal dynamics of F2 frequency over the same 300 ms interval. Based

428 on our ΔHeight results, the acoustic peak for /l/ occurs after 200 ms. Coda /l/s following /æ/

429 and /ɪ/ (rose-colored line and turquoise line) have higher peaks than onset /l/s following the

430 two vowels (purple line and green line). The F2 values of the two peaks (rose-colored line

431 and turquoise line) are quite similar, and they remain roughly constant.

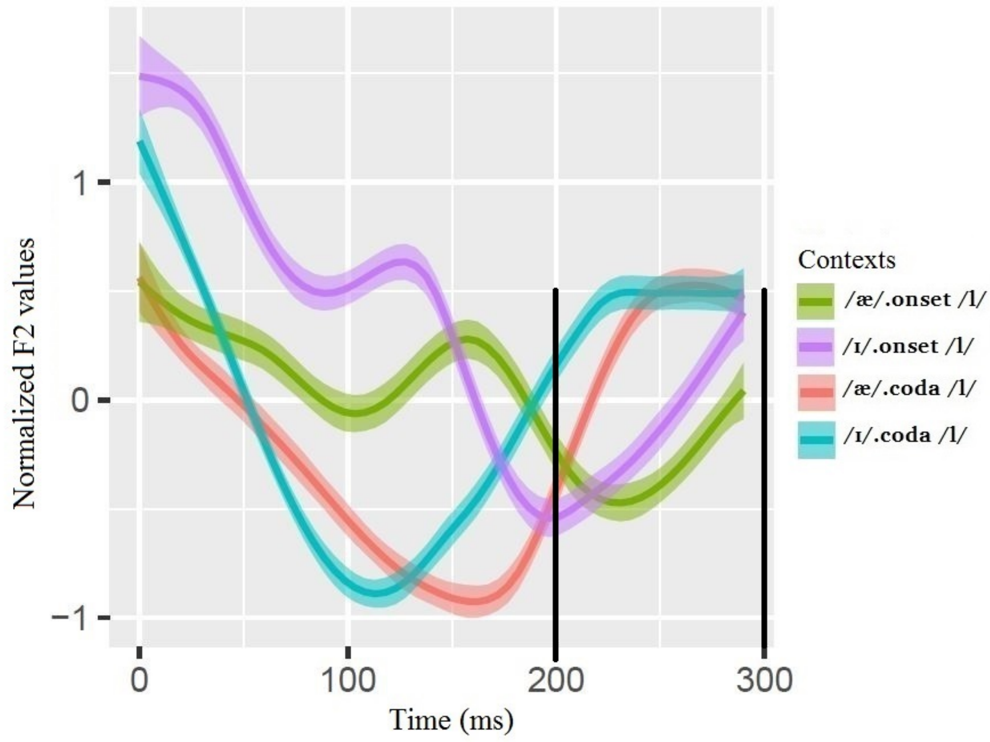


Figure 7: (Color online) Temporal dynamics of F2 frequency. The x-axis shows the time window of 300 ms from the onset of V_1. The y-axis shows normalized F2 values. Black lines on the x-axis indicate the target time window.

3. *F3 frequency*

Figure 8 shows the temporal dynamics of F3 frequency over the same 300 ms time window. The target time window is between 200 to 300 ms. The F3 peaks for onset /l/s following /æ/ (green line) and /ɪ/ (purple line) occur at about 230 ms and 200 ms, respectively. Coda /l/s following both vowels (rose-colored line and turquoise line) have a sharp drop during the target time window.

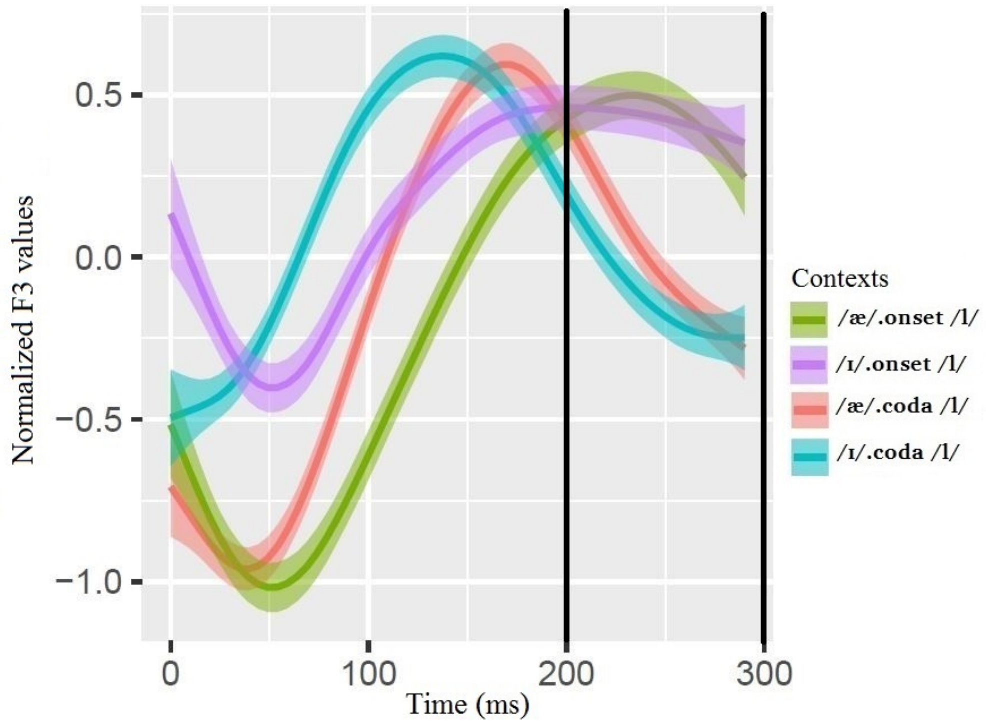


Figure 8: (Color online) Temporal dynamics of F3 frequency. The x-axis shows the time window of 300 ms from the onset of V_1. The y-axis shows normalized F3 values. Black lines on the x-axis indicate the target time window.

C. Statistical analysis of articulatory-acoustic relationships

To evaluate the reliability of the descriptive results above, three linear mixed effect (LME) models (one each for F1, F2 and F3 relationships to ΔHeight) were fit to the normalized ΔHeight values using the lme4 package (Bates, Maechler, Bolker, Walker, 2014) in Rstudio Version 3.6.1. The following models were tested:

F1 ~ ΔHeight + vowel*syllable + (1|speaker) + (1|item)

F2 ~ ΔHeight + vowel*syllable + (1|speaker) + (1|item)

F3 ~ ΔHeight + vowel*syllable + (1|speaker) + (1|item)

The Bonferroni-adjusted significance level was set to 0.017, for comparison across the three models. The vowel context (/æ/ and /ɪ/) and syllable position (onset and coda) were included as fixed effects. Random intercepts were included for speaker and item (i.e., token order) in each model. Estimates for t-statistics and p-values were generated using Satterthwaite approximation in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016). Model comparison was conducted by back-fitting along the Akaike Information Criterion (AIC) to measure quality of fit. All models were conducted with a three-way interaction term (ΔHeight*vowel*syllable). If a three-way-interaction term and a two-way-interaction term were both significant, then the three-way-interaction term and the two-way-interaction term were compared using analysis of variance (ANOVA) models with a Chi-square test to determine whether the reductions in the residual sum of squares are significant. If the two models did not differ significantly, then the simpler model was retained.

In order to examine which formant is best predicted by tongue lateralization, we conducted a relative importance analysis (RIA) in R (Grömping, 2006). The analysis was performed using calc.relimp function in R. The function calculates several relative importance metrics for linear models using a method (lmg) developed by Lindemann, Merenda and Gold (1980). lmg calculates the relative contribution of the predictor to the R2. According to Lindemann, Merenda and Gold (1980, p119), lmg is the R2 contribution averaged over ordering among predictors. The proportion of the variance is represented by R2. A larger R2 means that the predictor is more important to explain the outcome variable.

1. Linear mixed effect models

Table III shows the results of the LME model of normalized F1 frequency (normalized tongue lateralization). The main effect of ΔHeight is significant with a negative estimate (β=-0.057). This means that the normalized F1 frequency and ΔHeight have a systematic and inverse relationship as shown in Figure 9. As ΔHeight increases, F1 frequency decreases. The other two main effects (vowel and syllable) are significant as well. Vowel has a negative estimate (β = -0.738), which indicates that F1 frequency is significantly lower for /ɪ/ than /æ/. Syllable also has a negative estimate (β = -0.152), which indicates that F1 frequency is significantly lower for onset than coda /l/s (see Figure 9). There is a significant positive interaction (β = 0.167) between vowel and syllable, which indicates that the difference in F1 between /ɪ/ and /æ/ is greater for onset than coda /l/s.

TABLE III: Results of LME model on the relationship between F1 and ΔHeight.

Main effects and interactions	β	S.E.	t value	Pr (>|t|)
Intercept	0.324	0.120	2.720	0.040
ΔHeight	-0.057	0.014	-4.105	<0.001
Vowel	-0.738	0.026	-28.382	<0.001
Syllable	-0.152	0.029	-5.196	<0.001
ΔHeight * Vowel	0.080	0.020	3.926	<0.001
ΔHeight * Syllable	-0.023	0.022	-1.046	0.296
Vowel /ɪ/ * Syllable	0.167	0.041	4.024	<0.001
ΔHeight * Vowel * Syllable	0.060	0.033	1.793	0.070

(Bonferroni-adjusted significance 0.017)

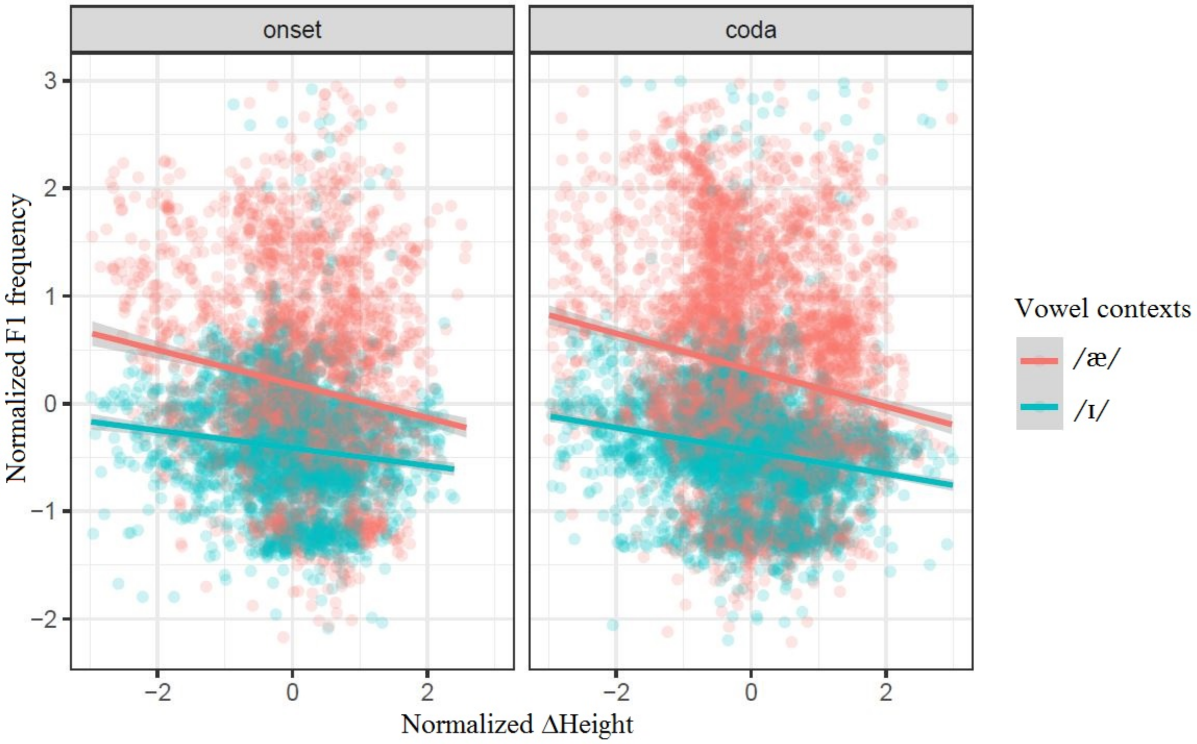


Figure 9: (Color line) The relation of F1 frequency and ΔHeight. The x-axis shows ΔHeight (normalized tongue lateralization); the y-axis shows the normalized F1 frequency (all the data points are within three standard deviations). Vowel context is separated by color and syllable position is separated by plots (left panel: onset position; right panel: coda position).

Table IV shows the LME model results on the relationship between normalized F2 frequency and ΔHeight. There was a modest overall relationship between ΔHeight and F2, as reflected in the borderline significance of the main effect for ΔHeight ($p = 0.017$). Both vowel and syllable effects are significant; both have positive estimates ($\beta = 0.199$ and $\beta = 0.246$, respectively). F2 frequency is higher for /ɪ/ than /æ/, and it is higher for onset than coda /l/s (see Figure 10). The interaction between ΔHeight and vowel was significant with a positive estimate ($\beta = 0.071$), indicating that the relationship between F2 frequency and ΔHeight is more strongly positive when the preceding vowel is /ɪ/ than /æ/. The three-way interaction (ΔHeight, vowel and syllable) was also significant. The negative estimate ($\beta = -0.140$) shows that the relationship between F2 and ΔHeight are different in onset /l/s compared to coda /l/s: when /l/s are in onset position, the relationship is more negative for both vowels; but when /l/s are in coda position, the relationship is less negative for both vowels.

TABLE IV: Results of the LME model on the relationship between F2 and ΔHeight.

Main effects and interactions	β	S.E.	*t* value	Pr (>|t|)
Intercept	-0.301	0.126	-2.399	0.060
ΔHeight	0.038	0.016	2.379	0.017
Vowel	0.199	0.028	7.163	<0.001
Syllable	0.246	0.031	7.882	<0.001
ΔHeight * Vowel	0.071	0.023	3.080	0.002
ΔHeight * Syllable	-0.058	0.026	-2.263	0.023
Vowel * syllable	0.087	0.044	1.960	0.050
ΔHeight * Vowel * Syllable	-0.140	0.038	-3.714	<0.001

(Bonferroni-adjusted significance 0.017.)

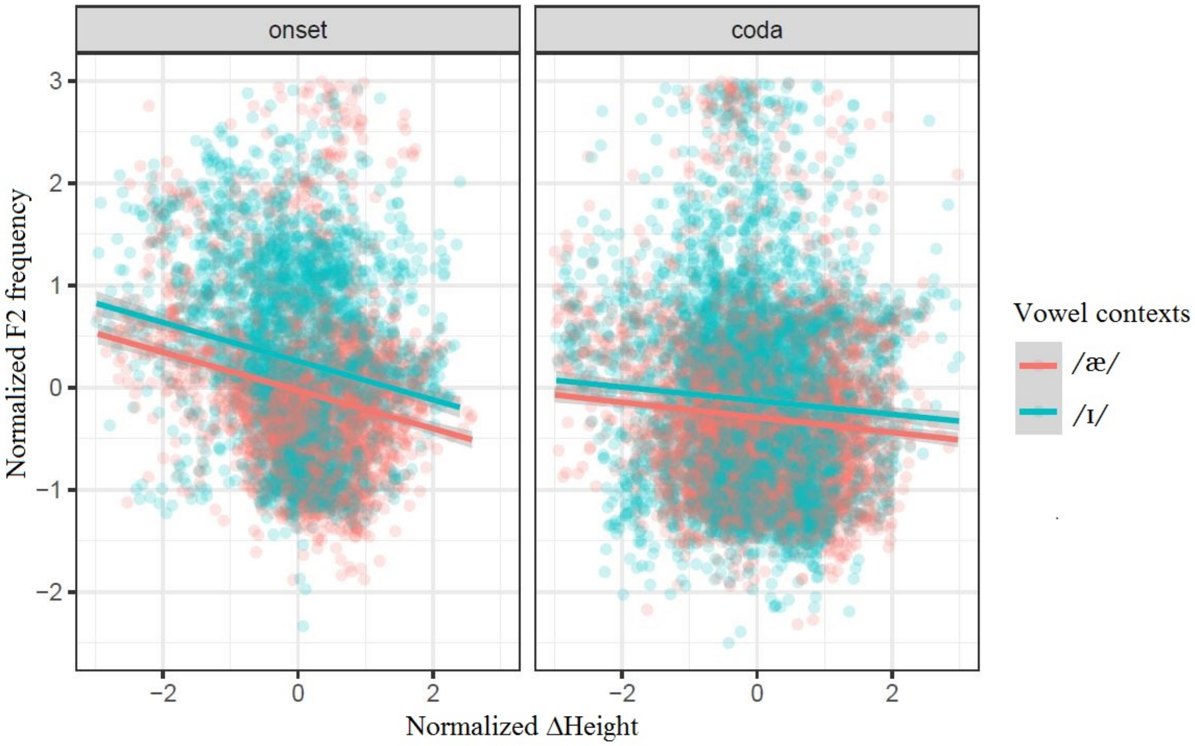


Figure 10: (Color online) The relation of F2 frequency and tongue lateralization. The x-axis shows ΔHeight (normalized tongue lateralization); the y-axis shows the normalized F2 frequency (all the data points are within three standard deviations). Vowel context is separated by color and syllable position is separated by plots (left panel: onset position; right panel: coda position).

Table V shows the results of the LME model on the relationship between normalized F3 frequency and ΔHeight. The main effect of ΔHeight is significant, with a positive estimate ($\beta = 0.079$), indicating a systematic positive relationship between ΔHeight and F3. The main effect of vowel is also significant with a positive estimate ($\beta = 0.262$), which indicates that F3 frequency is significantly higher for /ɪ/ than /æ/ (See Figure 11). The three-way interaction among ΔHeight, vowel and syllable is significant, and has a positive estimate ($\beta = 0.097$). The relationship of F3 to ΔHeight for /ɪ/ is more positive than that for /æ/. The vowel difference is more positive in onset /l/s.

TABLE V: Summary of the relation of F3 frequency and tongue lateralization.

Main effects and interactions	β	S.E.	*t* value	Pr (>|t|)
Intercept	-0.136	0.301	-0.449	0.671
ΔHeight	0.079	0.016	5.043	<0.001
Vowel	0.262	0.034	7.561	<0.001
Syllable	-0.044	0.039	-1.138	0.256
ΔHeight * Vowel	0.012	0.023	0.516	0.606
ΔHeight * Syllable	0.029	0.025	1.155	0.248
Vowel * Syllable	0.038	0.055	0.683	0.495
ΔHeight * Vowel * Syllable	0.097	0.037	2.604	0.009

(Bonferroni-adjusted significance 0.017.)

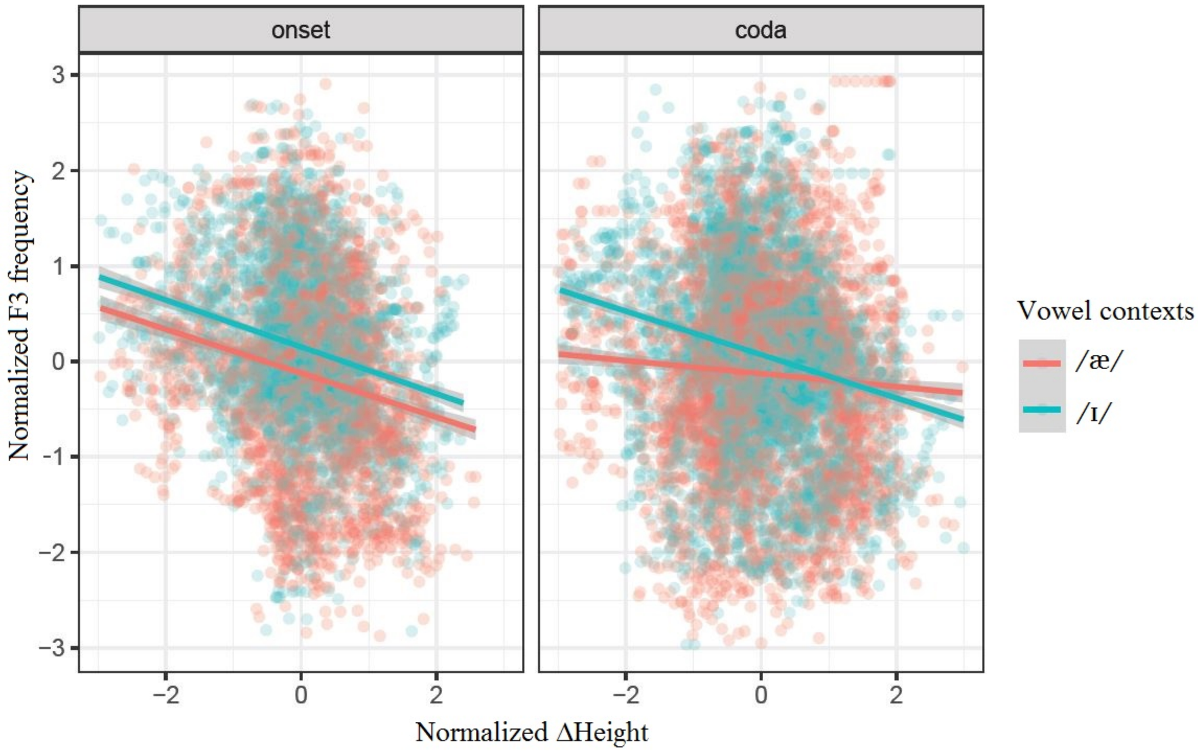


Figure 11: (Color online) The relation of F3 frequency and tongue lateralization. The x-axis shows ΔHeight (normalized tongue lateralization); the y-axis shows the normalized F3 frequency (all the data points are within three standard deviations). Vowel context is separated by color and syllable position is separated by plots (left panel: onset position; right panel: coda position).

2. Relative importance analysis

Given that tongue lateralization (ΔHeight) is a strong predictor of both F1 and F3 frequencies, a relative importance analysis was performed. In this case, tongue lateralization was modelled using the formants (F1, F2, and F3 frequencies). The following model was tested: lm (ΔHeight ~ F1 + F2 + F3). The proportion of variance explained by the model is about 6.32%. Metrics are normalized to sum to 100%. The relative importance metrics are: 27% for F1 frequency, 21% for F2 frequency, and 52% for F3 frequency. As mentioned earlier, a predictor with greater proportion means that the predictor is more important. Therefore, tongue lateralization predicts F3 values most strongly and F2 values least strongly.

IV. DISCUSSION

In this study, the relationship between the acoustics and the articulation of /l/ was examined. The normalized (z-score) frequency values of F1, F2 and F3 frequency were compared to the normalized (z-score) ΔHeight, an articulatory measure of tongue lateralization. We made four predictions: 1) tongue blade lateralization will be inversely correlated with F1 frequency; 2) tongue blade lateralization and F3 frequency will be positively correlated; 3) tongue blade lateralization will be less related with F2 frequency than it is to F1 and F3 frequency; and 4) the effect of syllable position on F1 and F3 frequencies to tongue blade lateralization will be minimal. Statistical analyses were conducted to examine each of these aspects of acoustic-articulation relationships.

Prediction 1, 2, and 3 were upheld; however, Prediction 4 was not supported. The statistical analyses show that tongue blade lateralization, as indexed by our novel measure ΔHeight, was inversely correlated with F1 frequency, supporting Prediction 1. As tongue blade lateralization increases, F1 frequency decreases. F3 frequency and the tongue lateralization show a strong positive correlation, supporting Prediction 2. Compatible with Prediction 3, the relationship between tongue blade lateralization and F2 frequency was less robust, only showing borderline significance. The data on vowel effects was at odds with Prediction 4, however. The effect of syllable position affects F1 frequency, but not F3 frequency. In order to determine which formant is most strongly affected by tongue lateralization, a relative importance analysis was conducted. The results show that tongue lateralization is a stronger predictor of F3 frequency than F1 frequency for the /l/ data in this study.

Previous studies (Giles & Moll, 1975; Espy-Wilson, 1992; Huffman, 1997) have reported a systematic and consistent inverse correlation between F1 frequency and tongue tip height. In addition to this, vowel context also has an effect on /l/ production. Figure 4.9 illustrates that /ɪ/ contexts condition lower F1 frequencies for both onset and coda /l/, compared to /æ/ contexts. This is because /ɪ/ is produced with a raised tongue tip – i.e., lower F1 – while /æ/ is produced with a lower tongue tip, i.e. higher F1. In terms of the syllable position effect, onset /l/s tend to have lower F1 values compared to coda /l/s.

F2 frequency is traditionally considered to be associated with the front-back movement of tongue body. A retracted tongue body usually has low F2 frequency. An advanced tongue body usually has high F2 frequency. In most varieties of English, onset /l/s are produced with a less retracted tongue body than coda /l/s. Therefore, onset /l/s tend to have higher F2 frequency than coda /l/s (Bladon, 1976; Recasens, 2012). Since tongue lateralization occurs mostly in the blade, which is the front portion of the tongue, while front-back positioning of the tongue is primarily associated with tongue body movement, this could explain why F2 has only a borderline relationship to tongue blade lateralization.

As for F3 frequency, Fant (1960) suggests that it is associated with the front cavity anterior to the closure location. Recasens (2012) reported that onset /l/s in Catalan tend to have lower F3 frequency than coda /l/s. However, as our data suggest, the relationship between syllable position and formant frequency are likely to be affected by adjacent vowel context. Figure 4.11 shows that onset /l/s have higher F3 frequency than coda /l/s. Our statistical model shows a significant relationship between F3 frequency and ΔHeight, and the model reveals a strong relationship between vowel and ΔHeight. We also found a complex three-way interaction among ΔHeight, vowel context and syllable position in the F3 LME model. This interaction might possibly reflect individual speaker variation in lateral channel formation, affecting anti-formants in the lateral channel. During /l/ production, a central constriction is formed at the dental place of articulation. This constriction has an important consequence: it traps a pocket of air. This pocket of air is usually considered to be the main source of anti-formants. Since the main airflow is also behind the central constriction, the anti-formants will weaken any acoustic energy of the cavity. As observed in past studies (Bangayan et al., 1996; Narayanan & Alwan, 1996), lateral channel asymmetry can also give rise to anti-formants, which may vary from speaker to speaker. Overall, the left and right lateral channels are unequal. The area(s) of the lateral channel(s) start increasing behind the linguo-alveolar contact, and start decreasing as the lateral channel(s) gets close to linguo-velar contact. Those authors speculated that the area increase is due to tongue lateralization. The anti-formants created by the lateral channel can be observed in the higher formants (usually F3 ~ F5). Anti-formants absorb energy and weaken the signal in these frequency regions.

Additional anti-formants can be formed at the sublingual cavity. The sublingual cavity is formed by linguo-alveolar contact in /l/ articulation. Zhou (2009) summarized that anti-formants could be produced in the following three situations: (1) The length of the lateral channel is short; (2) The lateral channels are asymmetrical; (3) the linguo-alveolar constriction is not narrow enough, thereby increasing the sublingual cavity. Relatedly, Narayanan and Alwan (1996) found that the sublingual cavity can predict F3 frequency.

Specifically, it reduces the frequency associated with the front cavity as the volume of the front cavity increases.

The earlier studies we have reviewed can be divided into two categories: 1) mid-sagittal tongue movements and their resulting acoustics on /l/ production; 2) lateral channel(s) formation on static /l/ production. Neither of these two categories covers the topic of the effect of active lateral channel(s) formation on the acoustics. The present study fills that empirical gap. We presented dynamic data on the lateral channel formation and its resulting acoustics. One of the most important findings in this study is that unlike F1 and F2, there is no effect of syllable position on F3. Typically, F1 and F2 frequencies are associated with the mid-sagittal tongue movements: F1 is related with tongue tip height and F2 is related with tongue body backness. The values of F1 and F2 frequencies show specific pattern depending on the syllable position of the /l/ (onset vs. coda). In terms of F3 values, we found that it is related with tongue lateralization. While F1 and F2 values vary across syllable positions, F3 value remains constant across different positions. In our previous study (Ying et al., submitted), we found that lateral channel formation show stability across syllable position and vowel context. This means that lateralization is an actively controlled gesture; even as the mid-sagittal gestures vary across syllable positions. Context-independent control is the hallmark of gestures as phonological units. Stevens (1972) observed that the acoustics are insensitive to changes in the articulation over a part of its range, whereas at the other part of its articulatory range the acoustics can change rapidly, and at another part the acoustics can be insensitive again. The insensitive region of acoustics is known as the stable region, and the rapid change region is unknown as the unstable region.

According to Stevens (1972), this type of articulatory and acoustic relation defines a distinctive feature. Sproat and Fujimura (1993) proposed a [+lateral] feature for /l/. They speculated that tongue tip gesture and tongue blade narrowing gesture are in active control during /l/ production. They further postulated that tongue body retraction is a consequence of these two gestures resulting from the volume-preserving nature of the tongue. In our previous study (Ying et al., submitted), we found that all the three gestures (tongue tip gesture, tongue blade gesture and tongue dorsum gesture) are actively controlled. The tongue body retraction is not simply a secondary articulation. Evidences from coda /l/s show that the tongue body gesture occurs prior to the tongue tip gesture in the mid-sagittal plane. In terms of articulatory-acoustic relation, we found that the constant characteristics of the tongue blade gesture and F3 frequency are the defining attributes of /l/. Since most articulatory data on /l/s have revealed that there are multiple gestures (tongue tip raising and tongue body retraction) in the mid-sagittal plane, and these two gestures are typically associated with F1 and F2 frequencies, these findings on tongue lateralization and the resulting acoustics deepens our understanding of /l/ articulation.

V. Conclusion

In this study, relationships between articulatory and acoustic properties of Australian English /l/ have been examined in new detail, by exploring the timecourse of lateral articulation and the acoustic consequences for formant frequencies. We found that degree of tongue lateralization – the relative height of the sides of the tongue compared to midsagittal blade height – is a strong predictor of F3 frequency. It remains to be seen if this relationship holds true for /l/ in other varieties of English and in other languages. More research needs to be conducted to examine the reliability of F3 frequency as an index of tongue lateralization.

Acknowledgments

This work was supported by a PhD Scholarship from the MARCS Institute, Western Sydney University. Special thanks to the thesis examiners for their comments on an earlier version of this paper as a thesis chapter.

References

Bladon, R. A., & Al-Bamerni, A. (**1976**). “Coarticulation resistance of English /l/,” J. Phon. **4**, 135-150.

Espy-Wilson, C. Y., Narayanan, S., Boyce, S. E., & Alwan, A. (**1997**). “Acoustic modelling of American English /r/,” In Fifth European Conference on Speech Communication and Technology.

Fant, G. (**1960**). Acoustic theory of speech production. The Hague: Mouton.

Huffman, M. K. (**1997**). “Phonetic variation in intervocalic onset/l/’s in English,” J. Phon. **25**(2), 115-141.

Ladefoged, P., & Johnson, K. (**2014**). A course in phonetics. Nelson Education.

Ladefoged, P, & Maddieson, I. (**1996**). The sounds of the world’s languages. Malden, MA (USA): Blackwell Publishing.

Lehiste, I. (**1964**). “Acoustical characteristics of selected English consonants. Indiana research centre in anthropology,” folklore and linguistics, **34**, 10–50.

Lin, S. (**2011**). “Production and perception of prosodically varying inter-gestural timing in American English laterals,” Ann Arbor, MI: University of Michigan dissertation.

Lin, S., Beddor, P. S., & Coetzee, A. W. (**2014**). “Gestural reduction, lexical frequency, and sound change: A study of post-vocalic /l/,” LabPhon. **5**(1), 9-36.

Lindblad, P., & Lundqvist, S. (**2003**). “[l] tends to be velarised, apical as opposed to laminal, and produced with a low jaw, and these features are connected,” In Proceedings of the 15th International Congress of Phonetic Sciences: Barcelona, 3-9 August 2003. Vol. 1/3 (pp. 1899-1902). Casual productions.

Lindeman, R. H. (**1980**). "Introduction to bivariate and multivariate analysis," No. 04; QA278, L553.

Narayanan, S., & Alwan, A. (**1996**). "Parametric hybrid source models for voiced and voiceless fricative consonants," In 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings (Vol. 1, pp. 377-380). IEEE.

Narayanan, S. S., Alwan, A. A., & Haker, K. (**1997**). "Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals," J. Acoust. Soc. Am. **101**(2), 1064-1077.

Narayanan, S., Byrd, D., & Kaun, A. (**1999**). "Geometry, kinematics, and acoustics of Tamil liquid consonants," J. Acoust. Soc. Am. **106**(4), 1993-2007.

Proctor, M. (**2011**). "Towards a gestural characterization of liquids: Evidence from Spanish and Russian," LabPhon, **2**(2), 451-485.

Recasens, D., & Espinosa, A. (**2005**). "Articulatory, positional and coarticulatory characteristics for clear /l/and dark /l/: evidence from two Catalan dialects," J. Int. Phon Assoc. **35**(1), 1-25.

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

Stevens, K. (**1998**). Acoustic phonetics, Cambridge.

Stone, M., Faber, A., & Cordaro, M. (**1991**). "Cross-sectional tongue movement and tongue-palate movement in [s] and [sh] syllables," In Proceedings of the 13th International Congress of Phonetic Sciences, Stockholm, Sweden (pp. 354-357).

Stone, M., Faber, A., Raphael, L., and Shawker, T. (**1992**). ''Cross-sectional tongue shapes and linguopalatal contact patterns in [s], [sh] and [l] syllables,'' J. Phon. **20**, 253–270.

Grömping, U. (**2006**). "Relative Importance for Linear Regression in R: The Package relaimpo," J. Stat. Softw, **17**(1), 1--27.

Ying, J., Shaw, J., Kroos, C, Best, C. (**2012**). “Relations between acoustic and articulatory measures of /l/,” Australasion International Conference on Speech Science and Technology. Sydney, Australia.

Ying, J., Shaw, J., Proctor, M., Carignan, C., Derrick, D., and Best, C. (submitted). “Para-sagittal dynamics in lateral /l/ production: Three-dimensional electromagnetic articulography data from Australian English,” Manuscript submitted for publication.