Temporal Dynamics of Lateral Channel Formation in /l/: 3D EMA Data from Australian English

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

[1]MARCS Institute, Western Sydney University, Australia
[2]Department of Linguistics, Yale University, USA
[3]Institute of Language, Brain & Behaviour, University of Canterbury, New Zealand
[4]Department of Linguistics, Macquarie University, Australia
[5]Haskins Laboratories, USA

`j.ying@westernsydney.edu.au, c.carignan@westernsydney.edu.au, jason.shaw@yale.edu,`
`michael.proctor@mq.edu.au, donald.derrick@canterbury.ac.nz, c.best@uws.edu.`

Abstract

This study investigated the dynamics of lateral channel formation of /l/ in Australian-accented English (AusE) using 3D electromagnetic articulography (EMA). Coils were placed on the tongue both mid-sagitally and para-sagitally. We varied the vowel preceding /l/ between /ɪ/ and /æ/, e.g., *filbert* vs. *talbot*, and the syllable position of /l/, e.g., /ˈtæl.bət/ vs. /ˈtæb.lət/. The articulatory analyses of lateral /l/ show that: (1) the mid-sagittal delay (from the tongue tip gesture to the tongue middle/tongue back gesture) changes across different syllable positions and vowel contexts; (2) the para-sagittal lateralization duration remains the same across syllable positions and vowel contexts; (3) the lateral formation reaches its peak earlier than the mid-sagittal gesture peak; (4) the magnitude of tongue asymmetrical lateralization is greater than the magnitude of tongue curvature in the coronal plane. We discuss these results in light of the temporal dynamics of lateral channel formation. We interpret our results as evidence that the formation of the lateral channel is the primary goal of /l/ production.

Index Terms: speech production, articulatory kinematics, lateral channel formation, lateral approximants

1. Introduction

Laterals exhibit complicated patterns of articulation. Most articulatory data on lateral /l/ has revealed that there are multiple gestural targets in the mid-sagittal plane [1, 2]. During the production of /l/, the tongue tip (TT) is raised, the tongue middle (TM) is lowered and the tongue back (TB) is retracted in the mid-sagittal plane. The sequence of these primary gestures varies systematically according to syllable position and vowel context. In syllable-onset position, the TT raising gesture and the TB retraction gesture tend to be roughly synchronous; whereas in syllable-coda position, the TB retraction gesture tends to precede the TT raising gesture. From some recent articulatory studies [3, 4, 5], we know that production of /l/ in American English also involves lateralization of the tongue. However, the formation of the lateral channel in the para-sagittal plane has not been studied.

It is unclear from these studies whether the lateral channel is the consequence of the two mid-sagittal gestural targets (TT

raising gesture and TB retraction gesture); or whether the formation of the lateral channel is the goal of the mid-sagittal gestural target. A more complete understanding of lateral production requires knowledge of lateral channel formation.

To investigate the temporal dynamics of lateral channel formation during /l/ production, we conducted a 3D EMA study making use of both mid-sagittal and para-sagittal articulographic data. Our materials varied the vowel preceding /l/, either /æ/ or /ɪ/, and the syllabic position of /l/, either in syllable onset or coda position. These vowels were chosen because of the different constraints they place on the shape of the tongue preceding /l/. The vowel /æ/ requires the opposite parasagittal configuration as /l/. It usually has a complete groove tongue shape in the coronal plane, such that the sides of the tongue are curved up (instead of curved down in laterals). A vowel like /ɪ/, in contrast, does not conflict with the tongue shape required for /l/ to the same degree. According to [5], TT and TM are raised during /ɪ/ production. The tongue shows a convexity in the coronal plane and a concavity in the tongue back. By varying /l/ in the context of these two vowels, we examine how local variation in tongue shape impacts the temporal dynamics of lateral channel formation. By varying /l/ in the different syllable positions, we investigate how syllable boundaries affect the formation of the lateral channel.

According to [1, 2], we identified two alternative hypotheses, specifically that (1) the formation of the lateral channel is the *consequence* of the mid-sagittal gestural targets that are the goal of /l/ production, in which case the timing of the mid-sagittal gestural targets should be consistent across vowels but the dynamics of the formation of the lateral channel should vary across vowels; or that (2) the formation of the lateral channel is the *goal* of the mid-sagittal gestural targets, in which case the timing of the mid-sagittal gestural targets should instead vary as a function of the preceding vowel while the dynamics of the formation of the lateral channel should be consistent across vowels.

2. Methods

2.1. Participants

Six monolingual Australian English speakers participated in this study. All the participants were born in Sydney, Australia

and raised in an environment in which general Australian English was spoken. The participants' average age was 21 years old. There were paid for the experiment. All participants reported that they had no speech impairments.

2.2. Stimuli

Target words for this experiment contained lateral /l/ in syllable-onset position (as in a CVb.lVC frame) or in syllable-coda position (as in a CVl.bVC frame). We used an intervening bilabial /b/ to minimize lingual coarticulatory effects on the /l/ and surrounding vowels. All target words (Table 1) were read in the carrier phrase "keep __ here". The /l/ was preceded by either /æ/ or /ɪ/. The stimuli were presented in a 10-word block with 10 repetitions in random order. The numbers of blocks and valid tokens vary from speaker to speaker.

Table 1. *A list of target words*

Tongue Shape	Complete Groove	Front Raising
Vowel	*æ*	*ɪ*
CVb.LVC	*tab.let*	*tib.let*
	cab.let	*kib.let*
(C)VL.bVC	*tal.bot*	*til.bert*
	cal.bert	*kil.bert*
	al.bert	*il.bert*

2.3. Procedure

Words were recorded at the Analysis of Human Articulatory Actions (AHAA) laboratory of the MARCS Institute for Brain, Behaviour and Development at Western Sydney University. Mid-sagittal and para-sagittal tongue movements were recorded at a rate of 100 Hz using an NDI Wave electromagnetic articulograph speech research system. Synchronized companion speech audio was recorded at a 22050 Hz sampling rate using a Schoeps Colette Series Supercardioid shotgun microphone and EURORACK UB802 preamplifier. Flesh point locations on tongue, lips and jaw were tracked over time. Three sensors were affixed mid-sagitally to the tongue tip (TT: ~5 mm from the apex), tongue back (TB: placed as far back on the tongue as the participant felt comfortable while speaking), tongue middle (TM; halfway between TT and TB) and another two sensors (para-sagittal tongue left and right were affixed to the sides of the tongue blade (PTL and PTR: ~5 mm in from the edge of the tongue). Figure 1 provides a schematic of tongue sensor placements. Sensors were also affixed to the lower jaw on the gum below the incisor, upper lip, and lower lip and to the left mastoid (LM), right mastoid (RM), and the nasion (NA). LM, RM and NA sensors are used for reference and allow for post-hoc correction of head motion. The occlusal plane was located by having speakers hold a triangular protractor between their teeth. Three sensors were attached to the corners of the protractor to define a rigid plane (between the upper and lower teeth in the transverse plane, immediately posterior to the incisors in the sagittal plane, and between the left and right upper incisors in the coronal plane).

Participants were familiarized with the word list before recording. After attaching the sensors, the recording session started. The microphone was placed approximately 10 cm in front of the participant to simultaneously record the acoustic data. Stimuli were presented on a computer monitor placed approximately 120 cm in front of the participant. The target word embedded in the carrier sentence was shown in the center of the screen. No time pressure was placed on participants, who were instructed to read the sentences at a comfortable speaking rate.

After data collection the articulatory signals were rotated so that X-axis is parallel to the occlusal plane; Y-axis was lateral to the occlusal plane and Z-axis was perpendicular to the occlusal plane. This procedure permitted the X, Y and Z coordinates to correspond to the same dimensions with the same origin across speakers.
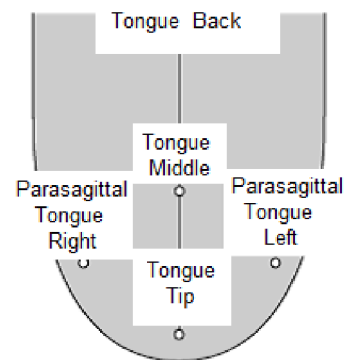

Figure 1. *Schematic placement of tongue sensors.*

2.4. Measurement

We first verified the time point of the maximum TT constriction by visual inspection of the movement trajectories. We observed that TT maximum constriction occurred at the /b/ segment for many of the CVl.bVC (syllable-coda) words. For CVb.lVC (syllable-onset) words, it mostly occurred at the /l/ segment. Based on this observation, we set the onset boundary at the onset of voicing following the consonant release, and the offset boundary at the /b/ offset for syllable-coda tokens. For syllable-onset tokens, that point was instead used as the onset boundary and the offset boundary was defined as the time point where F1 and F2 began diverging in the broadband spectrogram, as viewed in PRAAT. After we collected all the acoustic-based time information, we used this information to locate the articulatory data in Matlab.

2.5. Analyses

2.5.1. Mid-sagittal analysis

For the mid-sagittal analysis, we measured the delay between TT and TM (1), TT and TB (2):

$$\text{TM Delay} = \text{TT Extremum} - \text{TM Extremum} \qquad (1)$$

$$\text{TB Delay} = \text{TT Extremum} - \text{TB Extremum} \qquad (2)$$

The extremum is the point when the articulator is maximally constricted (corresponding to the minimum velocity signal during /l/ production). The TT and TM extrema were derived from the velocity of the z-dimension (i.e., corresponding to tongue raising), while the TB extremum was derived from the velocity of the x-dimension (i.e., corresponding to tongue body retraction). A positive value indicates that the TM/TB gestural target occurs earlier than the TT gestural target. A negative value indicates that the TT gestural target occurs earlier than that for TM/TB.

2.5.2. Para-sagittal analysis

For the para-sagittal analysis, we measured the time course of lateralization during /l/ production and the relative timing relation between the mid-sagittal sensor and the para-sagittal sensor directly.

We computed a lateralization index to capture the degree to which the sides of the tongue are higher/lower than the mid-sagittal plane. In order to do this, we estimated a tongue blade location (between TT and TM) in the mid-sagittal plane by first fitting a second-order polynomial to the x (horizontal) and z (vertical) dimensions of the mid-sagittal sensors. Then we located the intersection point between this polynomial and the horizontal dimension of the para-sagittal sensors (PTR and PTL). The x-, y-, z-values of this point represents the intersection of the mid-sagittal and coronal planes at the location of the para-sagittal sensors along the horizontal dimension. Then we then calculated the difference of the vertical position of this estimated blade sensor and the para-sagittal sensors (3):

$$\Delta \text{ Height = estimated tongue blade - PTL/PTR} \qquad (3)$$

We have only reported data from the dominant lateralization side (the side with the higher value) for each speaker. Four out of six speakers were right dominant. In order to identify the production goal, we extracted the gestural maxima of TM, TB and Δ Height. For TM and TB, the maximum is the time point associated with the minimum velocity in the lowering gesture (z dimension) and retraction gesture (x dimension), respectively. For Δ Height, the maximum is the time point associated with the greatest value in the lateralization index. Then we use this formula to measure the difference between the timing of the mid- and para-sagittal targets (4):

$$\text{Relative timing difference = TM/TB - }\Delta \text{ Height} \qquad (4)$$

A negative value indicates that the mid-sagittal gestural target (i.e., TM lowering/TB retraction) occurs first, while a positive value indicates that the para-sagittal gestural target occurs first.

2.5.3. Tongue curvature and tongue lateralization

In order to measure tongue curvature in the coronal plane, we logged the quadratic term of a second-order polynomial fit to the parasagittal sensors and the estimated tongue blade sensor. Then we applied smoothing spline ANOVA to the temporal trajectories of the quadratic term (used as an indication of tongue curvature in the coronal plane) and the Δ Height (used as an indication of tongue lateralization). Smoothing splines is a technique used to connect discrete data points and find the best fit of the data when they are noisy [6], and has been extensively employed in the speech production research [6,7]. In this analysis, we used the general smoothing splines (gss) package in R [8].

3. Results

3.1. Mid-sagittal analysis

Figure 2 shows the results of TT-to-TM delay. The delay is greater for the vowel /æ/ than the vowel /ɪ/. Across both vowels, the delay is longer for syllable-coda /l/ than for syllable-onset /l/. This indicates that TM gestural movement precedes TT gestural movement in syllable-coda position. In syllable-onset position, TT and TM gestural movements are almost synchronized following /ɪ/. For onset /æ/, there is a slightly long delay, but the delay in onset is still shorter than the delay in coda position.
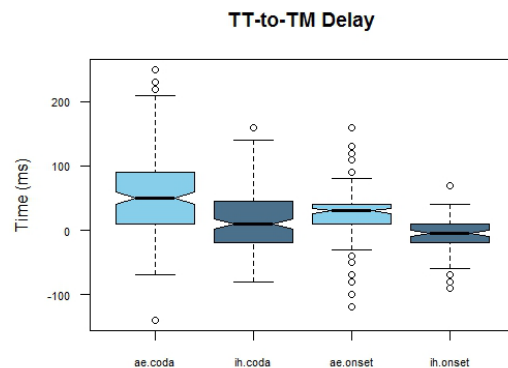

Figure 2. *The TT-to-TM delay of /l/ words in onset and coda position in /æ/ and /ɪ/ context.*

To evaluate the statistical reliability of the trends in Figure 2, we fit linear mixed effects models to the TT-to-TM delay and TT-to-TB delay using the lme4 package in R (version 3.0.2). The fixed effects were vowel (/æ/ and /ɪ/) and syllable position (onset and coda). A random effect was included for speakers. The models are summarized in Table 2. The results show significant syllable effect and vowel effect for the TT-to-TM delay. The results for the TT-to-TB delay were not significant, so we did not report it here. In the model, both syllable and vowel effects were significant. The negative t-value for vowel indicates that the TT-to-TM delay was shorter for /æ/ than for /ɪ/. The negative t-value for syllable indicates that the TT-to-TM delay was shorter for onset than for coda /l/. There is no significant interaction between syllable position and vowel.

Table 2. *Mixed effects model of TT-to-TM delay*

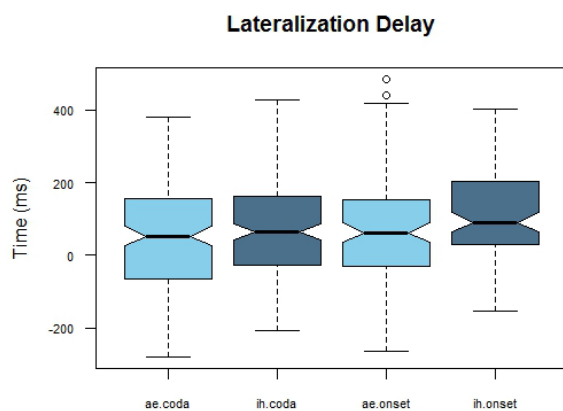| Effects	Estimated	t-value
Intercept	53.809	4.783
Vowel	-35.391	-2.783
Syllable position	-27.560	-4.352

Figure 3. *The lateralization delay of TM sensor to the para-sagittal sensor of /l/ words in onset and coda position in /æ/ and /ɪ/ context.*

3.2. Para-sagittal analysis

Figure 3 shows the time course of lateralization during /l/ production and the relative timing between the mid-sagittal sensor and the para-sagittal sensor. A negative value indicates that the mid-sagittal tongue gestural movement is prior to the para-sagittal tongue gestural movement, while a positive value suggests that the para-sagittal tongue gestural movement is prior to the mid-sagittal tongue gestural movement. Overall, the time course of lateralization remained stable. As for the relative timing relation between the mid-sagittal sensor and the para-sagittal sensor, we can see in this study that the para-sagittal tongue movement occurred prior to the mid-sagittal movement across syllable positions and vowel contexts.

Since the TM and TB results were almost identical, we have only reported the TM result. There was no significant effect of syllable position or vowel context for the TB results.

3.3. Tongue curvature and tongue lateralization

Figure 4 shows the temporal dynamics of tongue curvature in the coronal plane over the entire v-/l/ interval. Zero indicates a flat tongue shape. A negative quadratic term (sides of the polynomial are down relative to the apex) indicates a relatively doomed shape. A positive term (sides of the polynomial are up relative to the apex) indicates a relatively grooved shape. The tongue is grooved for the vowel in the very beginning (roughly correspondent to the vocalic portion). The tongue curves as it moves towards the consonantal portion. Overall, this measure shows a relatively flat tongue shape for the /l/ portion of the trajectory as the curvature values are close to zero. Figure 5 shows the temporal dynamics of tongue lateralization over the same interval. This figure confirms the tongue shape during the v-/l/ interval in Figure 4. A smaller value suggests a relatively grooved tongue shape (i.e., the middle is lower than the sides) and a greater value suggests a relatively doomed tongue shape (i.e., the middle is higher than the sides). In addition to this, we found that the tongue is asymmetrically lateralized during /l/ production by comparing Figure 4 and 5. In onset /æ/, the tongue curvature value is near zero at around 200 ms of the interval (indicating a flat tongue). However, the tongue lateralization value reaches its peak at the same time. This indicates that the tongue blade in the mid-sagittal plane is higher than one of the para-sagittal sides. The coda /æ/ shares the similar pattern in the lateralization figure, but the tongue is less asymmetrically lateralized as compared to the onset /æ/. In the case of onset and coda /ɪ/, the curvature values are all low (close to zero) during the /l/ portion; but the lateralization values are all raising in the same interval. This suggests that one of the para-sagittal sides is higher than the tongue blade (in other words, the tongue is tilted toward one side in the coronal plane).

4. Discussion and conclusion

Our results from the mid-sagittal analysis show that there is a difference between Australian /l/ in syllable-onset and syllable-coda position. This result is consistent with [1]. The TT gestural movement is prior to TM/TB gestural movement, or nearly synchronized with them in syllable-onset position, while in syllable-coda position, the TM/TB gestural movements are prior to TT gestural movement. In addition to this, we also found that /l/s adjacent to /æ/ vowel have a longer TT-TM/TB delay than the ones adjacent to /ɪ/ vowel.
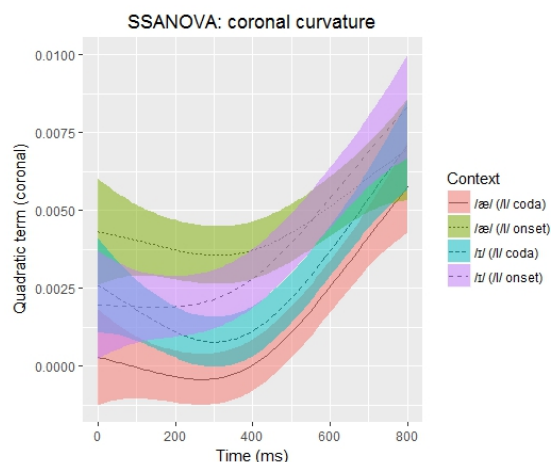

Figure 4. *The temporal dynamics of tongue curvature over the entire v-/l/ interval.*
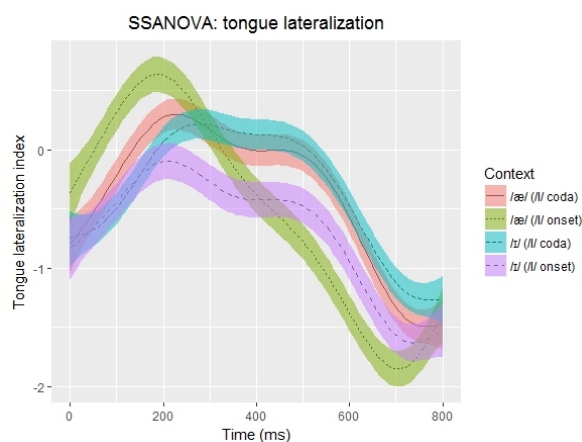

Figure 5. *The temporal dynamics of tongue lateralization over the entire v-/l/ interval.*

For the para-sagittal analysis, we found that (1) the lateralization interval remains stable across syllable position and vowel context; (2) the para-sagittal gestural movement reaches its peak earlier than the mid-sagittal gestural movement. Thus, we can say that the formation of the lateral channel is the speech production goal in our data. Across syllable position and vowel context, the lateralization duration is consistent. In order to achieve the syllable-onset and coda distinction, the mid-sagittal gestures have to coordinate with each other accordingly.

For the tongue curvature and lateralization analyses, we found that a measurement of tongue lateralization is able to capture gestural information that is not captured by tongue curvature in the coronal plane, suggesting that asymmetrical lateralization plays an important role in lateral formation.

To our knowledge this is the first study that has measured the para-sagittal lateralization by observing both mid-sagittal and para-sagittal planes. Future study will focus on /l/ lateralization in languages other than English.

5. References

[1] R. Sproat and O. Fujimura. "Allophonic variation in English /l/ and its implications for phonetic implementation*," Journal of phonetics*, vol. 21, no. 3, pp. 291-311, 1993.

[2] C. P. Browman and L. M. Goldstein, "Articulatory phonology: An overview," Phonetica, vol. 49, no. 3-4, pp. 155-180, 1992.

[3] S. Lin, P. S. Beddor, and A. W. Coetzee, "Gestural reduction, lexical frequency, and sound change: a study of post-vocalic /l/," *Laboratory Phonology*, vol. 5, no. 1, pp. 9-36, 2014.

[4] M. Proctor, "Towards a gestural characterization of liquids: Evidence from Spanish and Russian," *Laboratory Phonology*, vol. 2, no. 2, pp. 451-485, 2011.

[5] M. Stone and A. Lundberg, "Three-dimensional tongue surface shapes of English consonants and vowels," *Journal of the Acoustical Society of America*, vol. 99, no. 6, pp. 3728-3737, 1996.

[6] L. Davidson, "Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance a." *The Journal of the Acoustical Society of America* 120.1 (2006): 407-415.

[7] J. O., Ramsay et al., "Functional data analyses of lip motion." *The Journal of the Acoustical Society of America* 99.6 (1996): 3718-3727.

[8] C. Gu, *Smoothing Spline ANOVA Models*, New York: Springer, 2002.