Complex Tongue Shaping in Lateral Liquid Production Without Constriction-Based Goals

Caitlin Smith[1]

[1]Department of Linguistics, University of Southern California, Los Angeles, California, USA
`smithcm@usc.edu`

Abstract

Liquids are complex consonants that employ multiple gestures of the tongue in order to achieve specific tongue shapes. In the case of the lateral liquid /l/, that specific shape involves lowering and curling of the tongue along the blade. Many models of speech treat /l/ simply as a combination of tongue tip closure and tongue body retraction gestures, with tongue shaping receiving little attention. Here we examine the sequencing of achievement of both the constriction gestures and complex tongue shaping of American English /l/. This will allow us to discern the interaction between constriction formation and tongue shaping, shedding light on the true goals of production for lateral liquids. The results suggest that the so-called primary tongue tip constriction gesture of /l/ is actually a form of tongue bracing that aids in the achievement of tongue curling in tandem with tongue body retraction, the true goals of production of this complex consonant.

Keywords: liquids, laterals, tongue shaping, speech tasks, syllable structure

1. Introduction

Liquids have long brought up questions regarding the representations of their goals of production. This is especially true within the framework of Articulatory Phonology (Browman & Goldstein 1986, 1992), in which the complex tongue shaping associated with liquids is modeled via the coordination of multiple constriction-based gestures. The lateral liquid /l/ is captured by two gestures, one for tongue tip constriction at the alveolar ridge, and one for tongue body constriction at the pharynx. The complex tongue shaping of /l/, characterized in particular by tongue blade lowering or curling in order to create lateral air channels, is considered only indirectly.

The sequencing of the gestures that make up /l/ and lend it its complex shape has been shown to vary by syllable position. In general, the consonantal tongue tip gesture precedes the vocalic tongue body gesture in onset position, but the tongue body gesture precedes the tongue tip gesture in coda position (Sproat & Fujimura 1993; Browman & Goldstein 1995). It remains unclear what the role of tongue shaping is in terms of syllable position, and how that tongue shaping transpires over time when the sequencing of the tongue tip and tongue body gestures of /l/ varies.

It is especially unclear how the sequencing of the constriction gestures and tongue shaping of /l/ interact when one of those gestures is not fully realized. It is common for /l/ in coda position to be realized without any discernible contact between the tongue tip and alveolar ridge, a phenomenon known as /l/ vocalization (Hardcastle & Barry 1989). Previous work has shown that even during the production of vocalized /l/, which is seen in all syllable positions but most prevalent in nucleus and coda positions, substantial curling of the tongue blade is achieved (Smith & Lammert 2013). Understanding how all of these events are structured and sequenced within the syllable may shed light on why /l/ vocalization is prevalent in some syllable positions and not others, as well as further informing its effect on tongue blade curling.

This work aims to discern the goals of production of the /l/ of American English by examining the production of its composite gestures across syllable positions in terms of their individual achievement, their relative coordination, and their relationship to tongue blade lowering/curling. This is done through the use of real-time magnetic resonance imaging (MRI), which provides dynamic information about the entire vocal tract, allowing us to examine both the relative timing of multiple gestures and changes in the shape of the tongue throughout the achievement of those gestures.

2. Methods

Real-time MRI is a methodology that is well suited to the examination of lateral liquid production. Electropalatography can only provide information about whether the tongue tip has made contact with the alveolar ridge and tells us nothing about the complex shaping of the tongue body. Ultrasound is able to image the tongue body, but is usually unable to capture the tongue tip. MRI overcomes this problem, providing a full view of the vocal tract from glottis to lips. However, static MRI requires speakers to maintain production of target phones for several seconds, essentially allowing us to see only the equivalent of a syllabic /l/. The use of real-time MRI allows us to embed /l/ in different contexts, particularly different syllable positions, as well as to examine how the production of /l/ unfolds over time in each of these positions.

2.1. Data

This work examines the speech of two subjects in the USC-TIMIT real-time MRI database, compiled by the University of Southern California's SPAN (Speech Production and Articulation kNowledge) group and publicly available for research purposes (Narayanan et al. 2011). Each speaker in the USC-TIMIT database produced a phonetically balanced set of 460 sentences. Between both subjects examined here, this yielded tokens of /l/ in various syllable positions, as listed in Table 1.

Table 1: *Number of /l/ tokens at each syllable position.*

Syllable position of target /l/	Number of Tokens
Onset	54
Intervocalic/Ambisyllabic	44
Coda	42
Nucleus	34

In order to minimize the influence of tongue shaping from nearby liquids on the liquid targeted for analysis, any instances of /l/ with an /ɹ/ in an adjacent syllable were not considered.

2.2. Image acquisition

A team of SPAN members performed the acquisition of real-time MRI images at Los Angeles County Hospital. Each subject lay supine in the scanner with his or her head held in place by a neck brace. A mirror was placed in the scanner allowing subjects to see out of the bottom of the scanner and read stimuli that were projected onto a screen. Stimuli were presented in five-sentence blocks, with a pause of approximately thirty seconds between each block in order to allow the scanner to cool. Subjects produced each sentence once.

Scans used a real-time MRI protocol specifically developed for the examination of speech production (Narayanan, Nayak, Lee, Sethy, & Byrd 2004). Images were acquired along the midsagittal plane with a complete view of the vocal tract: glottis, pharynx, and oral and nasal cavities. They were reconstructed at a rate of 23.18 frames per second with a field of view of 68 x 68 pixels, corresponding to 200 x 200 millimeters.

2.3. Image analysis

Edge tracking for the contours of the tongue and palate was performed using a custom-designed GUI (graphical user interface). The GUI introduces an image scaling factor of five, bringing the spatial resolution of each frame from 68 x 68 pixels to 340 x 340 pixels. A semi-polar grid of approximately thirty lines orthogonal to the vocal tract was laid over the video images. This grid was fit specifically to each speaker's vocal tract by specifying points at the speaker's glottis, palate, alveolar ridge, and lips.

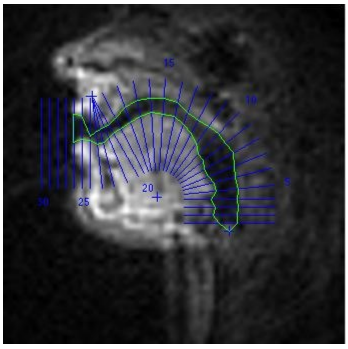

Figure 1: *One frame of midsagittal MRI with grid overlaid and upper and lower surfaces of vocal tract outlined*

Portions of the upper vocal tract surface's contour were fixed at the palate and pharynx; to account for head movement, the grid was shifted as necessary to line up these fixed portions of the contour with the relevant anatomical landmarks. For each frame of video, the air-tissue boundary for the upper and lower surfaces of the vocal tract were automatically detected at each gridline and corrected manually if necessary. Further details on this method of image analysis can be found in Proctor, Bone, Katsamanis, & Narayanan (2010).

2.4. Measurements

This study utilizes two types of measures: those that capture the formation of constriction or tongue shaping that may be seen as goals of production, as well as temporal measures of the achievement of those articulations. First, in order to see whether some syllable positions are more conducive to /l/ vocalization than others, rates of vocalization were compared across syllable positions. Rate of vocalization is calculated as the number of vocalized tokens of /l/ in a given syllable position divided by the total number of tokens of /l/ in that syllable position. A token of /l/ is considered vocalized if there is no contact between the tongue tip and the upper surface of the vocal tract during its production. Only non-coronal adjacent tokens of /l/ were considered for the calculation of rate of vocalization.

Additionally, a quantification of curvature of the tongue contour was performed. First, fifteen evenly spaced points were laid along the tongue contour. The radius of the circle passing through each set of three contiguous points was calculated, and its inverse taken and multiplied by 100 to yield a curvature score. If the circle between three points lay outside of the tongue area, the curvature score was recorded as negative. Curling of the tongue during the production of /l/ was signaled by a high-magnitude negative curvature score at a point along the tongue blade.

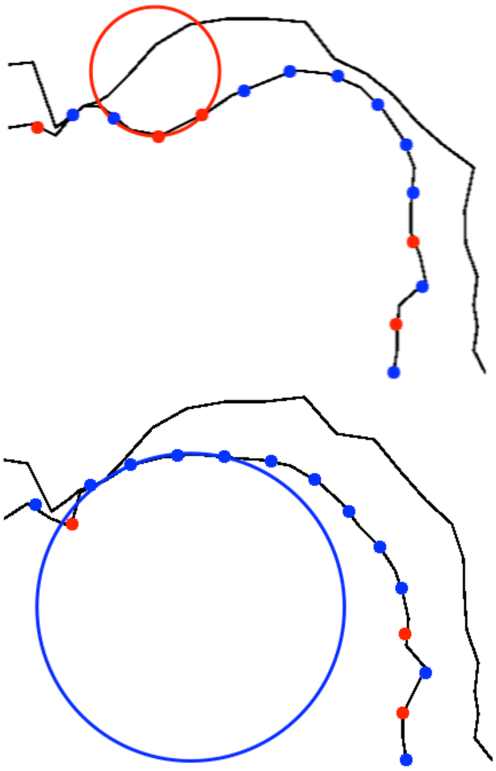

Figure 2: *Evenly spaced points along tongue contour during production of /l/ in 'excluded' (above) and /d/ in 'adult' (below). Blue indicates positive curvature; red indicates negative.*

In order to examine the sequencing of events during the production of /l/, several temporal measures were also made. In cases in which /l/ was not vocalized and tongue tip contact with the upper surface of the vocal tract was achieved, the frame of tongue tip contact was recorded. For all tokens, frame of greatest tongue body retraction and frame of greatest degree of tongue curling were recorded.

3. Results

3.1. /l/ vocalization

Vocalized tokens of /l/ were found in all syllable positions, but rate of /l/ vocalization was affected by syllable position. In onset and intervocalic positions, vocalized tokens were quite outnumbered by those in which some kind of tongue tip contact was made. However, in coda and nucleus positions the great majority of tokens were vocalized. These results are reported in Figure 3.

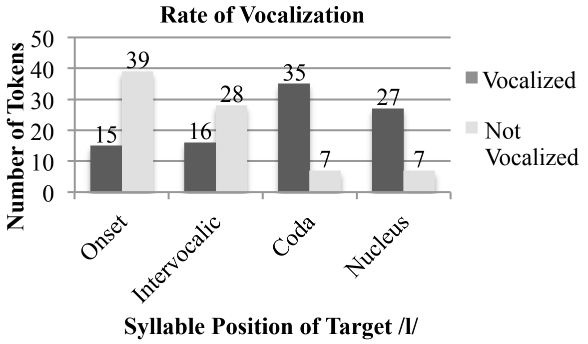

Figure 3: *Rate of /l/ vocalization across syllable positions*

3.2. Tongue curling & tongue tip contact

Lag times between the consonantal and vocalic tasks of /l/ were recorded to the nearest frame. Because many tokens of /l/ were produced with no discernable tongue tip constriction, achievement of the consonantal task was also measured by the achievement of maximal degree of tongue curling. The lag between the achievement of the vocalic gesture, measured by maximum tongue body retraction, and the achievement of the consonantal gesture, measured by both tongue tip closure and maximum degree of tongue curling, is reported in Figure 4.

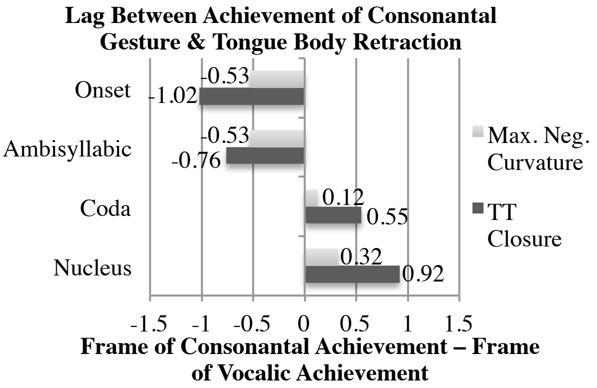

Figure 4: *Lag between achievement of consonantal gesture (either tongue tip closure or maximum degree of curling/negative curvature) and vocalic gesture (tongue body retraction)*

In onset and intervocalic positions, achievement of the consonantal gesture, whether measured by tongue tip contact or maximal tongue curling, preceded retraction of the tongue body. However, the reverse sequence was seen when /l/ is in coda and nucleus positions. Here, maximum tongue body retraction preceded the achievement of the consonantal gesture. In all syllable positions, the lag between achievement of the consonantal and vocalic gestures was substantially less when measured from the frame of maximal degree of tongue curling rather than the frame of tongue tip closure. This indicates that whatever the sequencing of tongue tip closure and tongue body retraction, achievement of maximal tongue blade curling occurred between these two events.

4. Discussion

There is a fairly clear-cut distinction between American English /l/ in onset and intervocalic positions versus coda and nucleus positions in terms of both rate of vocalization and temporal organization. The fact that tongue tip closure precedes tongue body retraction syllable-initially but follows this retraction syllable-finally is consistent with previous studies such as Sproat & Fujimura (1993) and Browman & Goldstein (1995). This is generally seen as a tendency for the vocalic gesture to occur closer to the syllable nucleus than the consonantal gesture. This study has also determined that maximal tongue curling occurs at some point between the achievements of these two constriction gestures, however they are arranged. Thus, the sequence of events in the production of /l/ appears to be tongue tip closure → maximal curvature → tongue body retraction in onset and intervocalic positions, but tongue body retraction → maximal curvature → tongue tip closure in coda and nucleus positions.

However, one should not think of these two possible sequences as simple mirror images of one another, or think of the tongue curling event as somehow intermediate in its vocalic/consonantal nature between tongue tip closure and tongue body retraction. One should instead consider the fact that in the case of onset and intervocalic /l/, tongue tip contact is followed by greater curling of the tongue, while in nucleus and coda /l/ the degree of tongue curling is diminished once tongue tip contact is made. Put another way, it appears that in onset and intervocalic positions the tongue tip contact has a facilitative effect on tongue curling. It could be, then, that the main production goal for /l/ is not the tongue tip closure, but the curling along the tongue blade, for which tongue tip contact is merely an assistive act of bracing. Figures 5 and 6 illustrate the sequencing of tongue tip closure and maximal tongue curvature in onset and coda positions.

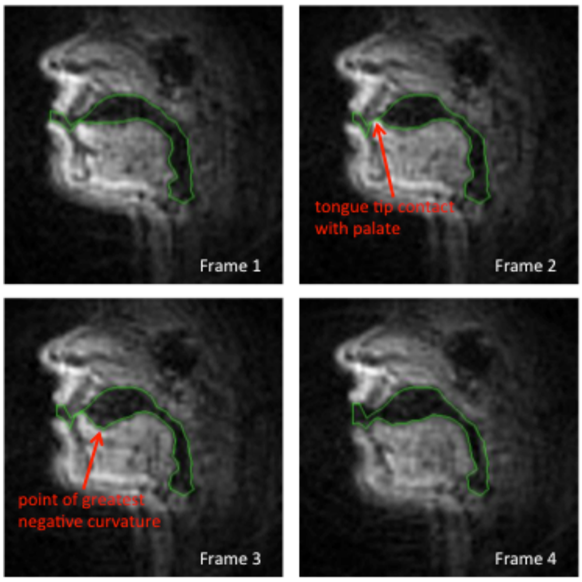

Figure 5: *Time course of production of /l/ in the word 'black,' demonstrating that tongue tip closure precedes maximal tongue curvature in onset position*

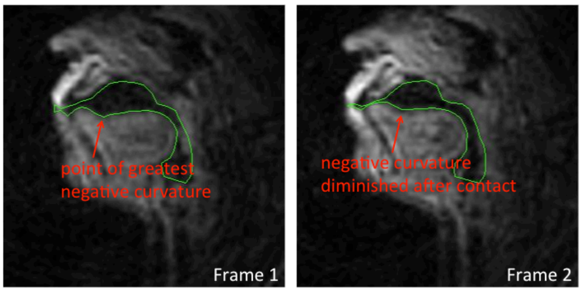

Figure 6: *Time course of production of /l/ in the word 'help,' demonstrating that maximal tongue curvature precedes tongue tip closure in coda position*

The idea of the tongue tip as a brace during the production of /l/ (and other consonants) in order to facilitate precise tongue shaping has been proposed previously by Stone (1990) and Stone, Faber, Raphael, & Shawker (1992). These studies suggest that tongue tip contact is not just facilitative but necessary to achieve shaping of the tongue beyond the repertoire of basic shapes available to unbraced vocalic postures. However, these studies deal only with tokens of /l/ that involved tongue tip contact. In fact, tongue curling occurs even when /l/ is vocalized, indicating that bracing is not strictly necessary for the production of /l/. Rather, there is a tendency to prefer tongue tip contact to assist in the achievement of tongue curling.

However, this tendency only seems to hold in onset and intervocalic positions, for which the rate of /l/ vocalization is low. The fact that /l/ vocalization varies according to syllable position indicates that bracing plays a different role depending on whether /l/ occurs in onset/intervocalic position or nucleus/coda position. This may be explained by the fact that the relative timing of tongue body retraction is different between these two sets of syllable positions. In onset and intervocalic positions, the tongue body is not yet retracted when tongue tip contact and maximal tongue curling are achieved. The more anterior position of the tongue body may make curling more difficult and in need of assistance in the form of a bracing maneuver of the tongue tip against the alveolar ridge. On the other hand, in coda and nucleus positions, for which the rate of /l/ vocalization is quite high, the degree of tongue curling is actually diminished once tongue tip contact is made. The tongue body is retracted earlier for /l/ in these syllable positions, bringing the tongue into a position that makes tongue curling easier. Thus, tongue tip contact is at best unnecessary for the achievement of tongue curling, and may even be inhibitory. This could explain, at least partially, why there is greater tendency toward vocalization of /l/ in coda and nucleus positions, consistent with the findings of Hardcastle & Barry (1989).

This theory is in stark contrast with previous accounts of the goals of production for liquids within Articulatory Phonology. In this framework, lateralization is captured indirectly through tongue tip and tongue body gestures that stretch the tongue, narrowing it enough to open lateral air channels, rather than the tongue blade curling that is evident in the data presented here. According to this view, the consonantal tongue tip constriction is the primary gesture of the liquid rather than a form of bracing that aids in the achievement of the true goals of production, tongue curling and tongue body retraction. In fact, a shaping-based goal of production is not something that can be currently implemented within the theory, which relies solely on constriction goals. However, this study provides evidence that in the case of lateral liquids this stance should be reevaluated.

5. Conclusion

The relative timing of the constriction gestures of American English /l/ and their relationship to the achievement of tongue blade curling indicate that /l/ has a production goal that is based on tongue shaping. Tongue tip contact, rather than being the primary gesture or goal of this consonant, appears to be a form of tongue bracing to achieve complex shaping of the tongue blade and body. Whether this bracing is facilitative to tongue curling depends upon whether the tongue body is retracted, and thus depends upon the syllable position of /l/. At the very least, these findings bring up interesting questions about models of speech that are based entirely on constriction tasks and largely set aside issues of tongue shaping.

6. Acknowledgements

This work was funded through NIH Grant DC007124 to the University of Southern California.

7. References

Browman, C. P., & Goldstein, L. (1986). Towards an Articulatory Phonology. *Phonology Yearbook*, *3*, 219–252.

Browman, C. P., & Goldstein, L. (1992). Articulatory Phonology: An Overview. *Phonetica*, *49*, 155–180.

Browman, C. P., & Goldstein, L. (1995). Gestural Syllable Position Effects in American English. In F. Bell-Berti & L. J. Raphael (Eds.), *Producing Speech: Contemporary Issues*. Woodbury, NY: AIP Press.

Hardcastle, W., & Barry, W. (1989). Articulatory and perceptual factors in /l/ vocalisations in English. *Journal of the International Phonetic Association*, *15*(2), 3–17.

Narayanan, S., Bresch, E., Ghosh, P., Goldstein, L., Katsamanis, A., Kim, Y., … Zhu, Y. (2011). A Multimodal Real-Time MRI Articulatory Corpus for Speech Research. In *Proceedings of the 11th Annual Conference of the International Speech Communication Association (INTERSPEECH)* (pp. 837–840). Florence, Italy.

Narayanan, S., Nayak, K., Lee, S., Sethy, A., & Byrd, D. (2004). An approach to real-time magnetic resonance imaging for speech production. *Journal of the Acoustical Society of America*, *115*(4), 1771–1776.

Proctor, M. I., Bone, D., Katsamanis, N., & Narayanan, S. (2010). Rapid Semi-automatic Segmentation of Real-time Magnetic Resonance Images for Parametric Vocal Tract Analysis. In *Proceedings of the 10th Annual Conference of the International Speech Communication Association (INTERSPEECH)* (pp. 1576–1579). Makuhari, Japan.

Smith, C., & Lammert, A. (2013). Identifying consonantal tasks via measures of tongue shaping: a real-time MRI investigation of the production of vocalized syllabic /l/ in American English. In *Proceedings of the 14th Annual Conference of the International Speech Communication Association (INTERSPEECH 2013)* (pp. 3230–3233). Lyon, France.

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, *21*, 291–311.

Stone, M. (1990). A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *Journal of the Acoustical Society of America*, *87*(5), 2207–2217.

Stone, M., Faber, A., Raphael, L. J., & Shawker, T. H. (1992). Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ], and [l]. *Journal of Phonetics*, *20*, 253–270.