

# Exploring temporal characteristics of lingual gestures in American English /l/

Mitsuhiro Nakamura

Nihon University, Tokyo, Japan  
nakamura.mitsuhiro99@nihon-u.ac.jp

## ABSTRACT

This study investigates interspeaker variation and variables in the production of the American English /l/. Using data extracted from the real-time MRI database, the current study aims to analyze the apical/laminal variation in the temporal domain and uncover articulatory tasks for generating the vocal tract shapes related to the apical/laminal distinction. The preliminary results reveal that (i) in the apical /l/ production, a ‘dip’ is formed on the anterior and middle part of the tongue before the tongue tip is raised to make a complete closure at the alveolar ridge and (ii) the apical /l/ and the laminal /l/ demonstrate the same pattern of intergestural coordination between the tip/blade and the dorsum (tip delay). It is reasonable to suppose that the dip formation facilitates the orientation angle of the tip to the alveolar ridge. The implications for articulatory organization and representation are discussed.

**Keywords:** Real-time MRI, USC-TIMIT, English lateral, apical vs. laminal

## 1. INTRODUCTION

Studying speech production from an articulatory phonetics perspective can be seen as the practical implementation of a ‘parametric approach’ [1], or parametric phonetics. This approach contrasts with, but is complementary to, a postural or segmental approach. Parametric phonetics focuses on the time-varying parameters of articulation. The current study applies this view to the study of the apical/laminal distinction of /l/ in American English.

The apical/laminal distinction refers to the point of the tongue involved in the constriction formation. The tongue tip is used for apical articulations while the tongue blade is used for laminal articulations. According to Dart [7], the highest percentage of the English /l/ is ‘apical,’ and the next most common pattern is ‘apicolaminal.’

Furthermore, the difference between these two articulations goes beyond the area of the tongue which makes a contact with the hard palate. Using the static MRI technique, Narayanan *et al.* [10] have shown that the overall tongue shape of the apical /l/ is concave with a groove in the anterior and/or middle part of the tongue, while that of the laminal /l/ reveals largely an overall convex shape. Stone & Lundberg

[14] identify ‘two-point displacement’ for /l/ which involves ‘two regions of displacement, an elevated anterior and posterior segment, with a short groove, almost a dimple, indicating compression of the middle segment.’ It can be asked how these characteristic tongue shapes are formed in the temporal domain. This question is related to the well-known intergestural timing pattern in /l/ production.

An influential study on the clear/dark allophony [13] has shown that in the production of the English /l/, maximum dorsal movement precedes maximum tip/blade movement for the dark variant, while the opposite temporal relationship occurs for the clear variant. This intergestural coordination is identified as ‘tip delay.’ The current study replicates the articulatory analysis to compare the intergestural timing pattern between the apical /l/ production and the laminal /l/ production.

The current study explores the question of how apical articulations are related to the concave shape of the tongue. We aim to (i) provide initial empirical data for the apical /l/ gesture, in comparison with the laminal /l/ gesture, (ii) capture the relationship between the articulatory tasks and the overall tongue shape, and (iii) compare the intergestural timing pattern of the apical /l/ with that of the laminal /l/. We will conclude with future directions.

## 2. DATA COLLECTION AND ANALYSIS

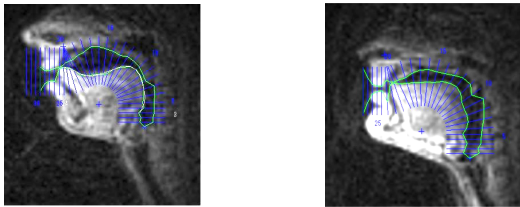
The speech materials were drawn from a database of multimodal speech production data, USC-TIMIT [11, 16]. This database comprises the real-time magnetic resonance imaging (rtMRI) data and acoustic data of 460 phonetically balanced sentences read by 10 native speakers of American English.

The lateral articulations in this study are categorized broadly into two groups, apical and laminal, based on a qualitative observation of the overall tongue shape. This is not to say that the tongue shape is a better indicator of apicality/laminality than the tongue-palate contact pattern. Rather, it works as one potential indicator. As mentioned earlier, the apical /l/ at the maximum constriction involves a groove on the anterior/middle part of the tongue while the laminal /l/ shows a largely flattened surface ([10]). Also, Dart [6] identifies 4 different contact patterns based on the linguograms and demonstrates with X-ray tracings that the overall tongue shape of the

‘apical’ and ‘upper apical’ categories is generally concave and that of the ‘apicolaminal’ and ‘laminal’ categories is largely convex (or slightly arched).

A custom-designed GUI [11, 12] was used for tracking tongue and palate edge location. An image frame was captured at the maximum constriction of /l/ closure (the maximum elongation of the anterior tongue) and its vocal tract outline was obtained by performing semi-automatic segmentation. Figure 1 shows vocal tract outlines with semi-polar analysis grids overlaid. When the program failed to locate the edges with sufficient accuracy, the vocal tract outline was manually corrected. Then, the overall tongue shape was visually inspected for categorization.

**Figure 1:** Segmented rtMRI frames, 2 different speakers  
(a) Apical /l/ (b) Laminal /l/



Also, in order to analyze how the characteristic shapes are created, the sequence of segmented tongue postures in the midsagittal plane was captured from the last frame of the preceding segment to the first frame of /l/ closure. The frame rate is 23.18 frames/s.

The temporal coordination between the tongue tip/blade and the dorsum was analyzed for the selected utterances in the corpus. Dataset 1 consisted of words with the onset singleton /l/ which were preceded by a schwa (e.g. *the latest, to leap*), and dataset 2 consisted of words with the coda singleton /l/ which were followed by a schwa or /h/ (e.g. *skill of, all holidays*). Other aspects of phonological contexts were not controlled. The number of tokens was 43 in total (20 onset /l/s + 13 coda /l/s).

The intergestural coordination was analyzed using the MATLAB-based programs. The VocalTract [3], which utilized a region of interest (ROI) technique, was used to detect the tongue movement over time. Pseudo-circular regions were placed over the location of the tongue tip (TT) constriction, tongue dorsum (TD) constriction, and middle-tongue groove formation. The average pixel intensity inside the region was calculated for each frame. Then, using the MVIEWRT’s [15] ‘find gesture’ routine, the tip gesture and the dorsum gesture were identified. The onset/offset of a whole gesture and a gestural plateau/nucleus was specified at 20% of the tangential velocity peak associated with movement toward, or away from, a target constriction. The times of the maximum constriction (i.e. velocity minimum), or MAXC in the interval of a gestural plateau, were used for calculating the index ‘tip delay.’ Following Sproat & Fujimura [13], tip delay in this study was defined as follows: [Tip Delay]=time(TT velocity minimum)-time(TD velocity minimum). The analysis was conducted for two speakers, F1 (female, apical /l/ speaker) and M2 (male, laminal /l/ speaker).

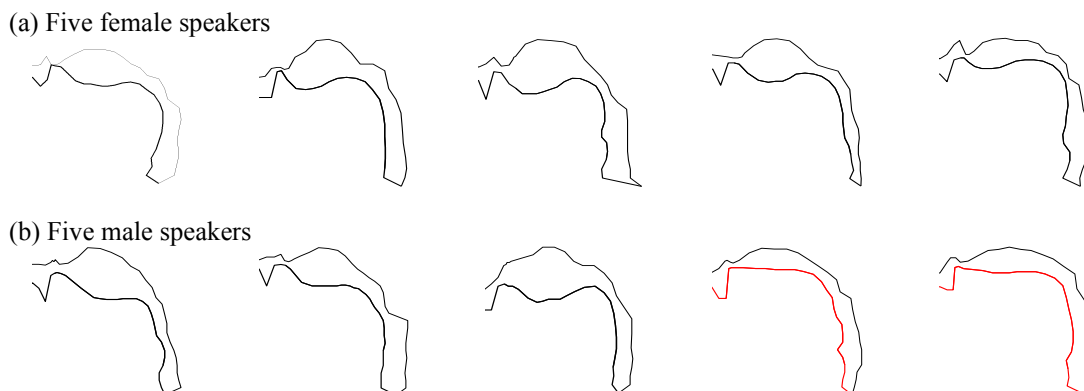
### 3. RESULTS

#### 3.1. Qualitative observation of apical and laminal /l/

The first task is to identify the apical/laminal differences among 10 speakers in the database. Figure 2 demonstrates the midsagittal vocal tract profiles captured in the production of the onset singleton /l/ in the sentence ‘The cow wandered from the farmland and became *lost*.’ Figure 3 shows the vocal tract profiles captured in the production of the coda singleton /l/ in the sentence ‘The big dog loved to chew on the old rag *doll*.’

It can be seen from Figures 2 and 3 that all the tongue profiles colored in black involve a raised tongue tip and dorsum with a middle-tongue groove, the degree of which varies across these 8 speakers.

**Figure 2:** The midsagittal profiles of the vocal tract of 10 speakers producing the onset singleton /l/



**Figure 3:** The midsagittal profiles of the vocal tract of 10 speakers producing the coda singleton /l/

(a) Five female speakers



(b) Five male speakers



In contrast, the two tongue profiles colored in red are largely flattened. Thus, the 8 speakers are categorized as apical and the two speakers as laminal.

Now we will move on to the analysis of how the characteristic shapes are created. Lateral articulation is shown in Figure 4, which compares apical and laminal production of the onset /l/ in the sentence ‘The cow wandered from the farmland and became lost.’ The apical /l/ speaker in (a) reveals a certain amount of ‘dipping’ in the anterior/middle part of the tongue (indicated by the red line). This dip formation consistently occurs before the raising gesture of the

tongue tip. After making a dip, the tongue dorsum is retracted, at the same time, the anterior tongue extends to make a closure at the alveolar ridge. These create a groove on the tongue surface and the overall tongue shape becomes concave. In contrast, the laminal /l/ speaker in (b) simply raises the tongue tip/blade for the /l/ closure from the immediately preceding position (indicated by the gray dotted line) and the dorsum is retracted.

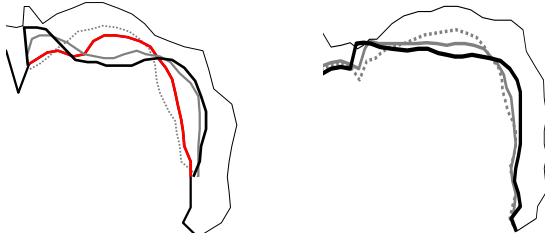
Figure 5 illustrates the production of the coda /l/ in the sentence ‘A roll of wire lay near the wall.’ Similar to the onset /l/, the dip (the red line) is formed from the immediately preceding position (the gray dotted line) and precedes the raising gesture of the anterior tongue in the production (a). It seems that the posterior tongue remains relatively stable: this could be the coarticulatory effect of the preceding back vowel and needs further research.

We now focus on the case where the apical /l/ in the coda position is vocalized. Figure 6 illustrates the vocalized /l/ within and across a word boundary. The relevant rtMRI images were identified with reference to spectrograms of the utterance recorded. The case in (a) is extracted from the sentence ‘When all else fails use force’ and the case in (b) is from the sentence ‘Should giraffes be kept in small zoos?’ In both cases, the dip is formed and in these particular vocalized

**Figure 4:** Production of onset /l/ in ‘lost’

(a) Apical /l/

(b) Laminal /l/

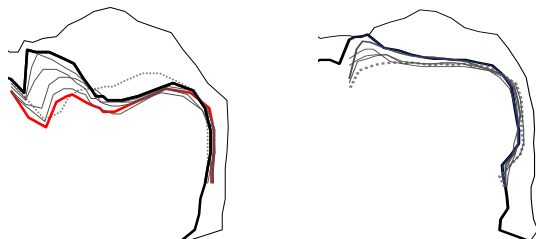


Gray dotted = the last frame of the preceding segment  
Red = making a ‘dip’ on the tongue  
Black = the first frame of the /l/ closure

**Figure 5:** Production of Coda /l/ in ‘wall’

(a) Apical /l/

(b) Laminal /l/

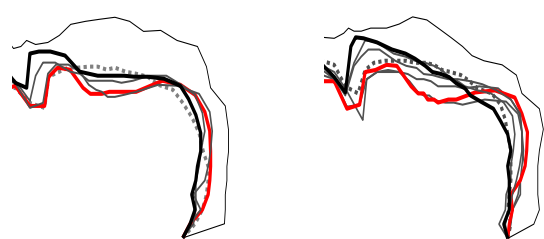


Gray dotted = the last frame of the preceding segment  
Red = making a ‘dip’ on the tongue  
Black = the first frame of the /l/ closure

**Figure 6:** Vocalized /l/ production (apical /l/)

(a) ‘fails’

(b) ‘small zoos’



Gray dotted = the last frame of the preceding segment  
Red = making a ‘dip’ on the tongue  
Black = the /z/ constriction

cases, the tongue tip is raised for the following voiced alveolar fricative /z/.

### 3.2. Analysis of intergestural coordination: tip delay

Figure 7 shows the mean values of tip delay for the onset and coda /l/ averaged for the two speakers, F1 (apical /l/ speaker) and M2 (laminal /l/ speaker). The negative value of the onset /l/ means that the maximum constriction of the tongue dorsum is reached *after* that of the tongue tip/blade. Conversely, the positive value of the coda /l/ means that the maximum constriction of the tongue dorsum is reached *before* that of the tongue tip/blade. Regardless of the two types of /l/, the timing pattern of intergestural coordination is consistent.

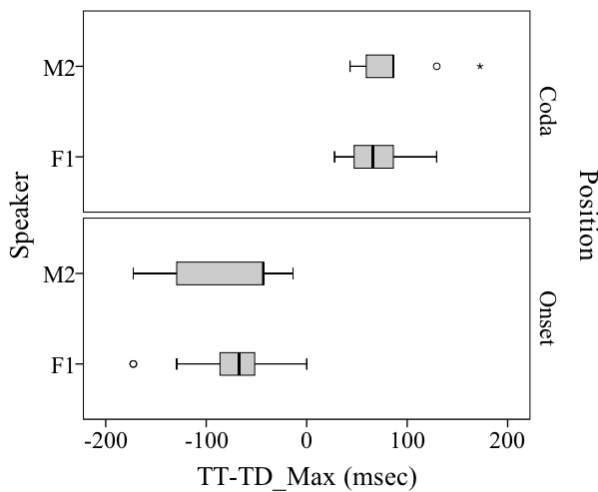


Figure 7: Mean tip delay for the onset and coda /l/

## 4. DISCUSSION

The preliminary results reveal that three major phases are identified in the apical articulation: (i) a ‘dip’ formation on the anterior and middle part of the tongue; (ii) retraction and raising of the tongue dorsum; (iii) elongation of the anterior tongue to make a complete closure at the alveolar ridge.

It is reasonable to suppose that the dip formation found in the apical /l/ production facilitates the orientation angle of the tip to the alveolar ridge. This then effectively leads to the achievement of the overall concave shape of the tongue. Furthermore, the ‘dip’ appears even when the /l/ is vocalized. This suggests that the dip formation could be considered as a preparatory action for raising the tongue tip toward the alveolar ridge and for initiating the execution of the apical /l/ gesture.

The results of comparing the temporal index ‘tip delay’ are compatible with previous studies ([5], [13]). The intergestural coordination shows a uniform trend between the two types of /l/, namely apical and laminal. While the formation process of the complete

closure in apical /l/s differs from that of the laminal /l/s, the two /l/s maintain the timing pattern of intergestural coordination.

The current study focuses on the overall tongue shape for the apical/laminal distinction of the English /l/. This methodology merits further improvements. It is worth noting that there are articulatory studies on the production of ‘obstruents’ which are relevant to the current study. In an EMA study of Wubuy (Australian) coronals, Best *et al.* [2] have shown that in the apicals [t, t̪], the tongue body (TB) gesture is relatively stable to support the lever-like motion of the tongue tip (TT), whereas the TB and TT gestures move in tandem in the laminals [t̪, c]. These patterns are also observed for some English obstruents by Derrick *et al.* [8] who compared the apicals /d, dʒ/ with the laminal /ð/. These apical and laminal consonants are simplex segments, articulatory coordination of which can be characterized as the given consonantal gesture being superimposed on the vocalic gesture. In contrast, the /l/ is considered as a complex segment which involves the consonantal gesture (i.e. the tip/blade raising) and the vocalic gesture (i.e. the dorsum retraction). Taken together, the articulatory nature of the given segment and the phonological status of the apical/laminal distinction should be considered carefully. It is supposed that identification of both ‘spatial’ and ‘temporal’ parameters can lead to a better understanding of the apical/laminal distinction of the English /l/.

The explorations of the apical/laminal distinction raise questions about how a speaker chooses one or the other. Recent studies on the biological grounding of phonology (Fuchs *et al.* [9]) have suggested that ‘biology determines the limits, the frame of reference for phonology.’ This leads us to ask a further question of how individual differences impact on phonology. Answers to these questions await more data and analysis.

## 4. CONCLUSION

The current study has explored the apical/laminal distinction of /l/ in American English using the data extracted from the real-time MRI database. It is hoped that the parametric phonetic analysis presented in this study provides another interesting description of the lateral articulation and contributes to a better understanding of individual differences and articulatory organization.

ACKNOWLEDGEMENTS: I would like to thank three anonymous ICPhS reviewers for their critical and constructive comments. All remaining errors are of course mine. This research is supported in part by a grant from Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C) 2018, Grant no. 18K00548.

## 7. REFERENCES

- [1] Abercrombie, D. 1965. Parameters and phonemes. In: Abercrombie, D. *Studies in Phonetics and Linguistics*. Oxford. Oxford University Press, 120-130.
- [2] Best, C.T., Kroos, C.H., Bundgaard-Nielsen, R.L., Baker, B., Harvey, M., Tiede, M., Goldstein, L. 2014. Articulatory basis of the apical/laminal distinction: Tongue tip/body coordination in the Wubuy 4-way coronal stop contrast. *Proc. of the 10th ISSP, Cologne, Germany*, 33-36.
- [3] Blaylock, R. 2014. VocalTract: software for describing properties and behaviors of a vocal tract obtained with rtMRI. Los Angeles, CA: University of Southern California.
- [4] Browman, C.P., Goldstein, L. 1989. Articulatory gestures as phonological units. *Phonology* 6-201-51.
- [5] Browman, C.P., Goldstein, L. 1995. Gestural syllable position effects in American English. In: Bell-Berti, F. & Raphael, L.J. (eds) *Producing Speech: contemporary issues*. New York: AIP Press, 19-33.
- [6] Dart, S.N. 1991. Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics*, No. 79, 1-155.
- [7] Dart, S.N. 1998. Comparing French and English coronal consonant articulation. *J. Phonetics* 26, 71-94.
- [8] Derrick, D., Fiasson, R., Best, C.T. 2014. Coordination of tongue tip and body in place differences among English coronal obstruents. *Proc. of the 10th ISSP, Cologne, Germany, 5-8 May*, 94-97.
- [9] Fuchs, S., Pompono-Marschall, B., Perrier, P. 2007. Is there a biological grounding of phonology? Determining factors, optimization, and communicative usage. *Proc. of the XVI ICPhS, Saarbrücken, Germany*, 219-223.
- [10] Narayanan, S., Alwan, A.A., Hanker, K. 1997. Towards articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals. *J. Acoust. Soc. Am.* 101(2), 1064-1077.
- [11] Narayanan, S., Toutios, A., Ramanarayana, V., Lammert, A., Kim, J., Lee, S., Nayak, K., Y-C., Zhu, Goldstein, L., Byrd, D., Bresh, E., Ghosh, P., Katsamanis, A., Proctor, M. 2014. Real-time magnetic resonance imaging and electromagnetic articulography database for speech production research. *J. Acoust. Soc. Am.* 136(3), 1307-1311.
- [12] 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. *Proc. of Interspeech 2010, Chiba, Japan*, 1576-1579.
- [13] Sproat, R., Fujimura, O. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *J. Phonetics* 21, 291-311.
- [14] Stone, M., Lundberg, A. 1996. Three-dimensional tongue surface shapes of English consonants and vowels. *J. Acoust. Soc. Am.* 99(6), 3728-3737.
- [15] Tiede, M. 2008 MVIEWRT: multiple concurrently recorded data streams viewer (a version intended for display of pixel intensity data derived from rtMRI. New Haven, CT: Haskins Laboratories.
- [16] Toutios, A. Narayanan, S.S. 2016. Advances in real-time magnetic resonance imaging of the vocal tract for speech science and technology research. *APSIPA Transactions on Signal and Information Processing*, 5, 1-12.