Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Speech Communication
Session 5aSCb: Production and Perception II: The Speech Segment (Poster Session)

5aSCb39. On distinguishing articulatory configurations and articulatory tasks: Tamil retroflex consonants

Caitlin Smith*, Michael Proctor, Khalil Iskarous, Louis Goldstein and Shrikanth Narayanan

***Corresponding author's address: Linguistics, University of Southern California, 3601 Watt Way, Los Angeles, CA 90089, smithcm@usc.edu**

Speech production can be described in multiple coordinate frames: articulatory configurations, gestural tasks, and acoustic patterns. Examination of the achievement of retroflex stops and liquids in Tamil suggests that we must consider separately the gestural task of apical post-alveolar constriction and the articulatory maneuver to achieve the task. The maneuver of the tongue during retroflex consonants varies across vowel contexts. Specifically, in the symmetrical intervocalic contexts between back vowels /a/ and /u/, an apical post-alveolar constriction is achieved by curling back the tongue. In the context of high front vowel /i/, a laminal post-alveolar constriction is achieved by bunching the tongue. However, the location of retroflex consonant constriction within the vocal tract is consistent across all of these vowel contexts, suggesting that the constriction task remains the same. Variation in the articulatory configuration of the retroflex in the two contexts was quantified through Gaussian curvature functions at fourteen points along the tongue, sampled at evenly spaced points throughout the vocal tract, on every other gridline of a polar-rectangular grid in every frame in each utterance. The empirical results support the notion that the articulatory configuration coordinate frame and the gestural task frame provide separate, but related, descriptions of speech production.

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4800521]
Received 22 Jan 2013; published 2 Jun 2013

1. INTRODUCTION

In the task dynamic model of speech production, articulatory gestures specify the location and degree of a constriction. The actual articulatory configurations necessary to achieve this constriction are not specified, but rather determined primarily by a set of contextual factors that may govern the state of the vocal tract. It is interesting, then, that there is an entire class of consonants, the retroflexes, which are apparently characterized instead by an articulatory maneuver employed in order to achieve contact between the tongue tip and the post-alveolar region of the palate. Does such a classification suggest that retroflexion of the tongue is achieved invariably using that particular articulatory maneuver? If so, what does it mean for the task dynamic model if we can classify these consonants by something other than their constriction task, instead appealing to how an articulator is configured?

Tamil (Dravidian) is an ideal language to examine in trying to answer these questions, as its phoneme inventory includes several retroflex consonants (Keane, 2004). By using real-time MRI to examine the so-called retroflex consonants in varying vowel contexts in which the tongue body is employed in different ways, we can see how and to what extent the articulatory maneuvers of these consonants will be affected. Our aim is to use measures of constriction location and tongue shaping to shed some light on whether we can actually characterize a class of consonants according to the articulatory maneuver of retroflexion rather than a target constriction location.

Previous work on Tamil indicates that we may indeed find variability in the articulatory maneuvers of retroflex consonants. Wiltshire & Goldstein (1997) found that retroflex consonants showed more variability in tongue tip orientation across vowel contexts than their dental/alveolar counterparts. Our use of real-time MRI rather than EMA allows us to further investigate the particulars of tongue shaping that lead to these differences in orientation, as well as constriction location and degree. Narayanan, Byrd, & Kaun (1999) established MRI as a viable method for the study of tongue shaping during dental/alveolar and retroflex consonants in Tamil.

2. METHODS

2.1. Elicited Speech

Target words in this study were all nonce words produced in isolation, and were all of the form pV__Vm, with the target consonant being produced between the two vowels. The vowel contexts in which the target consonant occurred were symmetrical. There were a total of nine target consonants (dental/alveolar stops, nasals, and rhotics, and an alveolar lateral approximant) produced in three vowel contexts (/a/, /i/, and /u/) for a total of twenty-seven unique forms. These were divided into two sets that were produced in alternation for a total of five repetitions for three speakers, and three repetitions for Speaker 4. Here, we will only consider the retroflex oral stops and rhotics.

2.2. Image Acquisition

Images were acquired for four speakers (three male, one female) with a real-time MRI protocol developed specifically for use in examining speech production (Narayanan et al., 2004). Images of speakers' vocal tracts (glottis, pharynx, and oral and nasal cavities) were taken on the midsagittal plane and reconstructed with a frame rate of 33.18 frames per second and a resolution of 68 x 68 pixels.

2.3. Image Analysis

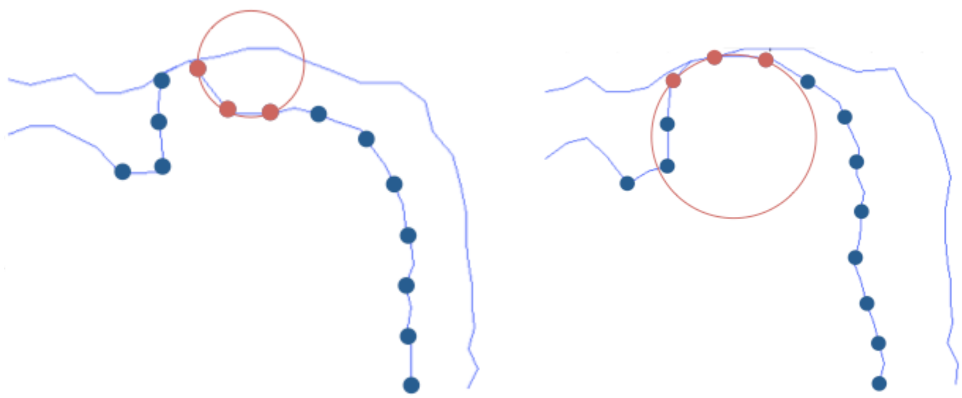
MRI videos were inspected using a custom graphical user interface which overlays a grid of approximately thirty lines orthogonal to the vocal tract (Narayanan et al., 2011; Proctor, Bone, & Narayanan, 2010). Detection of the upper and lower edges of the vocal tract were generated by marking where the vocal tract surfaces crossed each of these gridlines. Video frames corresponding to the beginning of movement toward the consonant target, initial consonant closure, and the end of movement away from the target were selected. Here, we will only discuss frames corresponding to initial achievement of closure for a target consonant.

Constriction location for each consonant was measured to the nearest gridline. In order to quantify tongue shaping during consonant articulation, measurements of Gaussian curvature were calculated. First, the curvature of the lower surface of the vocal tract, measured along the approximately thirty overlaid gridlines, was upsampled, then downsampled to fourteen evenly spaced points. For each set of three adjacent points, the diameter of the circle going through these three points was calculated. Then, the inverse of that circle diameter was taken so that a sharp curvature would have a high Gaussian curvature rating, and a gradual curvature would have a low rating.

3. RESULTS & DISCUSSION

3.1. Shaping of Tongue

Analysis of tongue shape during the articulation of retroflex stops and rhotics revealed quite a bit of difference depending on whether they were produced in the context of high front vowel /i/ or the back vowels /a/ and /u/. As discussed in Section 2.2, a Gaussian curvature analysis was applied to the curvature of the tongue in the three vowel contexts. The results show that in the /i/ vowel context, so-called retroflex consonants have little or no negative Gaussian curvature, which would signal curling back of the tongue tip. On the other hand, these same consonants in /a/ and /u/ contexts show sharp negative curvature at some point in the region of the tongue blade. Figure 1 compares tokens of the articulation of the retroflex stop /ɖ/ in the back vowel context /a/ and front vowel context /i/.



(a) (b)

FIGURE 1. Retroflex stop /ɖ/ in /a/ (a) and /i/ (b) vowel contexts. The three highlighted points along the curve of the tongue in (a) display negative curvature, while the three points along the curve in (b) show positive curvature.

Thus, in the /i/ context it appears that the post-alveolar consonant is achieved through bunching of the tongue, with little or no curling back of the tongue tip. However, in /a/ and /u/ contexts the post-alveolar consonant is achieved through the canonical articulatory maneuver of retroflexion.

3.2. Constriction location

Despite the fact that tongue shaping for the ‘retroflex’ consonants was quite variable across vowel contexts, constriction location in the post-alveolar region was fairly stable across vowel contexts. Recall that constriction location was measured by the superimposed gridline crossed by the tongue at the point at which it made contact with the palate. For all speakers, retroflex/post-alveolar consonants usually occurred at one specific gridline, and never varied farther toward the front or back by more than one gridline, no matter the vowel context. Thus, constriction location was relatively stable across vowel contexts, even though the actual maneuver to achieve contact between tongue tip and palate was variable.

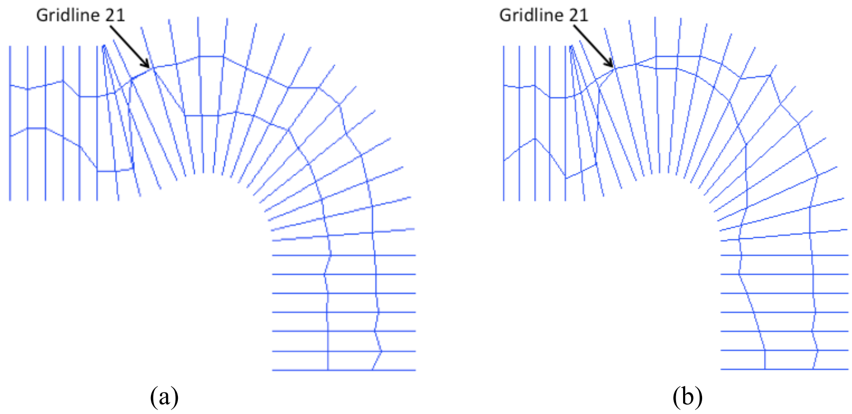


FIGURE 2. Retroflex stop /ɖ/ in /a/ (a) and /i/ (b) vowel contexts. Both show contact between the tongue tip and the palate at Gridline 21.

This result is compatible with the view that consonant gestures are superimposed over vowel gestures, as in (Ohman, 1966). Because the consonant is in a symmetrical vowel context, the tongue body is engaged throughout the articulation of the consonant. In the case of the /i/ context, the speaker is unable to compromise the position of the tongue body necessary for producing the high front vowel in order to achieve the tongue backing and lowering necessary for retroflexion. However, as we can see, the post-alveolar constriction location is not abandoned in favor of something more anterior; rather, the maneuver for achieving post-alveolar contact is altered, resulting in bunching of the tongue rather than retroflexion.

4. CONCLUSION

This study is a rather nice confirmation of the predictions of the task dynamic model of speech production: articulatory tasks are constant, but the actual articulatory maneuvers necessary (or possible) to achieve those tasks will vary based upon the state of the vocal tract. In the case of so-called retroflex consonants, this may mean that post-alveolar contact will occur, but may be achieved by bunching, rather than retroflexion, of the tongue. Our findings are reminiscent of those of Gick, Iskarous, Whalen, & Goldstein (2003), in which it was found that American English /r/ varied dialectally in its articulation despite the fact that constriction location remained stable. It seems that no matter the source of variation in articulatory maneuvering (dialectal differences, vowel context), constriction tasks remain stable. Taken together, these studies provide good evidence for the task-based nature of speech production.

ACKNOWLEDGMENTS

This work was funded through NSF Grant 1246750 to the University of Southern California.

REFERENCES

Gick, B., Iskarous, K., Whalen, D. H., & Goldstein, L. (2003). Constraints on Variation in the Production of English /r/. In S. Palethorpe & M. Tabain (Eds.), *Proceedings of the 6th International Seminar on Speech Production* (pp. 73–78).

Keane, E. (2004). Tamil. *Journal of the International Phonetic Association*, *34*(1), 111–116. doi:10.1017/S0025100304001549

Narayanan, Shri, 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.

Narayanan, Shrikanth, Bresch, E., Ghosh, P. K., Goldstein, L., Katsamanis, A., Kim, Y., Lammert, A., et al. (2011). A multimodal real-time MRI articulatory corpus for speech research. *Proceedings of the International Conference on Speech Communication and Technology*.

Narayanan, Shrikanth, Byrd, D., & Kaun, A. (1999). Geometry, kinematics, and acoustics of Tamil liquid consonants. *The Journal of the Acoustical Society of America*, *106*(4), 1993–2007. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10530023

Ohman, S. E. G. (1966). Coarticulation in VCV Utterances: Spectrographic Measurements. *Journal of the Acoustical Society of America*, *39*(1), 19–36.

Proctor, M. I., Bone, D., & Narayanan, S. (2010). Rapid semi-automatic segmentation of real-time Magnetic Resonance Images for parametric vocal tract analysis. *Proceedings of the International Conference on Speech Communication and Technology*.

Wiltshire, C., & Goldstein, L. (1997). Tongue Tip Orientation and Coronal Consonants. *Proceedings of the Eastern States Conference on Linguistics*.