ARTICULATORY COMPLEXITY AND LEXICAL CONTRAST DENSITY IN MODELS OF CORONAL COARTICULATION IN MALAYALAM

Indranil Dutta[1], Charles Redmon[2], Meghavarshini Krishnaswamy[1], Sarath Chandran[1], Nayana Raj[1]

[1]The English and Foreign Languages University, [2]University of Kansas
indranil@efluniversity.ac.in, redmon@ku.edu, meghavarshini@gmail.com, sarathchandran90@gmail.com, nr.nayanaraj@gmail.com

ABSTRACT

Measures of coarticulation from two models of consonant-vowel formant transitions – locus equations (LEs) and target-locus scaling (TLS) – are presented for Malayalam coronals. For both VC and CV transition models, LE slopes exhibited an ordering inconsistent with expectations based purely on estimates of articulatory complexity (i.e., retroflex < alveolar ≤ dental). In the CV context, LE slopes were found to be: tt < ʈʈ < t̪t̪, whereas in the VC context the ordering of dental and retroflex reversed to yield tt < t̪t̪ < ʈʈ. Alveolars consistently had the flatter slopes, a result predictable from their sparse representation in the lexicon. The retroflex < dental ordering in CV context is consistent with predictions based on articulatory complexity. The VC results suggest an asymmetry in coarticulation of retroflex stops with preceding and following vowels. These results suggest that coarticulatory resistance is mitigated by lexical imbalance. We discuss the production dynamics underlying these patterns through TLS models of the VC and CV formant trajectories.

Keywords: coarticulatory resistance, coronals, Malayalam, locus equations.

1. INTRODUCTION

Acoustic variation in the realization of phonetic segments in a language can be attributed to the size of the segmental inventory on the one hand [13, 14], and to resistance towards coarticulation on the other [16]. The resultant acoustic variation can be thought of as being a product of two constraints, contrastive and articulatory-motor. First, the degree of coarticulation resistance will shape the spread of coarticulation into-and-from neighbouring segments. Secondly, languages with large phoneme inventories, compared to those with relatively smaller ones, will have less diffuse realizations of phonemes in the same acoustic space. This variation in speech production has to be modeled by both human speech perception or any automatic speech recognition system in order to retrieve meaningful segmental contrasts.

Malayalam exhibits a number of coronal segments that makes it interesting to test the predictions of both the above-mentioned tendencies in languages; namely articulatory-motor and auditory-perceptual. Malayalam has a three-way contrast in place of articulation among plosives involving the tongue-tip (and also the tongue dorsum in the case of the retroflex), namely dental, alveolar and retroflex [8]. Thus, Malalyalam poses interesting questions on both these counts due to the fairly large number of contrasting segments in the dental/alveolar region.

It has been shown that segments with greater articulatory and motor constraints, especially those involving the tongue-dorsum, tend to resist coarticulatory influence from neighbouring segments more than those involving constrictions where the tongue-dorsum may not be involved [1, 17]. More recently, it has also been shown that in American English the retroflexed rhotic is highly resistant to coarticulation from neighbouring consonants and vowels [11]. Looking at the above mentioned constraints that affect phonetic variability it can be expected that the coronals in Malayalam will exhibit varying degrees of coarticulation resistance as predicted by the Degree of Articulatory Constraint (DAC) model [17]. Given that retroflex articulation of coronals must also involve tongue-dorsum gestures in addition to sub-apical gestures, the DAC will predict a higher degree of coarticulation resistance in Malayalam retroflexes (relative to dentals and alveolars). Similarly, the dentals, due to a blade articulation, should also exhibit higher degree of coarticulatory resistance when compared with alveolars. This tendency of dentals and retroflexes will predict variable degrees of coarticulatory resistance. Additionally, the acoustic variability in the density of contrasts in Malayalam (alveolars being the most sparse, with dentals and retroflexes more equivalent) would result from an interaction of the auditory-perceptual dimensions and the articulatory-motor constraints. Hence, it would be interesting to test whether the predictions of the DAC are borne out given the varying degrees of acoustic diffusion predicted by auditory-perceptual constraints.

As a further complication, Malayalam retroflex stops and nasals tend to exhibit a generally low F2 [8]. These findings lead to the assumption that F2 height may be a language-particular phenomenon [10]. While, there may not be a consensus on the role of F2 in the production of retroflexes, F3 measures, universally, tend to show lower values [8, 18]. Within this context findings from [12, 21, 9] are relevant in positing F3 as an important cue, especially in dense coronal systems. In that respect both the DAC and the auditory-perceptual constraints will predict variable interactions of F1, F2 and F3 values at consonant release and at the mid-point of the following vowels.

While segmental properties of coronals have been looked at to some extent in Hindi [15] and in Malayalam [8], it has yet to be shown what impact, if any, the density of the contrasts has on phonetic variability. In this paper, we present results from F2 locus equations [20] as derived from vowel onsets/offsets and vowel midpoints, and discuss the import of the locus equation slopes as an index of coarticulatory resistance. We find that while locus equation slopes offer complete discrimination between the alveolar-dental-retroflex contrast in Malayalam, there are narrower differences between retroflexes and dentals, consistent with cross-linguistic data on retroflex-dental stop contrasts. The inability of locus equations to capture differences in transition shape, suggests examination of an additional model – *target-locus scaling* [6, 5, 4] – that may offer clarifying information to the problem of modeling coarticulation in dense coronal systems.

2. METHODS

2.1. Participants

Ten female and ten male speakers between the ages of 17 and 23 were chosen for this study from nine districts of Kerala, with four or less (>20%) speakers from each district. All participants were students enrolled in Bachelors or Masters programs at the university and were chosen for having been raised and educated in Kerala at least till the age of 15. Participants reported no known speech or hearing difficulties. Of the 20 speakers, fifteen had studied Malayalam till grade 10, three received less than four years of formal instruction, and six reported having studied Malayalam at a junior college/Bachelors level.

2.2. Materials

Seventy-six items were chosen for this study, the majority of which were bisyllabic words. All items contained a $V_1C{:}V_2$ sequence, where V_1 was from the set /ə, o, i, u, e/, and V_2 was from the set /ə, a:, i, o, u, e/. The items were placed in a carrier sentence, and the block of sentences was repeated four times.

2.3. Recording

All recordings were carried out using a head-worn cardioid condenser microphone (Shure Beta 53), and digitized with a sampling rate of 22.5 kHz. The recordings were analyzed in Praat 6.0.4 [2]. Audio files were annotated manually, and verified by multiple annotators. For each VCV sequence, F2 was calculated from the Burg formant tracker in Praat, with formant tracks checked visually and hand-corrected for errors. For locus equations F2 was calculated at onset/offset (5% and 95% of the vowel, respectively) and midpoint. For the target-locus scaling model F2 was calculated at 9 equidistant points between vowel midpoint and consonant onset/offset (for VC and CV, respectively).

3. RESULTS

3.1. Locus equations

Locus equations were derived for each stop POA in each position (CV, VC) for each speaker according to Equation 1.

(1) $F2_c = \beta + \alpha F2_v$

where $\alpha = \rho\left(\frac{\sigma_c}{\sigma_v}\right)$ is the LE slope which serves as an index of V-to-C coarticulation, with steeper slopes corresponding to greater coarticulation. In the case of CV transitions, $F2_c$ corresponds to F2 5% into the following vowel, whereas for VC transitions $F2_c$ is measured at the 95% normalized time point from the preceding vowel. In both cases $F2_v$ is taken at vowel midpoint for either the following (V_2) or preceding (V_1) vowel.

Figure 1 displays the locus equations for CV transitions from dental, alveolar, and retroflex stops in Malayalam. Both female and male speakers show clear separation between alveolar stops and dentals/retroflexes, with alveolars having the flattest slopes, followed by retroflexes, then dentals.

In a mixed-effects model with POA predicting LE slopes, and Participant as a random intercept, significant differences between all three stops were obtained, with alveolars the most resistant to coarticulation (α = 0.431; α_{a-d} = -0.215, t = -10.39, p

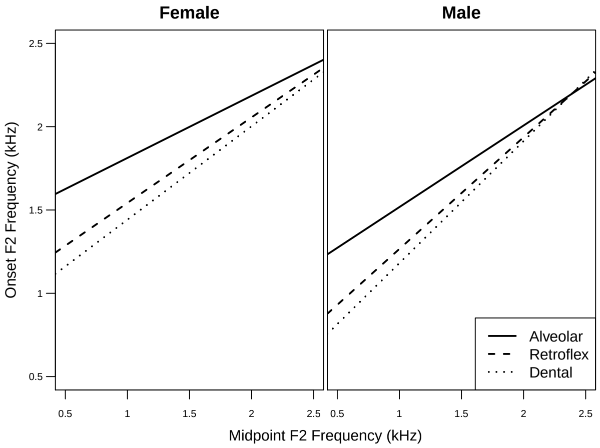

Figure 1: CV locus equations for female and male speakers. Aggregate lines are plotted in black.

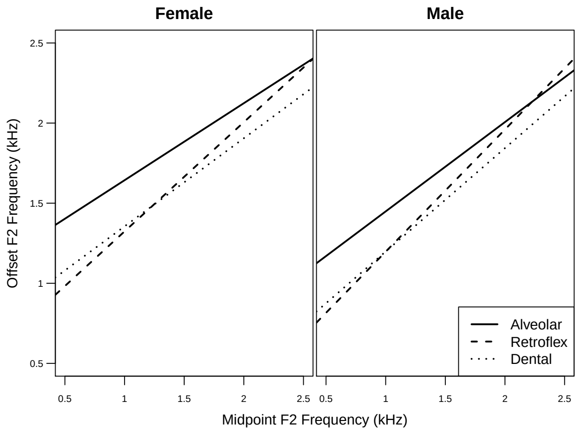

Figure 2: VC locus equations for female and male speakers. Aggregate lines are plotted in black.

< 0.001; α_{a-r} = -0.161, t = -7.761, $p < 0.001$), retroflexes the second-most resistant (α = 0.592, α_{d-r} = 0.055, t = 2.632, p = 0.012), and dentals the least resistant (α = 0.647). This pattern is consistent with both distributional and gestural (DAC) accounts of coarticulatory propensity, as the alveolar pattern is predicted from their extreme sparsity in the lexicon, while the retroflex pattern (relative to dentals) is expected given that all else equal, retroflex articulations are more complex and subject to less coarticulatory influence from the vowel.

In VC position (Figure 2), the same alveolar pattern was obtained (α = 0.519; α_{a-d} = -0.0785, t = -3.606, $p < 0.001$; α_{a-r} = -0.203, t = -9.329, $p < 0.001$), but retroflexes and dentals reverse order, with retroflexes coarticulating the most with the preceding vowel (α = 0.723), and dentals intermediate between the two (α = 0.598, α_{d-r} = -0.125, t = -5.724, $p < 0.001$). The reversal of relative degree of coarticulation between dentals and retroflexes in VC transitions is predictable both internally from a distributional account (retroflexes appear more commonly post-vocalically and word-finally than do dentals), and cross-linguistically from general typological preferences for post-vocalic retroflex stops, which are motivated both perceptually and articulatorily [10].

3.2. Target-locus scaling

While the locus equations do yield significant separation between all three places of articulation, in both CV and VC transitions, the close similarity in dental and retroflex slopes, as well as the general density of the coronal system, motivates further modeling of the dynamics of the transitions themselves. For this we employ the target-locus scaling approach of Broad and Clermont [3, 4, 5, 6], which is unlike the LE model in that it models formant transition patterns based on a vowel axis defined as

(2) $F_{cv}(n) = L_c + (T_v - L_c)K_c(n)$

where $F_{cv}(n)$ is the scaled formant transition (F = F1, F2, F3, etc.) in a given CV context, L_c is the 'locus' frequency of consonant C, T_v is the 'target' frequency of vowel V, and $K_c(n)$ is a scaling function that captures the similarity of formant transitions to/from a given consonant across a range of vowel contexts, as well as controlling the degree to which transitions tend to deviate from vowel target or consonant locus. The above parameters were calculated according to the procedure outlined in [3, 5, 6]. And while Equation 2 is defined for explicitly CV transitions, it is equally applicable to VC position.

To understand the coarticulatory characteristics of each stop implied by the model, we first examine the scaling function $K_c(n)$, which in addition to summarizing F2 transitions into and out of constrictions at each place of articulation, provides information on the degree to which vowel formants are compressed or rarefied (at a given point in the transition) relative to their target values. Then, we employ a variation on Broad and Clermont's (2010) derivation of the LE slope through the formula $\alpha = a_c(1)/a_c(10)$, which accounts for the transition shape in addition to onset and offset characteristics; namely, we take the mean of $a_c(n)$. For example, in Figure 3 we simulate two different families of formant transitions with identical onset and offset formant frequencies but different shapes, and consequently different coarticulatory profiles.

This measure, while necessarily related to the scaling function $K_c(n)$, depends on fewer assumptions about the ultimate form of the model, and as a summary of the function $a_c(n)$ that ultimately con-

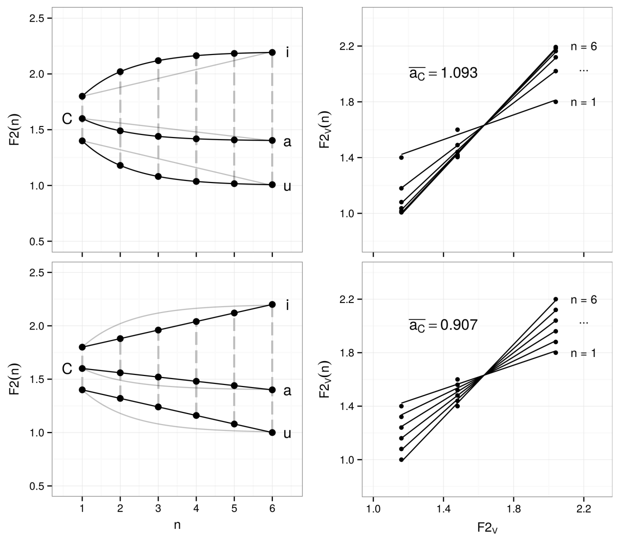

Figure 3: Simulation of the effect of changing contours on the scale parameter in the TLS model for trajectories with equivalent endpoints (and thus equivalent locus equations). With a lower $\overline{a_c}$, the bottom panel exhibits greater coarticulatory resistance from the consonant.

taines the locus equation slope as a special case, represents a useful link between the two approaches.

Figure 4 shows VC and CV transition scaling functions $K_c(n)$ per consonant place of articulation, aggregated across participants. For the transition into the stop closure, alveolars and dentals are strikingly similar, though alveolars exert a greater influence on the vowel near the point of closure (as evidenced by their greater compression of the vowel formant ensemble, VFE; see [5, 6] for detail). This difference is consistent with the slightly shallower locus equation slope for the alveolars in VC position.

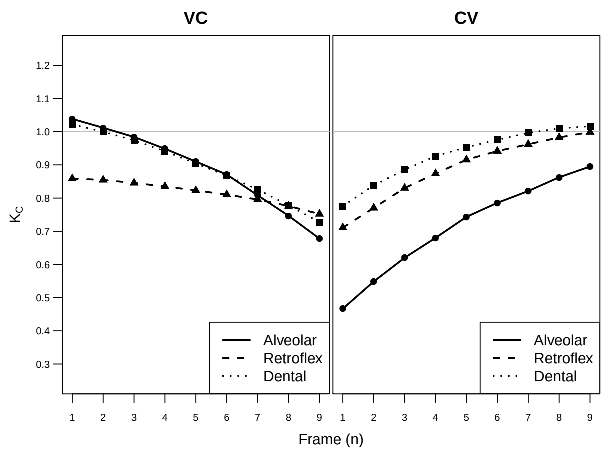

Figure 4: Scaling functions $K_C(n)$ by stop POA, for VC and CV transitions, aggregated across participants.

Retroflexes, on the other hand, show a notably different pattern. Even at midpoint, vowel formants are substantially compressed relative to their targets, which means that the steeper LE slope for retroflexes was due not to less coarticulatory resistance from the consonant, but the opposite. This result highlights the critical alternative way in which vowel midpoints and offsets/onsets can be highly correlated (i.e., yielding steeper LE slopes): Vowel midpoints can be driven significantly by the adjacent consonant. This point is related to the discussion of vowel *aggression* in [7], though Chen and colleagues utilize a different form of the locus equation to address C-to-V coarticulation. On the contrary, what the retroflex VC transition illustrates is that such correlations between $F2_v$ and $F2_c$ (as in Figure 2) may still be obtained for vowels that have generally been shifted to be closer to the consonant locus. Such a case merely implies that the domain and range of F2 values are compressed relative to other contexts.

The transition profiles in CV position are more in line with expectations, with retroflexes and dentals similar and both showing little deviation in vowel formants from their targets by midpoint. By contrast, alveolars show considerably less convergence on vowel targets, but with a larger K_c range, unlike in the retroflex VC case above.

Across participants, $\overline{a_c}$ was consistent with LE slopes in CV position (i.e., alveolar $<$ retroflex $<$ dental; ps < 0.01), but in VC position reflected the fact that preceding vowels were shifted substantially toward the retroflex (i.e., retroflex $<$ alveolar = dental), thus indicating that the two models diverge in certain critical areas with regard to their coarticulatory predictions.

4. DISCUSSION

Models of F2 transitions and derived locus equations have successfully modeled both place of articulation discrimination [20] and coarticulatory resistance [16] due to tongue-tip, tongue-blade, and tongue-body engagement. However, dense coronal systems, such as in Malayalam, have posed challenges for such models [19, e.g.]. In this paper, we first show that such systems can in some cases be accounted for in the LE model, but that Malayalam's unique lexical distribution may be partly responsible for this result. Contrary to the predictions of the DAC, Malayalam alveolars resist coarticulation more than the retroflexes in CV position. However, the VC data, though consistent in LEs, reveal the opposite pattern in the TLS model, providing greater motivation for attention to the full formant transition in coarticulation models.

5. REFERENCES

[1] Bladon, R. A., Al-Bamerni, A. 1976. Coarticulation resistance in English /l/. *Journal of Phonetics* 4, 135–150.

[2] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9/10), 341–345.

[3] Broad, D. J., Clermont, F. 1987. A methodology for modeling vowel formant contours in CVC context. *The Journal of the Acoustical Society of America* 81(1), 155–165.

[4] Broad, D. J., Clermont, F. 2002. Linear scaling of vowel-formant ensembles (VFEs) in consonantal contexts. *Speech Communication* 37(3-4), 175–195.

[5] Broad, D. J., Clermont, F. 2010. Target–locus scaling methods for modeling families of formant transitions. *Journal of Phonetics* 38(3), 337–359.

[6] Broad, D. J., Clermont, F. 2014. A method for analyzing the coarticulated CV and VC components of vowel-formant trajectories in CVC syllables. *Journal of Phonetics* 47, 47–80.

[7] Chen, W.-r., Chang, Y.-c., Iskarous, K. 2015. Vowel coarticulation: Landmark statistics measure vowel aggression. *The Journal of the acoustical society of America* 138(2), 1221–1232.

[8] Dart, S. N., Nihalani, P. 1999. The articulation of Malayalam coronal stops and nasals. *Journal of the International Phonetic Association* 29, 129–142.

[9] Dave, R. A formant analysis of the clear, nasalized, and murmured vowels in gujarati. *Indian Lingustics* 28, 1–30.

[10] Hamann, S. 2003. *The Phonetics and Phonology of Retroflexes*. Utrecht: LOT Press.

[11] Kochetov, A. 2011. Alveolar-to-rhotic coarticulation in North American English: A preliminary epg study. *Proceedings of the 17th International Congress of Phonetics Sciences (ICPhS 2011)* 1118–1121.

[12] Krull, D., Lindblom, B. 1996. Coarticulation in apical consonants: acoustic and articulatory analyses of hindi, swedish and tamil. *Quarterly Progress and Status Report of the Dept. for Speech, Music and Hearing* 37(2), 73–76.

[13] Manuel, S. 1990. The role of contrast in limiting vowel-to-vowel coarticulation indifferent languages. *Journal of the Acoustical Society of America* 88, 1286–1298.

[14] Manuel, S. 1999. Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In: Hardcastle, W., Hewlett, N., (eds), *Coarticulation: Theory, data and techniques*. Cambridge University Press.

[15] Ohala, M., Ohala, J. 2001. Acoustic vc transitions correlate with degree of perceptual confusion of place contrast in Hindi. *Travaux du Cercle Linguistique de Copenhague* 31, 265–284.

[16] Recasens, D., Espinosa, A. 2009. An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America* 125, 2288–2298.

[17] Recasens, D., Pallarès, M. D., Fontdevila, J. 1997. A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America* 102, 544–561.

[18] Shalev, M., Ladefoged, P., Bhaskararao, P. 1993. Phonetics of Toda. *UCLA Working Papers in Phonetics* 84, 89–123.

[19] Sussman, H. M., Hoemeke, K. A., Ahmed, F. S. 1993. A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation. *The Journal of the Acoustical Society of America* 94(3), 1256–1268.

[20] Sussman, H. M., McCaffrey, H. A., Matthews, S. A. 1991. An investigation of locus equations as a source of relational invariance for stop place categorization. *The Journal of the Acoustical Society of America* 90(3), 1309–1325.

[21] Tabain, M., Breen, G., Butcher, A. 2004. VC vs. CV syllables: a comparison of aboriginal languages with english. *Journal of the International Phonetic Association* 34(2), 175–200.