ISSN 0973-3302

THE JOURNAL OF ACOUSTICAL SOCIETY OF INDIA

Volume 51 Number 4 October 2024

A Quarterly Publication of the ASI
https://acoustics.org.in

The Journal of Acoustical Society of India
Volume 51 Number 4 October 2024 pp 212-275

The Journal of Acoustical Society of India : Vol. 51, No. 4, 2024 (pp. 268-275)

Fundamental frequency perturbation is a consequence of the biomechanics of voicing and not due to phonological status

Indranil Dutta[1*] and Molly Varghese[2]
[1]*School of Languages and Linguistics, Jadavpur University, Kolkata, India*
[2]*Department of Malayalam, Sree Krishna College, Guruvayur, India*
e-mail: indranildutta.lnl@jadavpuruniversity.in

[Received: 11-10-2024; Accepted: 20-01-2025]

ABSTRACT

Voicing is known to lower f_0 in the following vowel[18,21,23]. This lowering has been attributed to biomechanical constraints[22,28-30]. On the basis of evidence from typologically, aspiration languages, such as English, where the [voice] feature doesn't manifest itself phonetically, yet is accompanied by f_0 lowering, it has been argued that this lowering serves to enhance the phonological [voice] contrast. We present results from an acoustic study of f_0 perturbation in Malayalam, which exhibits predictable voicing in lexical items from its Dravidian stratum. Voiceless stops appear initially, and are voiced intervocalically. Malayalam also contains a Sanskrit lexical stratum, where phonetic voicing is tied to its stratal affiliation. Results from two experiments on voicing and f_0 perturbation show that even with predictable voicing, in the Dravidian stratum, Malayalam exhibits the crosslinguistic tendency to lower f_0 in the following vowel. We conclude that f_0 lowering is an automatic consequence of the biomechanics of voicing in Malayalam. At the representational level, such lowering could indeed be enlisted for enhancement, however, that is not the only, or the primary purpose of f_0 perturbation.

1. INTRODUCTION

Voicing and f_0, universally, exhibit a covariation, where voiced obstruents are known to lower onset f_0 in the following vowel, while voiceless obstruents tend to raise the f_0. In typologically aspiration languages, such as English, despite there being a phonological contrast of the feature [voice], phonetically, at least word-initially, this contrast is implemented through aspiration. The accompanying f_0 lowering prompts many to argue that f_0 lowering in aspiration languages serves to enhance the phonological [voice] contrast. In typical, voice languages, that is the ones that indeed have closure voicing, as well, universally onset f_0 perturbation is found. In this paper, we investigate the nature of f_0 perturbation in Malayalam, where in the Dravidian stratum lexical items voicing is predictable, while in the Sanskrit stratum lexical items phonetic voicing does accompany the underlying voiced stops. Our results indicate that voicing and f_0 covariation is a consequence of the biomechanics of voicing and hence automatic. We report on two experiments that help us make this claim. Phonological representations, however, may indeed enlist micro-fluctuations of f_0 to enhance a [voice] contrast.

In section 2 of this paper, we describe in detail the two major accounts that help us understand the nature and role of f_0 perturbation. Following that, in section 3 we detail our experimental methods and the materials used for our study. In section 4, we present the results from two experiments on f_0 perturbation in Malayalam, and following that in section 5, we stress on the need to understand the biomechanical bases of f_0 perturbations.

2. VOICING AND f_0 PERTURBATION

While voicing is universally known to depress onset f_0 at least in the first 10 ms of the following vowel, voicelessness as a laryngeal setting is known to have the opposite effect[18,21,23]. Regardless of the phonetic manifestation of the nature of the voicing contrast, i.e., languages with voiceless stops accompanied by aspiration such as English, and those with stop voicing such as French and Spanish[19,25]. Within the literature, the covariation between voicing and f_0 has been shown to function as an enhancement feature for the feataure [voice], where even phonetic voicing may be absent or mitigated by contextual influences[31]. On the other hand, this covariation has also been understood as governed by automatic biomechanical constraints, regardless of the phonological status of voicing in a given language[22]. While the former view ascribes to the notion that voicing related f_0 perturbation in the onset of the vowel cues voicing contrast where even primary voicing cues may be absent or obliterated, the latter view maintains that the acoustic perturbation of f_0 accompanying voicing is a consequence of physiological constraints that have to do with the loosening of the vocal folds[22]. A somewhat unified account is provided by[24 and 20] who show that phonetic covariation could at once be used for enhancement of contrasts as well as be physiologically governed.

For the intents of this paper, we refer to the covariation as enhancement argument for f_0 perturbation as the phonological account, and the physical-biomechanical explanations for f_0 lowering following voiced obstruents, and raising following voiceless obstruents as the phonetic account. This characterization is operational, rather than representational. Crucial to both these accounts is the phonological status of the feature [voice], since f_0 lowering following [voice] obstruents and concomitant raising following voiceless obstruents implies an underlying [voice] contrast regardless of the phonetic manifestation of such a contrast. While English contrasts the feature [voice], the phonetic manifestation of this contrast is largely cued by variable Voice Onset Times (VOT), in addition to the presence of [voice] related f_0 perturbation. French and Spanish[19,25], on the other hand, represent languages where the phonological feature [voice], phonetically manifests itself in closure voicing, essentially -VOT, where voicing related f_0 perturbation patterns itself with universal tendencies; mainly lowering following voiced, and raising following voiceless obstruents. In as much as these languages could be referred to as 'voice' languages compared to English, which could be characterized as an 'aspiration' language, the voicing related f_0 perturbation could either be understand as [voice] feature enhancement or due to biomechanical constraints. Hence, typological 'voice' languages may indeed pose difficulty in unequivocally attributing f_0 per-turbation due to enhancement of a known contrast or biomechanical constraints. Within this context, the phonetic manifestation and phonological status of the feature [voice] in Malayalam provides a potential third typological variant, distinct from aspiration and voice languages, due to the presence of Dravidian and Sanskrit vocabulary lexical items. While speakers may not have explicit knowledge of the etymology of Dravidian and Sanskrit source lexical items, they may indeed group these words according to the variable application of phonological rules to Dravidian and Sanskrit strata lexical items. To that end[26] proposes the features [+Dravidian] and [+Sanskrit] that are specified in the lexicon for the application of rules sensitive to these features. For instance, in Malayalam compounding, if the stem-initial stop of the second stem in the compound has the feature [+Dravidian] then that stop is geminated while if it is [+Sanskrit] it is not. This reflects the ubiquitous presence of intervocalic voiceless geminates in Malayalam and absence of voiced geminates in this environment. The Dravidian stratum and vocabulary as expressed in[26] consists of predictable, non-contrastive voicing, the Sanskrit stratum consists of several high-usage lexical items that exhibit phonetic voicing and also share the four-way laryngeal contrast seen in typical Indo-Aryan languages. The status of the feature [voice] in Malayalam, therefore, is tied to the stratal affiliation of the lexical item. In this paper, we would like to offer a two-fold view; first that in order to ascertain the phonetic and phonological basis for f_0 perturbation, it is crucial to look at languages such as Malayalam, where the phonological status of the feature [voice] is ambiguous at best, or stratally determined. Secondly, the presence of f_0 perturbation in Malayalam voiced and voiceless obstruents would lend support to the claim that f_0 lowering has physiological and biomechanical basis as has been shown by[22].

Stiffening and loosening of the vocal folds is governed by three muscles, namely, Cricothyroid (CT), vocalis, and Thyroartynoid (TA)[29]. Voicing during stop closure is enabled by lowering the larynx which helps expand the supraglottal cavities in order to allow voicing during oral closure. During this period, the relaxation in the Cricothyroid (CT) muscle resets the vocal fold tension and results in f_0 lowering[22]. Cricoid cartilage (CC) moves downward making the posterior plate parallel to the curvature of the cervical spine. This rotation of the Cricoid cartilage CC leads to shortening of the vocal folds.

3. MATERIALS AND METHODS

3.1 Experiment 1

Data were collected from 6 native speakers of Malayalam, 3 male and 3 female. The target words consisted of word-initial (I) and word medial (M) consonants: 6 C(vl)V and 6 Nasal+Vowel (NV) words for the I context, 3 $V_1C(vd)V_2$, 4 $V_1C{:}(vl)V_2$ words and 6 $V_1N{:}V_2$ words for the M context. The 25 target words along with distractors were embedded in a carrier sentence, with 4 repetitions of each word. Time normalized f_0 was measured at 10 intervals into the vowel using ProsodyPro[32], 5 ms offset from the vowel onset. Total 600 items made up the corpus for this experiment. Since the Dravidian stratum lexical items do not exhibit initial voicing, nor intervocalic voiceless singletons, singleton nasals were used in place of voiced stops. In an effort to control for any effects of gemination on f_0 nasal geminates were used to contrast with voiceless geminates.

3.2 Experiment 2

The data were collected from 6 native speakers of Malayalam, 3 male and 3 female. Target words consisted of word-initial (I) and word medial consonants (M): 9 C(vl)V and C(vd)V words for the I context, where C(vd)V words were mainly words from the Sanskrit stratum that are in common use among Malayalam speakers (as confirmed from a Malayalam corpus). 9 $V_1C(vl)V_2$ and $V_1C(vd)V_2$ words for the M context. 2 prosodic conditions, strong and weak, were also used. This led to a total of 27 words* 2 (I and M) * 2 prosodic conditions * 4 repetitions * 6 subjects = 2592 items in all. Malayalam is reported to exhibit intervocalic lenition of voiced stops[27] such that they become frictionless approximants. These lenited stops were excluded from our study, hence the final number of items analysed were 2056. Since intervocalic stops get voiced in Malayalam, $V_1C(vl)V_2$ items were taken from compounds where the second stem is a borrowed word from Sanskrit. According to compounding rules in Malayalam, second stem Sanskrit compounds, unlike second stem Dravidian compounds, do not undergo gemination. Instead, the vowel at the end of the first stem is lengthened[26], and the initial consonant of the second stem remains singleton voiceless, as observed from the examples from[26]. As in Experiment 1, the target words along with distractors were embedded in a frame sentence, with 4 repetitions of each word. Time normalized f_0 was measured at 10 intervals into the vowel with ProsodyPro[32], 5 ms offset from the vowel onset. For both the experiments data was manually segmented and annotated. In order to control the effects of gender and subject, all f_0 values in Hz were converted to their corresponding z-score normalized values grouped by subject for both experiments.

4. f_0 PERTURBATION IN MALAYALAM

Results from both experiment 1 and experiment 2 show that despite the phonological status of [voice] in Malayalam, voiced stops significantly lower f_0 in the following vowel. There is also a significant effect of segment context, with word-initial stops, lowering f_0 more than the word-medial stops. Results from Experiment 1 show that the z-score normalized f_0 (zf_0) values are lower for voiced stops and nasals compared to their voiceless counterparts (Fig. 1). A noticeable effect is also the segment position, *i.e.*, the word-initial stops and nasals have lower onset f_0 compared to the stops and nasals in word-medial position.

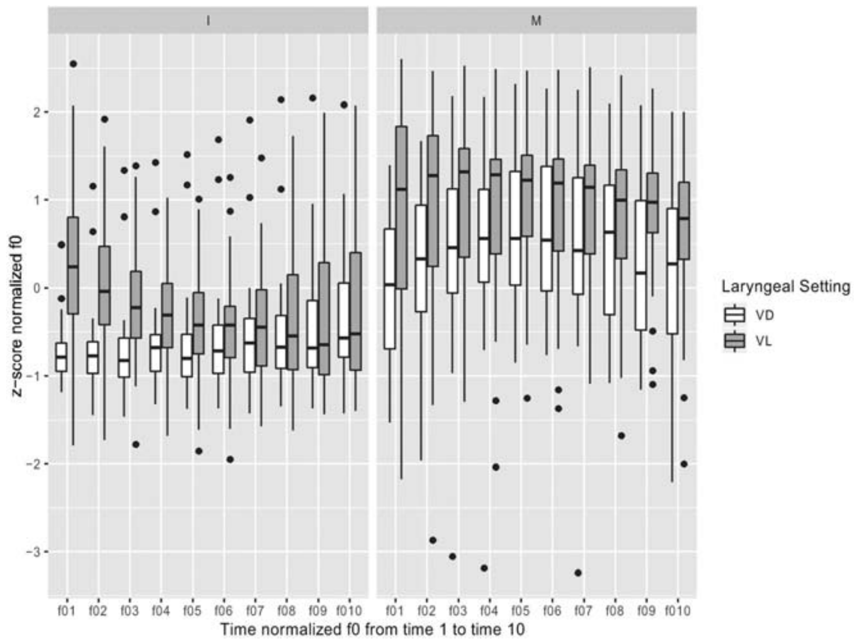

Fig. 1. z-score normalized f_0 values for 10 time normalized periods in the following vowel (Experiment 1).

As can be seen in Fig. 2, we do not find an effect of manner of articulation on the onset f_0, and nasals are not known to effect f_0 perturbation on account of nasal airflow or velic lowering. However, there is a significant interaction between nasal stops and context. Word initial nasals have lower f_0 compared to word medial nasals. We elaborate on this interaction and the implications for our study below.

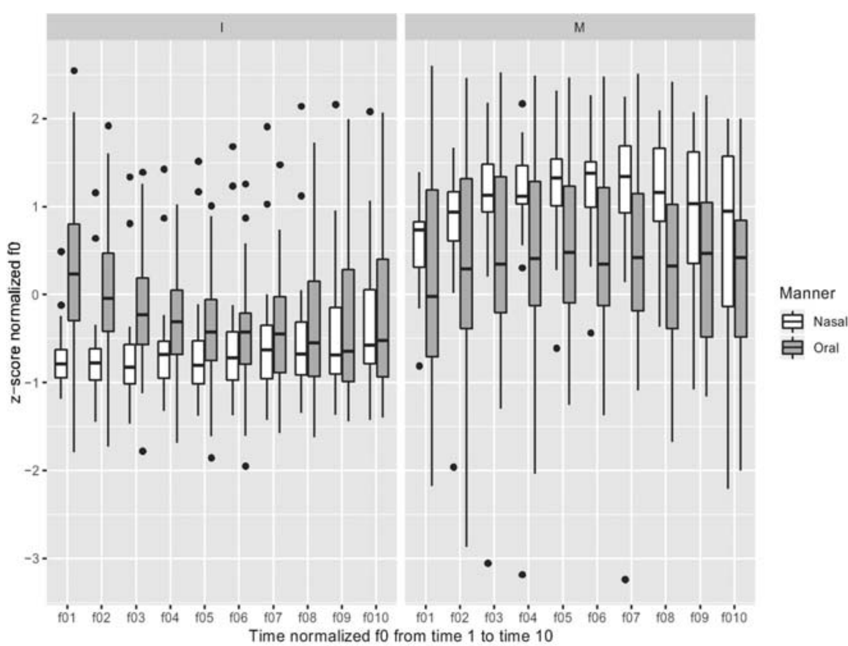

Fig. 2. z-score normalized f_0 values for 10 time normalized periods for oral and nasal manner of articulation (Experiment 1).

We performed a linear mixed effects analysis of the relationship between zf_0 and laryngeal setting, voiceless (VL) and voiced (VD), and with context; word initial (I) and word media (M)[17]. The fixed effects, laryngeal setting and context were fitted on the model. An interaction term between context and manner of articulation was also included in the model. The model also consisted of items (words), and iteration as random effects. We obtained p-values by using likelihood ratio tests, i.e., comparison of the full model with the null model. At T1, corresponding to f_0 at 5% of the vowel, zf_0 is significantly predicted by laryngeal setting, p=0.02643, while context and manner interaction is also found to have a significant effect, p=0.022. At T2, i.e., 20% of the vowel, zf_0 is not significantly predicted by laryngeal setting and neither

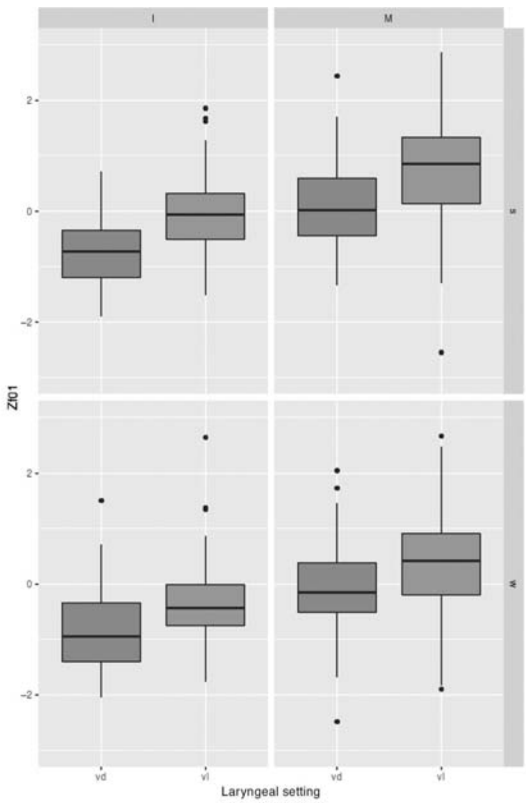

Fig. 3. z-score normalized f_0 values for time T1 that corresponds to 10% of the vowel.

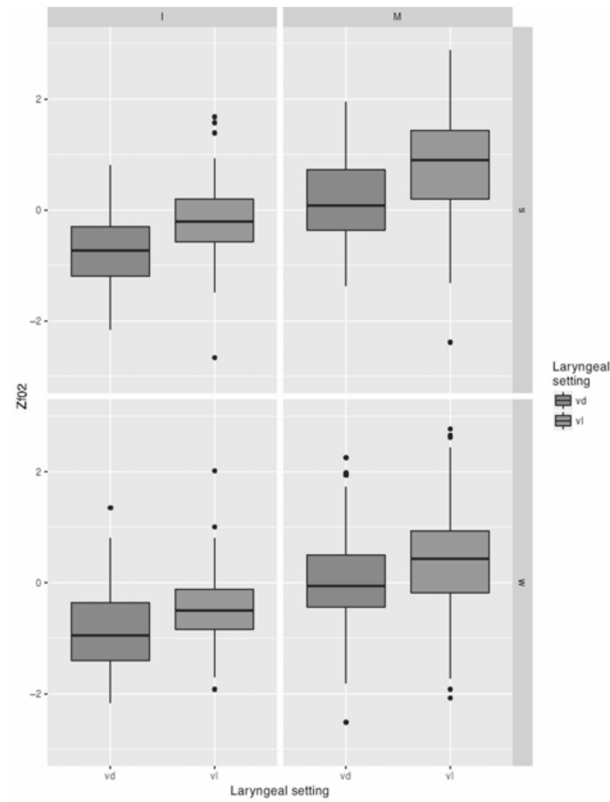

Fig. 4. z-score normalized f_0 values for time T2 that corresponds to 20% of the vowel.

context nor manner show significant effects on zf_0. We did not find any significant effect of laryngeal setting on zf_0 for the remaining time frames.

As can be seen in Fig. 4 the z-score normalized f_0 for voiced stops is lower than that of the voiceless stops at 10% of the vowel. Similar to the results in experiment 1, there is an effect of segmental context, in that word-initial stops lower f_0 in general, compared to word-medial stops. This pattern persists till about 30% of the vowel, with there being an interaction between the effect of the laryngeal setting and segmental context. The effect of the laryngeal setting wanes into the vowel, from it being the greatest at vowel onset, compared to the effect of the segmental context. We also do find a marginal effect of prosodic context; a focal-strong position and non-focal weak position.

For experiment 2, we performed a linear mixed effects analysis of the relationship between zf_0 and laryngeal setting with two levels; voiced (VD) and voiceless (VL). The fixed effects in our model were laryngeal setting and context; word initial (I) and word medial (M), prosodic position (strong and weak) and gender. The random effects had intercepts for subjects, items, and iterations, as well as by-subject, by-item and by-iteration random slopes for the effect of laryngeal setting. We obtained a p-value of p=0.0.011792 by conducting likelihood ratio tests of the full model with laryngeal setting against a model without laryngeal setting for the effect of laryngeal setting on zf_0 at vowel onset (T1). Introducing an interaction term between laryngeal setting and segmental context we obtained a p-value of p=0.0355 which was significant at the $p<0.05$. These results suggest the laryngeal setting, voiced or voiceless is significant in determining the onset f_0 of the vowel, regardless of the phonological status of voicing in a language. In addition we find that the effect of segmental context; word-initial and word-medial, is such that f_0 is significantly lower word initially than word medially. Prosodic context was also found to have a significant effect on zf_0 at vowel onset (T1) with significantly higher zf_0 values for stronger prosodic positions compared to the weaker prosodic context.

5. PHYSIOLOGICAL FACTORS BEHIND f_0 PERTURBATION

Microprosodic local f_0 fluctuations most definitely play a role in enhancement of segmental contrasts, in this case a laryngeal contrast (where presumably there is none). The processes involved, however, for these local fluctuations are patently biomechanical, and phonological representations recruit these available acoustic cues for discrimination. Cricothyroid (muscle) applies a sudden stretch to the vocal folds to stop vocal fold vibration for voiceless stops and hence raises f_0, the Cricothyroid due to the rotation shortens the length of the folds which lowers the f_0 in voiced stops. In this paper, we observe that Malayalam due to the variable nature of the status of the feature [voice] in its lexicon provides an excellent testing ground for the automatic versus controlled accounts of f_0 perturbation. Our results indicate that f_0 perturbations are biomechanical in nature, even in a typologically non-contrastive [voice] language, which also exhibits phonetic voicing in the Sanskrit stratum.

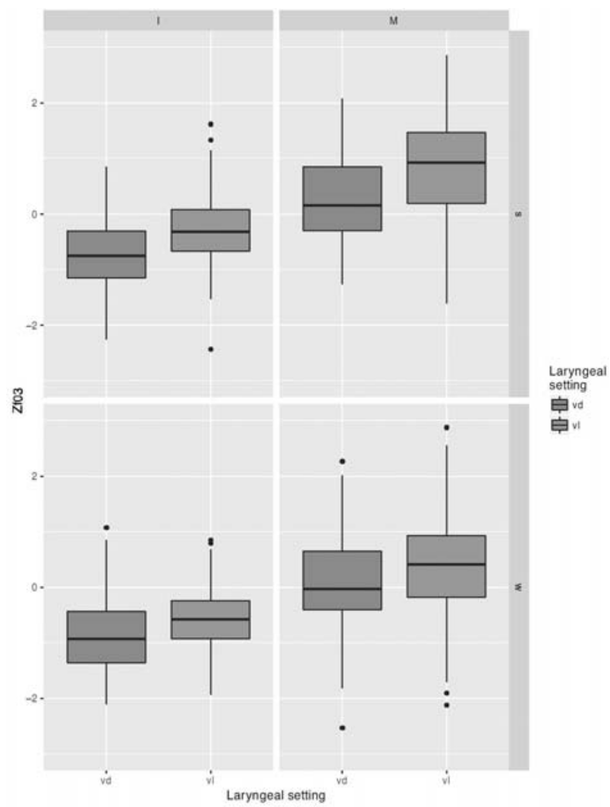

Fig. 5. z-score normalized f_0 values for time T3 that corresponds to 30% of the vowel.

6. ACKNOWLEDGMENT

We acknowledge the help provided by Jasmine Maria George and Minu Sara Paul during collection of acoustic data for experiment 1.

7. REFERENCES

[1] D. Bates, M. Mächler, B. Bolker and S. Walker, 2015. Fitting linear mixed-effects models using lme4 *Journal of Statistical Software*, 67(1), 1-48, doi: 10.18637/jss.v067.i01.

[2] L. Chistovich, 1969. Variations of the fundamental voice pitch as a discriminatory cue for consonants Soviet Phys. *Acoust*, 14, 372-378.

[3] O. Dmitrieva, F. Llanos, A. A. Shultz and A. L. Francis, 2015. Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English Journal of Phonetics 49, 77-95, http://www.sciencedirect.com/science/article/pii/S0095447014001053, doi: http://dx.doi.org/10.1016/j.wocn.2014.12.005.

[4] L. L. Holt, A. J. Lotto and K. R. Kluender, 2001. Influence of fundamental frequency on stop-consonant voicing perception: A case of learned covariation or auditory enhancement? *The Journal of the Acoustical Society of America*, 109(764), 131-174.

[5] J. M. Hombert, 1978. Consonant types, vowel quality, and tone in Tone: A linguistic survey, edited by V. A. Fromkin (Academic Press, New York), pp. 77-111.

[6] K. Honda, 2004. Physiological factors causing tonal characteristics of speech: from global to local prosody Speech Prosody .

[7] A. S. House and G. Fairbanks, 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America*, 25, 105-113.

[8] J. Kingston and R. L. Diehl, 1994. Phonetic knowledge in Language, edited by S. Everson (Cambridge University Press), pp. 419-454.

[9] A. Löfqvist, T. Baer, N. S. McGarr and R. S. Story, 1989. The cricothyroid muscle in voicing control. *The Journal of the Acoustical Society of America*, 85, 1314-1321.

[10] K. Mohanan, 1986. Studies in Natural Language and Linguistic Theory. *The Theory of Lexical Phonology* (Reidel), https://books.google.co.in/books?id=bGZiAAAAMAAJ.

[11] K. Mohanan and T. Mohanan, 1984. Lexical phonology of the consonant system in malayalam Linguistic Inquiry, 15, 575-602.

[12] R. N. Ohde, 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America*, 75, 224-230.

[13] K. Stevens, 2001. *Acoustic Phonetics* (MIT Press).

[14] J. O. Svantesson and D. House 2006. Tone production, tone perception and kammu tonogenesis. *Phonology*, 23, 309-333, http://journals.cambridge.org/article_S0952675706000923, doi: 10.1017/S0952675706000923.

[15] D. H. Whalen, A. S. Abramson, L. Lisker and M. Mody, 1993. F_0 gives voicing information even with unambiguous voice onset times. *The Journal of the Acoustical Society of America*, 93, 2152-2159.

[16] Y. Xu, 2013. Prosodypro - a tool for large-scale systematic prosody analysis. *In Proceedings of Tools and Resources for the Analysis of Speech Prosody*, pp. 7-10.

[17] Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48, doi: 10.18637/jss.v067.i01.

[18] L. Chistovich, 1969. Variations of the fundamental voice pitch as a discriminatory cue for consonants. *Soviet Phys. Acoust*, 14, 372-378.

[19] O. Dmitrieva, F. Llanos, A. A. Shultz and A. L. Francis, 2015. Phonological status, not voice onset time, determines the acoustic realization of onset f_0 as a secondary voicing cue in Spanish and English. *Journal of Phonetics*, 49, 77-95, http://www.sciencedirect.com/science/article/pii/S0095447014001053, doi: http://dx.doi.org/10.1016/j.wocn.2014.12.005.

[20] L. L. Holt, A. J. Lotto and K. R. Kluender, 2001. Influence of fundamental frequency on stop-consonant voicing perception: A case of learned covariation or auditory enhancement? *The Journal of the Acoustical Society of America*, 109(764), 131-174.

[21] J. M. Hombert, 1978. Consonant types, vowel quality, and tone in Tone: A linguistic survey, edited by V. A. Fromkin (Academic Press, New York), pp. 77-111.

[22] K. Honda, 2004. Physiological factors causing tonal characteristics of speech: from global to local prosody Speech Prosody .

[23] A. S. House and G. Fairbanks, 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America*, 25, 105-113.

[24] J. Kingston and R. L. Diehl, 1994. Phonetic knowledge in Language, edited by S. Everson (Cambridge University Press), pp. 419-454.

[25] A. Löfqvist, T. Baer, N. S. McGarr and R. S. Story, 1989. The cricothyroid muscle in voicing control. *The Journal of the Acoustical Society of America*, 85, 1314-1321.

[26] K. Mohanan, 1986. Studies in Natural Language and Linguistic Theory. *The Theory of Lexical Phonology* (Reidel), https://books.google.co.in/books?id=bGZiAAAAMAAJ.

[27] K. Mohanan and T. Mohanan, 1984. Lexical phonology of the consonant system in malayalam. *Linguistic Inquiry*, 15, 575-602.

[28] R. N. Ohde, 1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America*, 75, 224-230.

[29] K. Stevens, 2001. *Acoustic Phonetics* (MIT Press).

[30] J.O. Svantesson and D. House, 2006. Tone production, tone perception and kammu tonogenesis. *Phonology*, 23, 309-333, http://journals.cambridge.org/article_S0952675706000923, doi: 10.1017/S0952675706000923.

[31] D. H. Whalen, A. S. Abramson, L. Lisker and M. Mody, 1993. F_0 gives voicing information even with unambiguous voice onset times. *The Journal of the Acoustical Society of America*, 93, 2152-2159.

[32] Y. Xu, 2013. Prosodypro - a tool for large-scale systematic prosody analysis. *In Proceedings of Tools and Resources for the Analysis of Speech Prosody*, pp. 7-10.