Acoustic Measurement of Laryngeal Constriction in Thai Consonants

Jeremy Perkins (University of Aizu)
jperkins@u-aizu.ac.jp

1. Introduction

A production experiment was conducted with two main goals: (1) To use a new acoustic method to find laryngealized consonants, via a technique that has recently been used to identify creaky phonation in vowels; and (2) to answer a phonological question on whether certain Thai consonants are [+constricted glottis] ([+CG]), as posited in Perkins (2011).

Concerning the second goal, systematic gaps exist in Thai syllables, where high tone is absent following voiced and voiceless unaspirated stop consonants (Perkins, 2011; Lee, 2008). These gaps have been treated as phonological restrictions, due to native Thai speakers' grammaticality judgments (Ruangjaroon, 2006; Perkins, 2011). Perkins (2011) and Lee (2008) analyzed these via avoidance of a marked sequence of a consonant and a tone. There is widespread evidence cross-linguistically of voiced consonants lowering F0, with many languages phonologizing this sequence (Bradshaw, 1998). However, to explain why voiceless unaspirated stops would also be involved, Perkins (2011) posited that they are [+constricted glottis] ([+CG]). [+CG] consonants are often avoided adjacent to high tone, cross-linguistically.

Perkins's (2011) analysis was based on phonetic evidence from three Thai speakers. The present experiment expands the number of speakers to twenty. In addition to spectral tilt and F0 perturbations, psychoacoustic roughness is measured at vowel edges to identify coarticulatory laryngeal constriction induced by laryngealized consonants. Roughness correlates with laryngeal constriction in vowels, as it successfully identified creaky phonation in Burmese and Zhuang (Villegas et al., 2020). It is defined as a sensory attribute related to rapid changes on the amplitude envelope of a sound at specific frequencies between 15 to 300 Hz. These specific frequencies involved are excited to a greater extent during sounds that humans label as being "rough". As such, psychoacoustic roughness relates to the way a "rough" sound is perceived by humans, and may correlate with creakiness in languages, offering a unique measurement that focuses on the perception end of the speech chain, rather than the production end (as spectral tilt does). While Villegas et al. (2020) found that roughness discerns creaky phonation from modal phonation, it is not yet clear how roughness correlates with other non-modal phonations (i.e. breathy phonation). Similarly, creaky phonation is only one of the articulatory modes of laryngeal constriction, and so tense phonation, vocal fry and others (Esling et al., 2005) may not correlate with roughness. Hypothetically, since it focuses on the perception end of speech, roughness will mirror the phonological fact that languages do not allow these modes of laryngeal constriction to contrast. As such, it may be more appropriate as an indicator of phonological creakiness than spectral tilt.

This hypothesis is tested in the present experiment, by including fricatives and aspirated stops, which will induce breathiness in a following vowel. Unaspirated stop onsets are compared with contexts in which laryngeal constriction is known to exist: Glottal stops in both onset and coda and obstruent codas (Morén & Zsiga, 2006). On the other hand, nasals and glides are included to provide control contexts that lack laryngealization.

Section 2 summarizes the methods used in the production experiment and section 3 presents the results. Section 4 presents the discussion and conclusions.

2. Methods

2.1. Materials

In a production experiment, 20 native Thai speakers were recorded via acoustic recording. A list of 136 Thai monosyllabic words containing all syllable types and tones was repeated five times, yielding a total of 680 words per speaker. Onset consonants were restricted to bilabial and alveolar place and only the low vowel [a] was used.

2.2. Procedure

Twenty native (4 male) Thai speaking participants were recruited in Bangkok, Thailand. 19 participants spoke Central Thai natively; one participant from Nakhorn Phanom spoke northeastern Thai. Each participant was recorded in a quiet environment via an AKG C520 head-worn condenser microphone connected to a Marantz PMD661 MKII Professional Portable Flash Field Recorder. Wav files were created with a sampling rate of 48 kHz. Recording sessions took place over two consecutive days and were approximately 30 minutes each.

2.3. Analysis

The recordings were annotated in Praat with TextGrid segments for onsets, vowels, and codas. Measurement of F0, vowel-normalized spectral tilt (H1*–A1* was selected because of its relatively large effect size), and psychoacoustic roughness were made every 10 ms during vowels. All measurements were made in Matlab with F0, and spectral tilt analyzed via VoiceSauce (Shue, 2010) and roughness via the "Creakiness by roughness" package (Villegas et al., 2020). Measurements of F0, spectral tilt and roughness were made only during vowels, excluding onset and coda segments. Data was then analyzed as time series via R using ggplot2, with results plotted using the "gam" method. F0 measurements were normalized relative to each speaker's overall median F0 and was converted to cents to allow comparisons across speakers. For coda effects, normalized time was computed (time divided by vowel duration) to set the vowel-coda boundary as the reference point and to compensate for different vowel durations across tokens and speakers. Analyses were performed separately for syllables with different phonological vowel lengths and tones. All effects presented here were similar in both long and short vowels, but only results with short vowels are reported for ease of exposition. In cases where tone didn't affect an acoustic measurement, results are presented across all tones; otherwise, results are separated by tone.

3. Results

3.1. Coda Effects

As shown in Figure 1, spectral tilt was lowered preceding glottal stop codas, consistent with laryngeal constriction. However, while spectral tilt was not affected preceding obstruent codas, it was lowered preceding glide codas. The former observation is unexpected if obstruent codas are laryngealized (Morén & Zsiga, 2006); the latter observation is expected due to the high tongue position of the [j] and [w] glides, which raises A1 (Lotto, Holt & Kluender, 1997).

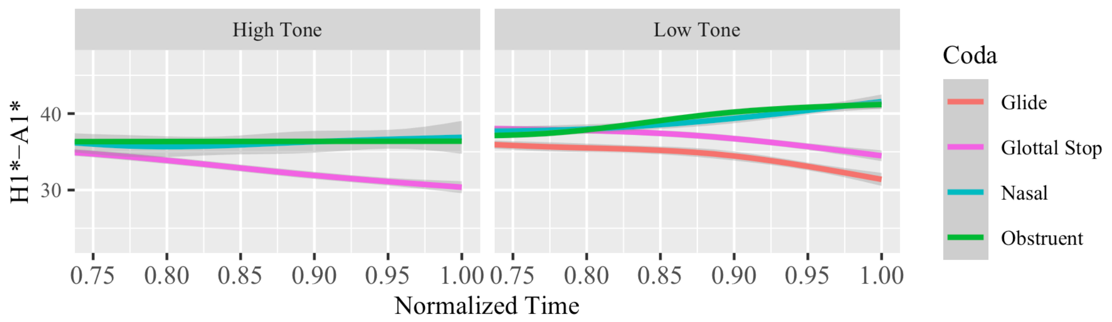

Figure 1: H1*–A1* in the final 25% of the vowel across all speakers.

Coda effects on F0 were most noticeable following glottal stops, and differed by speaker: Five speakers raised F0, eight speakers lowered F0, and for six speakers, there was no effect of coda on F0. Figure 2 shows these effects on the five speakers with raised F0 (left) and for the eight speakers with lowered F0 (right). F0 was lowered preceding obstruent codas in low tone, but to a lesser extent than glottal stops. In high tone, all speakers raised F0 preceding obstruent codas.

In general, F0 depends on the mode of laryngeal constriction: While creaky phonation lowers F0, tense phonation raises F0. Garellek (2019) noted that tense phonation is more commonly seen in high tones cross-linguistically. The fact that F0-raising was seen only in high tone here suggests that the laryngeal gesture associated with the consonant depends on the tone: Speakers produced laryngeal constriction via tense phonation with high tone, and creaky phonation with low tone.

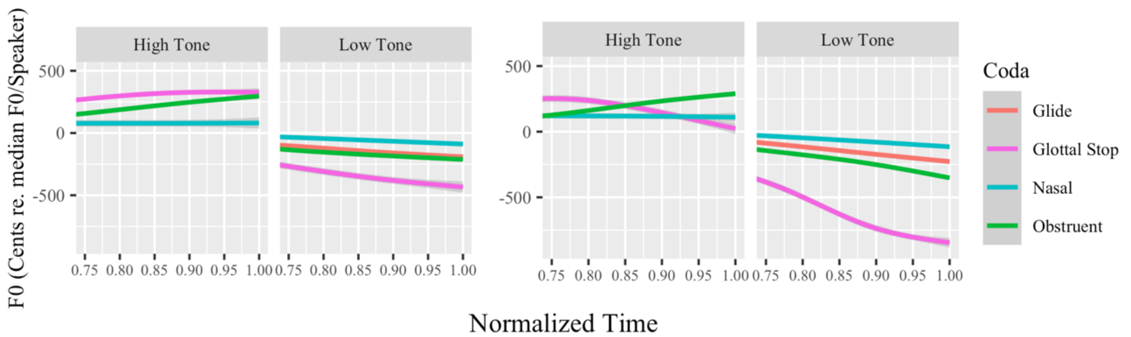

Figure 2: F0 in the final 25% of the vowel in speakers who raised F0 (left) and who lowered F0 (right).

Finally, psychoacoustic roughness was increased significantly at the end of a vowel preceding a glottal stop coda, and to a lesser extent preceding an obstruent coda. This is illustrated in Figure 3. Notably, while spectral tilt was not affected at all prior to obstruent codas, psychoacoustic roughness is increased. This suggests that obstruent codas are laryngealized, but that spectral tilt is not sensitive enough to pick this up.

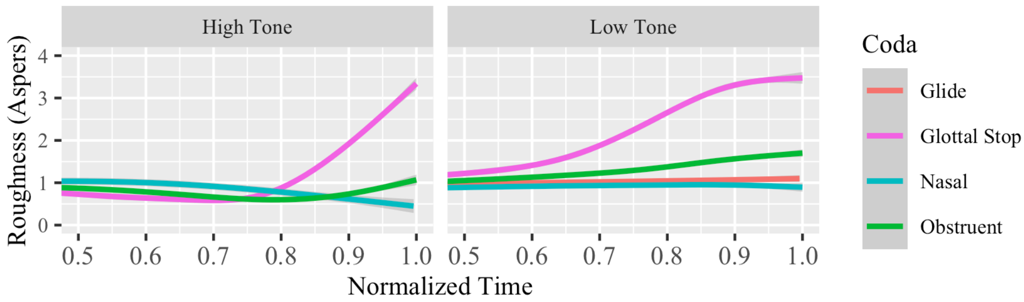

Figure 3: Psychoacoustic roughness in the final 25% of the vowel across all speakers.

3.2. Onset Effects

If unaspirated and voiced stops in onset are [+CG], then these onset consonants, along with glottal stops, should lower spectral tilt and F0 and raise roughness at the start of the following vowel. However, there was no effect on spectral tilt following glottal stops or unaspirated stops (see Figure 4). Spectral tilt was lowered significantly following voiced stops, but this effect was also present in nasals, and is a result of voicing, rather than laryngealization. Finally, spectral tilt was raised significantly following voiceless fricative and aspirated stop onsets. This effect is expected since both consonants are accompanied with breathiness, which is known to raise spectral tilt. These effects lasted until approximately 30-40 ms into the vowel.

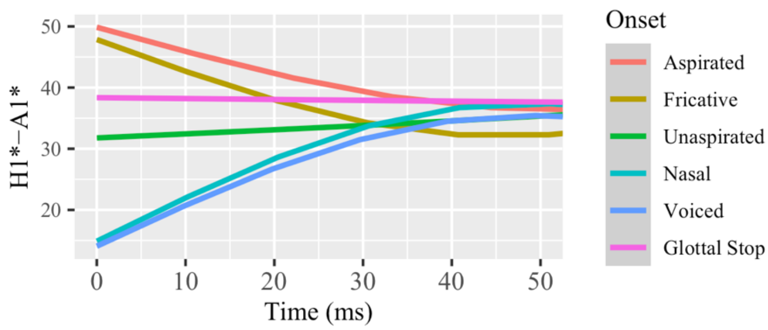

Figure 4: H1*–A1* in the initial 50 ms of the vowel across all speakers.

The results for F0 are shown in Figure 5. Glottal stop onsets significantly lower F0, but, in contrast, unaspirated stops raise F0. This suggests that unaspirated stops are not laryngealized. Like spectral tilt, F0 is also raised following voiceless fricatives and aspirated stops but is lowered following voiced stops and nasals. These effects are expected since breathiness raises F0 and voicing lowers F0. Rising and high tone do not occur in Thai words following unaspirated, voiced, and glottal stop onset consonants, and so they are not shown in Figure 5.

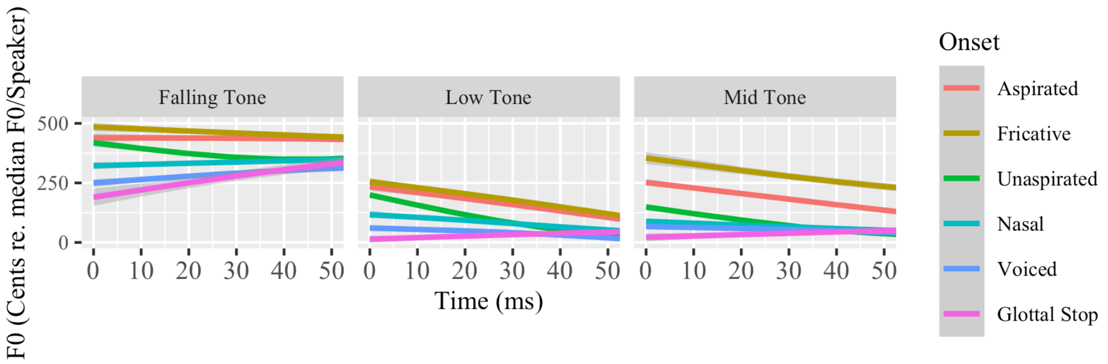

Figure 5: F0 in the initial 50 ms of the vowel across all speakers.

Finally, as seen in Figure 6, psychoacoustic roughness was higher following glottal stop onsets, but only two of the twenty speakers saw increased roughness following unaspirated onsets; for the remaining 18 speakers, there was no increase in roughness. Notably, roughness was not increased (and was in fact slightly decreased) following voiceless fricatives and unaspirated stops, both of which induce breathiness at the start of the following vowel. This indicates that psychoacoustic roughness yields a distinction between laryngealized phonations and breathy phonation. Roughness doesn’t seem to depend on the mode of laryngeal constriction, as it was uniformly increased by both tense and creaky phonation.

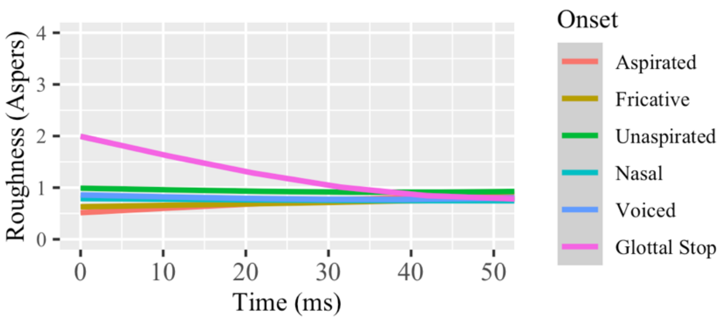

Figure 6: Psychoacoustic roughness in the initial 50 ms of the vowel across all speakers.

4. Conclusion

A production experiment showed that psychoacoustic roughness is a reliable measure of laryngeal constriction in consonants. Compared to spectral tilt, it yielded larger effect sizes and was able to identify laryngealization accompanying obstruent codas in Thai. In addition, psychoacoustic roughness successfully distinguished phonations with laryngeal constriction from breathy phonation. In fact, both tense and creaky phonations correlated with similarly increased roughness measurements. As such, roughness may correlate well with the phonological feature [+CG], since it does not distinguish between finer phonetic differences such as those of tense and creaky phonation. This behavior is not surprising since psychoacoustic roughness is defined based on the perception and not the production end of speech.

The results also contribute to the understanding of Thai featural phonology. Unaspirated stop and voiced stop onset consonants are not laryngealized in Thai, counter to the phonological analysis of Perkins (2011), where both voiced and unaspirated stops were posited as [+CG]. On the other hand, these findings are consistent with the analysis of Lee (2008), where the [–spread glottis] feature is used to account for gaps where high and rising tone are not found following voiced and unaspirated stop onsets.

References

Bradshaw, M. (1998) A Cross-Linguistic Study of Consonant-Tone Interaction. Ph.D dissertation, Ohio State University.

Esling, J., & Harris, J. (2005) “States of the glottis: An articulatory phonetic model based on laryngoscopic observations.” In W. Hardcastle, and J. Beck (eds.), A figure of speech: A Festschrift for John Laver (pp. 347-383). Lawrence Erlbaum Associates.

Garellek, M. (2019) “The phonetics of voice.” In The Routledge Handbook of Phonetics (pp. 75-106). Routledge.

Lee, S. J. (2008) “Consonant-tone interaction in optimality theory”, PhD dissertation, Rutgers University.

Lotto, A., Holt, L., & Kluender, K. (1997) “Effect of voice quality on perceived height of English vowels”, Phonetica 54, 76-93.

Morén, B., & Zsiga, E. (2006) “The lexcial and post-lexical phonology of Thai tones”, Natural Language & Linguistic Theory 24, 113-178.

Perkins, J. (2011) “Consonant-tone interaction in Thai”, PhD dissertation, Rutgers University.

Ruangjaroon, S. (2006) “Consonant-tone interaction in Thai: An OT analysis”, Taiwan Journal of Linguistics 4, 1-66.

Shue, Y. (2010) “The voice source in speech production: Data, analysis and models”, PhD dissertation, UCLA.

Villegas, J., Markov, K., Perkins, J., & Lee, S. J. (2020) “Prediction of creaky speech by recurrent neural networks using psychoacoustic roughness”, IEEE J. of Selected Topics in Signal Processing, 14(2), 355-366.