Consonant-Tone Interaction and Laryngealization in Thai

Jeremy Perkins
Rutgers University

1. Introduction

While a number of previous studies have reported laryngealization of Thai voiced and voiceless unaspirated stops, none have provided instrumental evidence of laryngealization. This paper provides acoustic evidence that Thai voiceless unaspirated and voiced stops are laryngealized via F0, spectral tilt and jitter measurements.

Many authors have made impressionistic claims that voiceless unaspirated stops in Thai are accompanied by a secondary laryngeal or pharyngeal articulation. Abramson (1962: 4) notes that "pre-vocalic /p t k/ are pharyngealized". Harris (1972: 11) labels the unaspirated series as glottalized. He describes them as "pronounced with simultaneous oral and glottal closures... so that the glottal release is not heard". Gandour & Maddieson (1976: 244) note that voiceless unaspirated /p/ is "often described as accompanied by glottal constriction". The latter authors conclude that the voiceless unaspirated series are tense stops and not ejectives.

Voiced stops have also been reported to be glottalized by Harris (1972: 14), who noted that "utterance initial voiced stops and approximants are usually preceded by weak glottal closures." Harris adds that even though there is some glottalization, these voiced stops are not produced as implosives. Ladefoged & Maddieson (1996: 55) describe voiced stops in Thai as occurring with "stiff, or even creaky voice". They add (p. 78) that voiced stops in Thai "are often accompanied by downward movement of the larynx that make them slightly implosive". None of the above authors presented any instrumental evidence of laryngeal activity however, instead relying on impressionistic claims.

Phonological facts also suggest that voiced and voiceless unaspirated stops are laryngealized. Thai has five contrastive tones (Abramson, 1962): High[1], low, mid, rising and falling. Stops and affricates can be voiced, voiceless unaspirated or voiceless aspirated. Tone occurs freely with one exception: High tone is not found in syllables with voiced or voiceless unaspirated onsets as shown in Table 1 below.[2] The only exceptions to this restriction occur in loans and onomatopoeia. In addition, [ʔ] and [h] do not co-occur with high tone.

---

[1] High tone is actually produced with a rising contour, starting at about the level of mid tone.

[2] This generalization holds of onset clusters as well. Thai allows only clusters of two consonants in onset position, the second of which is always l, r, or w (Abramson, 1962; Tumtavitikul, 1993).

Table 1
Onset-tone restrictions in Thai (examples from Ruangjaroon (2006) and Slayden (2009)). Shaded cells indicate combinations of onset and tone that are unattested in Thai native words.[3]

Consonant type	Mid tone	Low tone	Falling tone	Rising tone	High tone
Voiceless aspirated	pʰa: "take"	pʰàː "cut"	pʰâː "clothes"	pʰă: "a cliff"	pʰá: "knife"
Voiceless unaspirated	pa: "throw"	pà: "forest"	pâ: "aunt"	pě: "limp"	pá: "father" (Loan)
Voiced	baj "a leaf"	bà: "shoulder"	bâ: "crazy"	bŏ: “hollow”	None
Sonorant	ma: "come"	màj "new"	mâ:j "widow"	mă: "dog"	máj (question particle)
ʔ	ʔa: "aunt"	ʔà:w "a bay"	ʔâ: "spread"	ʔɔ̆: “yes”	None
h	haː “fun”	hàː “cholera”	hâː “five”	hăː “look for”	háː “ha!” (Loan)

Voiced stops lower F0 for about the first 50 ms in a following vowel in Thai (Gandour, 1974), suggesting a phonetic explanation for the phonological ban on high tone following voiced stops. The effect where voiced stops lower F0 has been documented widely even in non-tonal languages such as English (Hombert et al, 1979). Phonological accounts have even utilized a single feature for low tone and voicing (Bradshaw, 1999; Halle & Stevens, 1971). Halle & Stevens' system uses the features [stiff vocal cords] and [slack vocal cords] to refer to the vertical tension in the vocal cords. Increased stiffness raises F0 and inhibits voicing, while increased slackness lowers F0 and allows voicing to occur more easily. The stiff/slack distinction simultaneously explains the high/low tone distinction on vowels as well as the voiced/voiceless distinction in consonants, thus offering a possible explanation for the correlation between voicing and low tone and voicelessness and high tone, cross-linguistically. Thus, the phonological ban on voiced stops with high tone in Thai is motivated with or without glottal constriction.

However, glottal constriction, itself, can also affect F0 on a following vowel (Stevens, 1977; Tang, 2008). Unlike voicing however, glottal constriction can either raise or lower F0. Past studies on interaction between voiceless stops and F0 in Thai have yielded results that diverge in both directions. Erickson (1975) found that for eight of eleven native Thai speakers, F0 was raised following voiceless aspirated stops relative to voiceless unaspirated stops. However, the remaining three speakers showed the opposite pattern, with F0 raised following voiceless *unaspirated* stops. Gandour (1974) reported that voiceless aspirated stops also lower tone.

Halle & Stevens' model is designed to allow phonation to vary independently from tone across languages. However, phonation type and F0 do interact at the phonetic level as well (Stevens, 1977; Gordon & Ladefoged, 2001). For example, non-modal phonation (both creaky and breathy) is most commonly associated with lowered F0. While this is the case generally, there are reports of glottalization raising F0 (Maddieson, 1977; Hombert et al, 1979; Kingston, 2005), indicating that the two are phonetically

---

[3] The high tone restriction holds only of clusters whose first member is voiced or voiceless unaspirated. Therefore, the second member of the cluster has no effect on the tonal restriction.

independent, at least in part. Likewise, on the phonological side, Lee (2008) notes that the feature [+constricted glottis] in a preceding consonant can neutralize high tone to low tone in Burmese, but also it neutralizes low tone to high tone in Mulao. Downing and Gick (2001) presented evidence of two sets of aspirated stops in Botswana Kalang'a and two similar sets of fricatives in Nambya, one of which acted as a tone depressor, while the other did not, suggesting that spread glottis can also have two different effects on F0.

Despite this variation in the effects on F0, in the case of Thai, there are two observations to take into account that suggest a single hypothesis. The first are the impressionistic claims that voiced and voiceless unaspirated stops are laryngealized and sometimes creaky. The second is the phonological ban on high tone following these same consonants. If these two observations are related, then hypothetically voiceless unaspirated and voiced stops are laryngealized in Thai resulting in *lowered* F0 and creakiness at the onset of a following vowel. This paper describes a phonetic experiment that finds evidence of creakiness at the onset of vowels following voiced and voiceless unaspirated stops, that indicates that these consonants are laryngealized.

Measurements of F0 as well as jitter and spectral tilt in the onset of a vowel following the consonant are used to confirm laryngealization (Gordon & Ladefoged, 2001). Jitter is a measure of the degree of variation in the glottal period. When the glottis is constricted, the glottal pulses are less regular and so jitter is higher than in modal voice. Spectral tilt refers to the difference in amplitude between higher formants and one of the harmonics of F0. In modal phonation, there is a relatively larger amount of energy in F0 and relatively less energy in the higher formants. When the glottis is constricted, however, the higher formants gain energy relative to F0, resulting in lower spectral tilt. When the glottis is spread, on the other hand, the higher formants have considerably less energy, resulting in higher spectral tilt. Laryngealization is thus confirmed by higher jitter, lower spectral tilt, and lower F0 at the onset of a following vowel.

2. Methods

2.1 Experimental Design

A list of stimuli to be read by native Thai speakers and recorded for analysis was constructed as follows. Since creakiness is hypothetically associated with voiceless unaspirated and voiced obstruent onsets but not with voiceless aspirated onsets, the stimuli included words that differed only in which of these three onsets they contained. In addition to the oral stop series, [ʔ] and [h] were included in the study since these sounds also involve laryngeal articulation. Cross-linguistically, [ʔ] can be realized either with creaky phonation on an adjacent vowel or as a complete closure without any creakiness. (Ladefoged & Maddieson, 1996: 75). The fact that [ʔ] participates in the high tone restriction in Thai suggests that it might be realized as creaky phonation, thus lowering pitch in a following vowel as well. In this case, it should pattern with voiceless unaspirated and voiced onsets then. [h] is articulated with spread glottis. This configuration can result in aspiration or breathiness (Halle & Stevens, 1971), the latter of which usually lowers pitch (Laver, 1994: 477-8; Gordon & Ladefoged, 2001). Note that in Thai, aspirated onsets *can* occur with high tone but that [h] cannot. If the phonological tone restriction is based only on details of phonetic effects on F0, then [h] should be breathy, but the aspirated series should not be.

Table 2 below summarizes expectations for each onset class based on the discussion above.

Table 2
Phonetic hypotheses for each onset type[4]

Onset Type	Occurs With H Tone?	Phonetic Hypotheses
Voiceless unaspirated stops & affricates [p t k t͡ʃ]	No	Lowered F0, creaky phonation
Voiceless aspirated stops & affricates [pʰ tʰ kʰ t͡ʃʰ]	Yes	No effect on F0, modal phonation
Voiced stops [b d][5]	No	Lowered F0, creaky phonation
Nasal stops [m n ŋ]	Yes	No effect on F0, modal phonation
Glottal stop [ʔ]	No	Lowered F0, creaky phonation
Glottal fricative [h]	No	Lowered F0, breathy phonation

In building the experiment, paired comparisons were constructed by identifying near minimal pairs that differed only along a single laryngeal dimension. Each of the three oral stop series was paired with each other in one comparison, yielding the first three comparisons. Bilabial place is used here for illustrative purposes. While the [b] vs. [pʰ] comparison involves two laryngeal differences (voicing and aspiration), it was included since it directly addresses the hypothesis that [b] is laryngealized and [pʰ] is not. The [ʔ] vs. [h] comparison is included since it is the only one that removes oral place as a factor. It is not clear which of a creaky [ʔ] or a breathy [h] will lower F0 to a greater degree and so no prediction can be made for F0, however clearly [ʔ] should be creakier than [h]. The [ʔ] vs. [p] comparison is included to test the hypothesis that [p] is laryngealized, using [ʔ] as a baseline, since it surely *is* laryngealized. The [h] vs. [pʰ] comparison is included to test the hypothesis that [h] is breathy, while [pʰ] is not. Finally, the [pʰ] vs. [m] comparison is included as a baseline to confirm that [pʰ] (like [m]) is not laryngealized. These seven pairwise comparisons are summarized in table 3 below, along with the specific hypotheses for each concerning phonation and F0 differences induced at the onset of a following vowel.

---

[4] Other sounds that occur in Thai onsets include [l r j w f s]. These onsets were not considered in this study.

[5] There is no voiced velar stop [g] in Thai.

Table 3
Pairwise comparisons for onsets

Comparison	Hypothesis
[p] vs. [pʰ]	[p] creakier & lower F0
[p] vs. [b]	Same creakiness & F0
[b] vs. [pʰ]	[b] creakier & lower F0
[ʔ] vs. [h]	[ʔ] creakier but [h] breathier; unclear F0 prediction
[ʔ] vs. [p]	Same creakiness & F0
[h] vs. [pʰ]	[h] breathier & lower F0
[pʰ] vs. [m]	Same creakiness & F0

A single Thai frame sentence was used with the experimental word stimuli inserted. Morén & Zsiga (2004) used this frame sentence in their study of Thai coda-tone interaction. An example is given in (1), with the stimulus word underlined.

(1) Experimental sentences
nít bɔ̀k na: <u>pʰa:</u> kʰɯ: kamtɔ̀p
Nit tell Naa <u>take</u> be answer
"Nit told Naa that "<u>take</u>" was the answer"

These sentences place the stimuli words in stressed positions. The words both preceding and following the stimuli word were chosen with mid tone because mid tone has no coarticulatory effect on the tone of adjacent syllables (Morén & Zsiga, 2004).

Stimuli monosyllables were constructed varying onset type, which is the main independent variable of interest in this study. Voiced stops, voiceless unaspirated stops and voiceless aspirated stops were used as onsets. Additionally, nasal stops were also included as a non-glottalized baseline. Finally, [ʔ] and [h] were included since they are glottal. Bilabial onsets were used in building words because these have the smallest amount of coarticulation on the following vowel. Additionally, the Thai lexicon has a large number of words with bilabial onsets. Tone was also controlled for. For each of the six onsets, a vowel with each of the five tones was used in composing stimuli words, resulting in a total of thirty stimulus categories. Monosyllables were built using the long low vowel [aː], thus controlling for vowel quality. Codas were not used since they affect the pitch on the preceding vowel.

Not all of the thirty stimulus categories correspond to a Thai word. In cases where no Thai word existed, the nonce stimulus was still used, but in addition, Thai words were selected that differed only in that they contained a glide coda ([w] or [j]). In a few cases, it was necessary to use a short vowel with glide coda due to lexical restrictions. These Thai words were included in case the speakers had trouble producing the CV: nonce word versions, and were to be included in the analysis only in that case. Further, these Thai words with codas would only be included after a statistical test showed that the coda did not affect the pitch or creakiness at the onset of the vowel. The introduction of a glide coda or shortened vowel is not ideal but it is the best compromise that can be made in choosing words from the lexicon. In one case (pʰáː), a Thai word existed that contained an optional [r] trill following the [pʰ] onset. This word was included as part of the back-up stimuli (rather than a similar word with glide coda) since there was no word pronounced as [pʰá:] in Thai. However, there was a possibility that the speakers would pronounce the [r] and that this would affect the onset of the following vowel. Tokens with [pʰr] clusters would only be included in the analysis if statistical tests showed that [r] has no effect on the creakiness of the following vowel and if a larger number of errors were made on the nonce [pʰá:] syllable. A chart showing the stimuli words is given in table 4 below.

Table 4
Experimental word stimuli

Test Stimuli	Mid	Low	Falling	High	Rising
Aspirated	pʰa: "take"	pʰà: "cut"	pʰâ: "clothes"	pʰá: (nonce)	pʰǎ: "a cliff"
Unaspirated	pa: "throw"	pà: "forest"	pâ: "aunt"	pá: "father" (loan)	pǎ: "father" (loan)
Voiced	ba: “bar” (loan)	bà: "shoulder"	bâ: "crazy"	bá: (nonce)	bǎ: (nonce)
Sonorant	ma: "come"	mà: (nonce)	mâ: "grandma" (loan)	má: "mother" (loan)	mǎ: "dog"
ʔ	ʔa: "aunt"	ʔà: (nonce)	ʔâ: "spread"	ʔá: (nonce)	ʔǎ: (nonce)
h	ha: "fun"	hà: “cholera”	hâ: “five”	há: “ha!” (loan)	hǎ: “look for”

Back-ups	Mid	Low	Falling	High	Rising
Aspirated	N/A	N/A	N/A	pʰ(r)á: “knife”	N/A
Unaspirated	N/A	N/A	N/A	N/A	N/A
Voiced	baj "a leaf"	N/A	N/A	none	none
Sonorant	N/A	màj "new"	mâ:j "widow"	máj (question particle)	N/A
ʔ	N/A	ʔà:w "a bay"	N/A	none	none

The thirty CV: stimuli and the six back-up stimuli composed the complete experimental stimuli. Distractor stimuli were added that met the following conditions: First, they did not contain any of the onsets used in the experimental stimuli listed above. Second, they did not contain the low vowel [a]. Codas were allowed. They were all Thai monosyllablic words, randomly selected from Slayden’s (2009) online Thai dictionary. Twenty-six distractors were included, yielding a total of sixty-two token sentences. The sixty-two stimuli words were translated into Thai script, as was the host sentence. The stimulus word was separated from the rest of the sentence by spaces, so as to allow for the intended reading of the sentence. Eight slideshow files were made, each with all sixty-two sentences in a unique random order.

2.2 Recording

Three male native speakers of Standard Thai (speaker C, speaker T, and speaker K) were recruited via social networking. All three grew up in Bangkok speaking Standard Thai as their native language. All have parents who also spoke Standard Thai. They all listed English as a second language that they are able to use proficiently but not at a native-speaking level; the author's impression was that C & T have a greater degree of fluency in speaking and listening to English whereas speaker K had a very low fluency in both speaking and listening in English. Neither speaker had any physical or cognitive language impairment, nor any illnesses that would have affected their speech at the time of recording. Speaker C is thirty years old and moved to Los Angeles at age twenty. He had visited Los Angeles many times prior to moving there as well. He now resides in New Brunswick, NJ. Speaker T is thirty-four years old and lived in Thailand until he moved to the United States at age thirteen. He has visited Thailand three times for periods of about two weeks since then. He now resides in New Brunswick, NJ. Speaker K is thirty-nine years old and has lived in Nakhon Pathom and Nonthaburi, both on the outskirts of Bangkok. He has spent almost his entire life in Thailand and has only visited New Brunswick, NJ on two separate occasions for a total of three months.

Each speaker participated in a single recording session in the sound-attenuated booth with the door closed at the Rutgers University Phonetics Lab in New Brunswick, New Jersey. There were no sound sources within the booth. The speakers read all the sentences off a computer screen inside the booth that displayed the Thai sentences. An Audio-Technica AT4040 microphone with pop filter was used. It was connected via an XLR cable to an Applied Research & Technology Tube MP amplifier. A second XLR cable connected the amplifier to an M-Audio Delta 1010 sound card; digitization used ASIO drivers. The sound was digitized on a custom-built PC running Windows XP. Audio files containing the stimuli were created at a sampling rate of 44100 Hz using GoldWave version 5.06. The files were resampled to 16000 Hz in order to prevent overloading of the signal (Ladefoged, 2003: 95) prior to analysis in Praat (Boersma & Weenink, 2005).

A short practice session was done in order for the speakers to get used to the sentences. The speakers were also instructed that they may not recognize some of the Thai words (the nonce words), but that they should pronounce them as accurately as they could. For speaker C, only four or five tokens of each test stimulus were recorded due to an error that halted recording in the middle of the fifth randomized run-through of the stimuli. For speaker T, eight tokens of each stimulus were recorded. For speaker K, six or seven tokens of each stimulus were recorded.

2.3 Measurement

Vowels in the test stimuli words were segmented via Praat and saved in a text grid file. The edges of segmented vowels were determined based on the appearance and disappearance of the F1 formant. Figure 1 shows a segmented spectrogram that illustrates a typical example of how vowels were segmented.

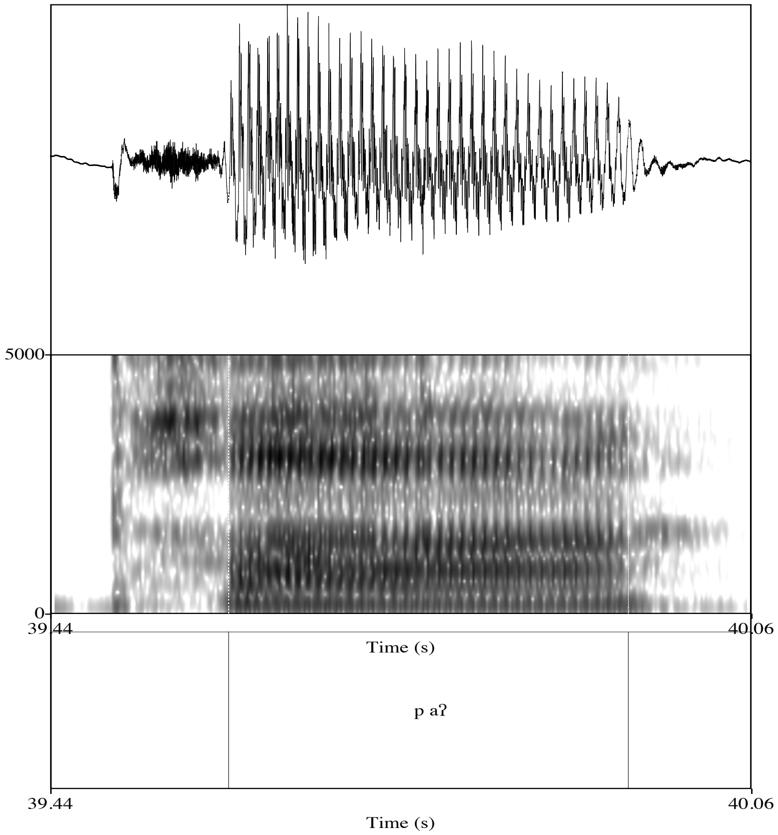

Fig. 1. Example of vowel segmentation via onset and offset of F1.

Creakiness of the vowel is determined via two measurements: Jitter and spectral tilt (Gordon & Ladefoged, 2001; Ladefoged, 2003: 169-181). Jitter(loc), calculated within Praat, is a measure of the variance of the time between successive glottal pulses. Higher jitter values indicate greater degrees of creakiness since creakiness induces more variation in the time between glottal pulses. Jitter is calculated over the first 70 ms of the following vowel since this is the smallest time interval that allowed the jitter calculation used in the software to work.

Spectral tilt is a second measure of creakiness. It is measured by taking the difference of the amplitude of either the first harmonic or the first formant with the amplitude of one of the higher formants. I follow Keating & Esposito (2007) in measuring H1–A1, the difference between the amplitude of the first harmonic of F0 (H1) and the amplitude of the first formant (A1). Creaky voice is typically produced such that the time that the glottis is open is less than for modal voice over a given glottal pulse period. Because of this, the amplitude of higher formants is relatively higher in creaky voice (Holmberg et al, 1995) and so the difference H1–A1 is close to zero typically. In order to measure spectral tilt, the vowels are broken into ten equal segments. For each of these segments, a long-term average spectrum is taken within Praat. The amplitudes of the first harmonic and first formants are then measured and their difference is calculated to get the spectral tilt for each segment of the vowel. Only the first of these ten segments is used in the data analysis, since this segment is closest to the consonant, and therefore is most likely to show effects of laryngealization due to the consonant.

F0 is measured via extraction of pitch values at 10 ms intervals over the entire length of the vowel. Gandour (1974) notes that consonant coarticulation with vowels in Thai occur over the first 50 ms of the vowel. However, only the first F0 measurement is used in the analysis, since it is closer to the onset consonant than the other measurements, and therefore most likely to include an effect, if present. All three of the above measurements are automated in a Praat script adapted from diCanio (2007). The measurements are automatically entered into a text file for further analysis.

Statistical analysis involved ANOVA's with speaker, tone and onset as independent variables. An ANOVA is a test used to determine whether a given independent variable (or combination of those variables) can explain the data significantly better than by chance. ANOVA's yield p-values, which are the probability that the partition of the data based on a given independent variable is explained by chance. A p-value smaller than a pre-determined significance level α, indicates a significant effect for a given independent variable then.

In this experiment, one ANOVA was used for each of the three dependent variables (jitter, spectral tilt, F0) to test for statistically significant effects due to onset type across speakers and tones. Significance level of α=0.05 was used. If a significant effect was found, then specific hypotheses were tested next. In this next layer of tests, filters were applied to the data to look for effects with a given two-way comparison from table 3 in mind. For example, in order to test for a significant effect between [p] and [p^h], the data were filtered to include only those tokens with [p] and [p^h] onsets and an ANOVA was performed on this subset of the data. Seven such comparisons were made as outlined in section 2.1. Bonferroni adjustments were made to account for the possibility of inflating the chance for Type-1 error by testing multiple hypotheses on the same data set. For example, since [p] is involved in three of the comparisons, the significance level α is adjusted to $\alpha/3=0.05/3= 0.0167$.

In cases where ANOVA's revealed significant interactions, two-tailed independent sample t-tests were conducted testing the specific hypotheses of the experiment. Whenever significant effects due to tone and/or speaker were discovered in the filtered ANOVA's, multiple t-tests were conducted over each speaker-tone group, of which there are fifteen in total (five tones * three speakers). Otherwise, if no significant effect was found for speaker or tone in an ANOVA, then those categories were ignored and t-tests were performed across speakers and/or tones.

2.4 Data Accuracy

Prior to statistical analysis, the recorded tokens were checked for accuracy. Of a total of 690 total tokens, 15% contained errors. 100 tone errors were discovered and 6 errors were discovered in the consonant productions. These tokens were excluded from the analysis since tone and onset are crucial factors in the experimental design. Table 5 illustrates that most of the errors were made on nonce words, with nearly as many made on loans, while only 4% of native Thai words were produced with errors. Table 6 reports errors as a function of speaker.

Table 5
Errors by Type

Type	# of Tokens	# of Errors	% Error
Nonce	134	57	43%
Loans	113	34	30%
Onomatopeia	18	0	0%
Native	425	15	4%

Table 6
Errors by Speaker

Speaker	# of Tokens	# of Errors	% Error
C	168	16	10%
T	288	57	20%
K	234	33	14%

These results indicate that the nonce words were not very effective at eliciting the intended tones. Two words accounted for 36 errors alone: The Thai interjection [háː] was consistently read with falling tone rather than high tone in all 20 tokens and so it was discarded completely. This finding seems too systematic to be an error, and so it is more likely that the Thai word used to elicit [háː], is pronounced with falling tone rather than high tone, at least for the three speakers in this study. While Ruangjaroon (2006) lists this word in her appendix without noting it as a loan word (she also transcribes it with high tone), one native Thai speaker[6] has informed me that it is a loan from the English interjection “ha!”. For this reason, it was classified as a loan word.

Likewise, the nonce word [pʰáː] was produced incorrectly in 16 of 18 tokens. Ten errors from speakers C and T were produced with falling tone. Another single error by speaker T was produced with high tone but with an [f] onset instead. The remaining five errors were made by speaker K, who inserted a liquid [l] following the initial [pʰ], yielding a [pʰl] cluster in all but one of his utterances. The two correct utterances were made by speakers T and K, and were retained. An additional three nonce words, two of which also contain high tone, also were produced with at least a 50% error rate. These words were [báː], [ʔáː] and [ʔàː]. In all cases, the most common error was for the tone to be produced as falling tone, although mid and rising tones were also produced. The large percentage of high-tone mispronunciation resulted in a very large number of high tone tokens being excluded, making statistical analysis within the high-tone category impossible in many cases. Notably, the three nonce words with the lowest error rate ([bǎː] 22%, [màː] 20% and [ʔǎː] 11%) also do not contain high tone.

Since the nonce words had such high error rates, the backup native Thai tokens were considered. A statistical test was conducted in order to test whether the presence of a glide coda had a significant effect on F0, jitter or spectral tilt. If no effect would be discovered, then the tokens [ʔàːw], [máj], [mâːj], and [màj] would replace their nonce correspondents without codas in the analysis, each of which was produced with a greater number of errors. The token [baj] was included in the experimental stimuli, but was not

---

[6] This same Thai speaker judged it as high and not falling tone.

considered as a replacent to the English loan [baː], since the latter was produced without any errors.

An ANOVA using onset, tone, speaker and coda as independent variables was performed, with the result that, while the coda had no effect on jitter [F(2, 469) = 1.2495, p = n.s.] or F0 [F(2, 467) = 2.7307, p = n.s.], it did affect spectral tilt [F(2, 462) = 166, p < 0.001] and so the tokens with codas were not used. A second test was conducted after first removing the tokens with short vowels, in order to allow for the possibility that only the codas following a short vowel were responsible for the previous result. The second test again confirmed that codas affected spectral tilt [F(2, 419) = 56.8, p < 0.001]. Additionally, an effect was discovered on F0 this time [F(2, 421) = 3.92, p < 0.05], while again no effect was discovered on jitter [F(2, 423) = 1.36, p = n.s.]. This indicates that codas do have a significant effect on the creakiness at the onset of the preceding vowel and so the tokens with codas were not included in analysis.

One further test was conducted to test the effect of a stop-liquid cluster on the onset of a following vowel. This test was conducted since many more errors were made in producing [pʰáː] than in producing [pʰráː]. If the liquid were to have no effect on any of the dependent variables, then [pʰráː] could replace [pʰáː] in the data set. Otherwise, the two correctly produced tokens of [pʰáː] would be used instead.

An ANOVA was conducted using onset and speaker as independent variables and using only [pʰáː] and [pʰráː] tokens in the data set. Jitter [F(1, 13) = 5.86, p < 0.05] and spectral tilt [F(1, 13) = 13.75, p < 0.01] were affected significantly by the liquid, while F0 [F(1, 13) < 1, p = n.s.] was not. This indicates that liquids affect the creakiness at the onset of a following vowel and so the [pʰráː] tokens were not included in the analysis.

3. Results

3.1 Jitter

Jitter is affected by onset type [F(5, 391) = 11.80, p < 0.001]. This result is consistent with the possibility that voiced and voiceless unaspirated stops in Thai are laryngealized. Figure 2 plots the mean jitter measurements for all three speakers categorized by tone and onset type. Jitter is also affected by the tone of the vowel [F(4, 391) = 10.28, p < 0.001] and differs depending on the speaker [F(1, 312) = 131.8419, p < 0.001]. For example, speaker T's vowels all contained larger jitter measures than both speaker C's and speaker K's vowels.

ANOVA's were conducted on the filtered subsets of the data that included only pairs of onsets being compared, as described in table 3. Table 7 summarizes the ANOVA results.

Table 7

ANOVA Results for Jitter with Onset Type

Comparison	ANOVA Results
[p] vs. [pʰ]	F(1, 141) = 1.92, p = n.s.
[p] vs. [b]	F(1, 141) = 2.68, p = n.s.
[b] vs. [pʰ]	F(1, 128) < 1, p = n.s.
[ʔ] vs. [h]	F(1, 119) = 7.99, p < 0.01
[ʔ] vs. [p]	F(1, 137) = 5.40, p = 0.022 (n.s.)
[h] vs. [pʰ]	F(1, 123) = 1.94, p = n.s.
[pʰ] vs. [m]	F(1, 131) = 37.28, p < 0.01

In two of the seven comparisons, onset type significantly affected jitter. The first difference, between [ʔ] and [h], was expected since [ʔ] should induce creakiness, while [h] should not. The second difference, between [pʰ] and [m], is unexpected since neither sound involves any creakiness. Figure 2 shows that [m] has consistently low jitter values, whereas all of the obstruents, including [pʰ] display occasional spikes where jitter is both high in magnitude and in variation, as can be seen from the larger 95% confidence intervals for the obstruent onsets. This result might indicate a difference between obstruents and sonorants more generally, rather than a difference between [pʰ] and [m].

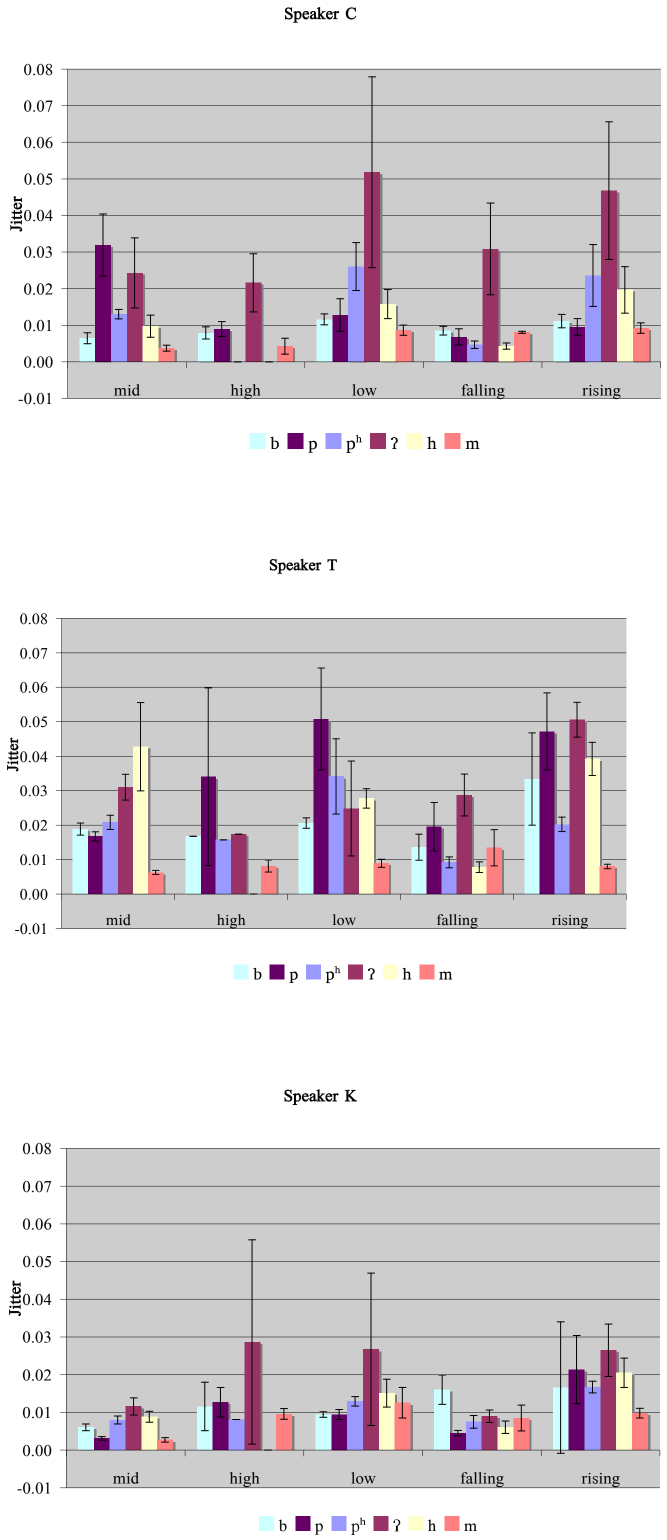

Fig. 2. Mean jitter measurements for speaker C (top), speaker T (middle) and speaker K (bottom) categorized by tone and onset. Error bars, here and throughout, indicate 95% confidence intervals.

There was no difference in jitter between [p] and [pʰ], nor between [b] and [pʰ], [pʰ] and [h], or [p] and [ʔ]. These results, coupled with the finding that [pʰ] has higher jitter than [m], suggest that all obstruents share similar jitter values, higher than those of sonorants. The [ʔ]-[p] comparison was nearly statistically significant (Bonferroni adjusted α was 0.017, while $p = 0.022$ was just above that level). [ʔ] yielded consistently higher and more variable jitter values than the other onset types (including [p]), suggesting that it induces creakiness slightly more often than other obstruents. The fact that an effect was detected between [ʔ] and [h] but not [ʔ] and [p] is consistent with the original hypothesis, since both [ʔ] and [p] are hypothetically creaky, while [h] is not. Next, t-tests were performed on the two comparisons where jitter *was* found to be significantly different. First, consider the [ʔ]-[h] comparison. Since tone [$F(4, 119) = 7.99$, $p < 0.01$] and speaker [$F(2, 119) = 9.13$, $p < 0.01$] also affected jitter, separate t-tests were performed for each tone and for each speaker. Table 8 summarizes the t-test results.

Table 8
T-test results for jitter comparisons between [ʔ] and [h]

Tone	Speaker C	Speaker T	Speaker K
Falling	$t(3.045) = 2.91$, p = n.s.	$t(7.895) = 2.62$, $p < 0.05$	$t(10) = 1.19$, p = n.s.
High	(not enough data)[7]	(not enough data)	(not enough data)
Mid	$t(4.505) = 1.64$, p = n.s.	$t(8.191) = -0.689$, p = n.s.	$t(2.324) = 0.900$, p = n.s.
Low	$t(3.139) = 1.88$, p = n.s.	$t(2.572) = -0.401$, p = n.s.	$t(2.454) = 1.10$, p = n.s.
Rising	$t(4.892) = 1.52$, p = n.s.	$t(11.996) = 1.41$, p = n.s.	$t(8.635) = 0.698$, p = n.s.

Of the fifteen t-tests, only one significant difference was discovered: Speaker T has significantly higher jitter following [ʔ] than [h] with falling tone. While not significant, figure 2 shows that jitter is also higher for rising tone following [ʔ] than [h] for speaker T. However, the situation is reversed for low and mid tones as speaker T has higher jitter following [h] than [ʔ] for these tones. When considering the other two speakers, however, jitter following [ʔ] is higher than [h] across all five tones. Thus while the t-tests do not provide conclusive proof, there is still some evidence that [ʔ] correlates with higher jitter values than [h], as expected if [ʔ] induces creakiness.

Next, consider the [pʰ]-[m] comparison. Again, both tone [$F(4, 131) = 6.46$, $p < 0.01$] and speaker [$F(2, 131) = 5.05$, $p < 0.01$] affected jitter and so t-tests were conducted for each tone and for each speaker, as before. Table 9 summarizes the t-test results.

---

[7] Here, and throughout, t-tests within the high tone category were usually impossible due to the small amount of data available with high tone. The high error rate in high-tone tokens made it necessary to remove a large number of high-tone tokens, leaving the sample size at 1 or 0 in many cases.

Table 9
T-test results for jitter comparisons between [pʰ] and [m]

Tone	Speaker C	Speaker T	Speaker K
Falling	t(3.487) = 4.45, p = 0.015 (n.s.)	t(6.314) = 0.721, p = n.s.	t(7.892) = 0.247, p = n.s.
High	(not enough data)	(not enough data)	(not enough data)
Mid	t(6.743) = –6.70, p < 0.001	t(8.045) = –5.35, p < 0.001	t(10.719) = 0.378, p = n.s.
Low	t(4.236) = –2.91, p = n.s.	t(5.112) = –2.20, p = n.s.	t(5.941) = 0.527, p = n.s.
Rising	t(4.307) = –1.85, p = n.s.	t(8.367) = –4.38, p < 0.01	t(10.665) = –3.13, p < 0.01

Of the fifteen t-tests, significant differences were discovered in mid tone for speakers C and T and in rising tone for speakers T and K. In all four of these comparisons, jitter is higher for [pʰ] than for [m]. In addition, even though the other comparisons didn't yield significant differences in jitter, all speakers have higher mean jitter in [pʰ] than in [m] with mid, low and rising tone. However, the situation is reversed for falling tone: All three speakers have lower mean jitter for [pʰ] than [m]. This interaction between onset and tone, while not significant in the t-tests for falling tone, was found to be significant in the ANOVA over the larger pool of data [F(4, 131) = 5.14, p < 0.001].

In conclusion, the jitter results did not indicate a difference in creakiness between any of [pʰ], [p], or [b]. Therefore, the jitter measurements do not support the main hypothesis that voiced and voiceless unaspirated stops are laryngealized while [pʰ] is not. The only significant difference discovered among obstruent comparisons was that [ʔ] had higher jitter than [h], as expected. Jitter was also slightly higher for [ʔ] than for [p], although not statistically so. Finally, [pʰ] has higher jitter than [m] in low, mid and rising tones, but the situation was reversed for falling tone where [m] has higher jitter than [pʰ].

3.2 Spectral Tilt

Spectral tilt is affected by onset type [F(5, 388) = 245.03, p< 0.001]. This result is consistent with the hypothesis that voiced and voiceless unaspirated stops are laryngealized. Significant effects were also discovered for tone [F(4, 388) = 109.7, p < 0.001], speaker [F(2, 388) = 13.10, p < 0.001], as well as interactions between onset and tone [F(19, 388) = 160.11, p < 0.001], onset and speaker [F(10, 388) = 4.47, p < 0.001], and tone and speaker [F(8, 388) = 3.59, p < 0.001].

ANOVA's were performed on seven separate subsets of the data consisting of only those onsets being directly compared as outlined in table 3. Table 10 illustrates the results of the ANOVA tests.

Table 10

ANOVA Results for Spectral Tilt with Onset Type

Comparison	ANOVA Results
[p] vs. [pʰ]	F(1, 140) = 258.01, p < 0.001
[p] vs. [b]	F(1, 139) = 14.39, p < 0.001
[b] vs. [pʰ]	F(1, 127) = 108.93, p < 0.001
[ʔ] vs. [h]	F(1, 118) = 86.17, p < 0.001
[ʔ] vs. [p]	F(1, 135) = 880.70, p < 0.001
[h] vs. [pʰ]	F(1, 123) = 28.12, p < 0.001
[pʰ] vs. [m]	F(1, 131) = 12.75, p < 0.001

In every comparison, spectral tilt measurements were significantly affected by onset. Spectral tilt measurements were not expected to differ in comparisons between [p] vs. [b], [ʔ] vs. [p], and [pʰ] vs. [m], since these pairs weren't thought to differ in the extent of creakiness or breathiness.

T-tests were conducted separately for each comparison to test for significant differences in the means for each onset type. First, consider the [p]-[pʰ] comparison. Spectral tilt was significantly different for each speaker [F(2, 140) = 6.87, $p < 0.01$]. Likewise, the effect of onsets on spectral tilt differed among the speakers [F(2, 140) = 11.44, $p < 0.001$]. Tone did not significantly affect spectral tilt [F(4, 140) = 0.85, p = n.s.] and so it was ignored as a factor in conducting t-tests. Table 11 summarizes the results for the t-tests for each speaker.

Table 11

T-test results for spectral tilt comparisons between [p] and [pʰ]

Speaker	T-Test Results
C	t(27.15) = –11.44, $p < 0.001$
T	t(59.79) = –8.31, $p < 0.001$
K	t(40.45) = –7.98, $p < 0.001$

For all three speakers a significant difference was found between spectral tilt measurements for [p] and [pʰ]. Figure 3 plots mean spectral tilt values for each onset across speakers. It is evident that the values for [pʰ] are higher than those for [p] for all three speakers. This is consistent with the original hypothesis that [p], but not [pʰ] should induce creakiness, resulting in lower spectral tilt, at the onset of a following vowel.

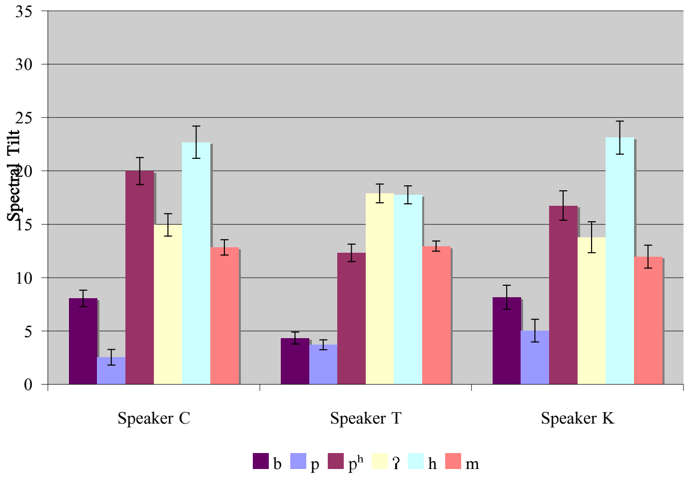

Fig. 3. Mean spectral tilt measurements for each speaker categorized by onset.

The [p]-[b] comparison is considered next. In addition to onset, an effect for speaker was also discovered in this comparison [F(2, 139) = 4.75, $p < 0.017$]. Three t-tests were performed then, one for each speaker, since tone [F(4, 139) = 0.23, p = n.s.] had no effect on spectral tilt. The results of the t-tests are summarized in table 12.

Table 12

T-test results for spectral tilt comparisons between [p] and [b]

Speaker	T-Test Results
C	t(38.52) = –4.91, $p < 0.001$
T	t(54.86) = –0.53, p = n.s.
K	t(40.11) = –2.51, $p < 0.017$

For two of the three speakers a significant difference was found between spectral tilt measurements for [p] and [b], a result that is unexpected. The original hypothesis that both [p] and [b] are laryngealized does not predict a difference in spectral tilt. The result indicates that the two are not laryngealized to the same degree. Figure 3 shows a lower spectral tilt for [p] indicating a higher degree of glottal constriction than [b] for speakers C and K.

Next, consider the [b]-[p^h] comparison. The three speakers differed in the degree of spectral tilt for each onset [F(2, 127) = 14.28, $p< 0.001$]. Tone had no effect on spectral tilt [F(4, 127) = 1.09, p = n.s.] and so it could be ignored in the t-tests. Again, three t-tests were conducted, one for each speaker, the results of which are summarized in table 13.

Table 13

T-test results for spectral tilt comparisons between [b] and [p^h]

Speaker	T-Test Results
C	t(31.37) = –7.38, $p < 0.001$
T	t(59.69) = –6.12, $p < 0.001$
K	t(49.22) = –5.08, $p < 0.001$

For all three speakers, a significant difference in spectral tilt was found. Figure 3 shows that the spectral tilt for [p^h] is higher than [b]. This supports the hypothesis that [b] is laryngealized, while [p^h] is not.

Next, consider the [ʔ]-[h] comparison. Significant effects on spectral tilt were also discovered for tone [F(4, 118) = 268.48, $p < 0.001$], speaker [F(2, 118) = 4.50, $p < 0.025$], the interaction between onset and tone [F(3, 118) = 423.59, $p < 0.001$], the interaction between onset and speaker [F(2, 118) = 6.66, $p < 0.001$], and the interaction between tone and speaker [F(8, 118) = 2.55, $p < 0.025$]. T-tests were conducted for each speaker and for each tone, yielding fifteen tests, whose results are summarized in table 14.

Table 14
T-test results for spectral tilt comparisons between [ʔ] and [h]

Tone	Speaker C	Speaker T	Speaker K
Falling	t(6.49) = –2.34, p = n.s.	t(12.47) = –1.05, p = n.s.	t(9.81) = –1.79, p = n.s.
High	(not enough data)	(not enough data)	(not enough data)
Mid	t(6.76) = –1.45, p = n.s.	t(12.00) = 1.17, p = n.s.	t(8.59) = –3.21, p < 0.025
Low	t(3.93) = 13.20, p < 0.001	t(2.11) = 8.50, p < 0.025	t(5.88) = 26.69, p < 0.001
Rising	t(7.19) = –2.88, p < 0.025	t(11.47) = 0.60, p = n.s.	t(6.95) = –2.36, p = n.s.

Figure 4 illustrates mean spectral tilt values for [ʔ] and [h] onsets, categorized by onset, tone, and speaker. Significant differences were found for all three speakers with low tone, with [h] having higher spectral tilt than [ʔ]. The remaining two significant findings were for speaker K, where [h] had higher spectral tilt than [ʔ] for mid tone, and for speaker C, where [h] had higher spectral tilt than [ʔ] for rising tone. While the other comparisons did not yield significant results, speakers C and K have higher spectral tilt for [h] than [ʔ] across all tones. Only speaker T had higher spectral tilt for [ʔ] than [h] (in mid, rising and falling tones), although these results were not found to be statistically significant. Thus, in nine out of twelve cases, [ʔ] has lower spectral tilt than [h]. The lower spectral tilt in [ʔ] is consistent with either increased creakiness in [ʔ] or increased breathiness in [h].

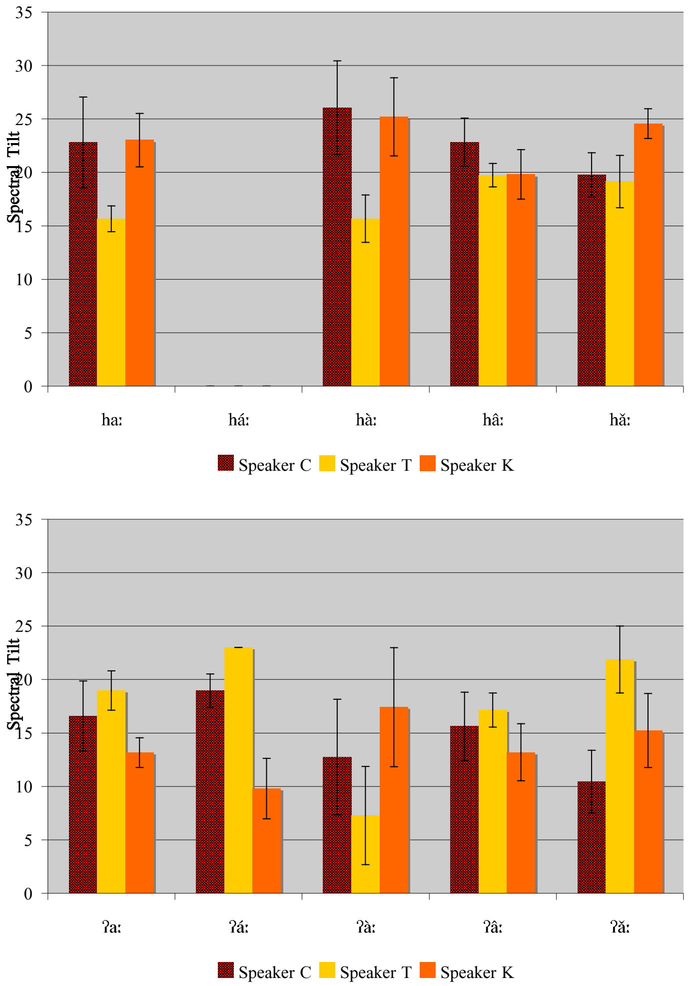

Fig. 4. Mean spectral tilt measurements for [h] (top) and [ʔ] onsets (bottom) for each speaker, categorized by tone.

Next, consider the [ʔ]-[p] comparison. Spectral tilt was affected by tone [F(4, 135) = 308.49, p < 0.001], by the interaction between onset and tone [F(4, 135) = 492.04, p < 0.001], by the interaction between onset and speaker [F(2, 135) = 4.84, p < 0.01], and by the interaction between onset, tone, and speaker [F(8, 135) = 2.64, p < 0.017]. However, no effect was detected for speaker [F(2, 135) = 1.93, p = n.s.]. Five t-tests were conducted then, one for each tone, the results of which are shown in table 15.

Table 15

T-test results for spectral tilt comparisons between [ʔ] and [p]

Tone	T-Test Results
Mid	t(31.66) = 7.67, p < 0.001
High	t(9.99) = 6.16, p < 0.001
Low	t(9.98) = 24.37, p < 0.001
Falling	t(7.12) = 29.69, p < 0.001
Rising	t(20.38) = 5.70, p < 0.001

Significant differences were discovered across all five tones. Figure 5 illustrates mean spectral tilt values for [ʔ] and [p] for each tone. [ʔ] has higher spectral tilt in all cases. For low tone, the difference is reduced: [ʔ] has a relatively more low spectral tilt and [p] has a relatively more high spectral tilt. This result does not support the hypothesis that both [ʔ] and [p] should induce creakiness in a following vowel. The high spectral tilt values for [ʔ] indicate that it is not inducing creakiness. This result is in contrast to the findings for jitter, which indicated that [ʔ] *does* in fact induce more creakiness than other obstruents. However, the jitter results were not statistically significant, suggesting that [p] induces creakiness in a following vowel, while [ʔ] does not.

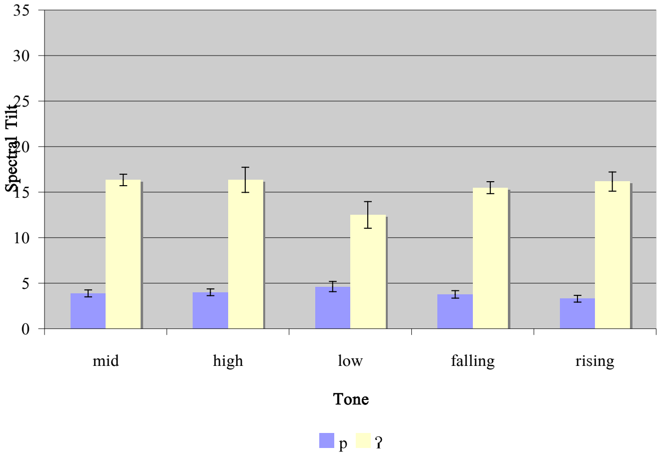

Fig. 5. Mean spectral tilt measurements for [p] and [ʔ] onsets across speakers, categorized by tone.

The next comparison is [h] versus [p^h]. In this case, an effect existed with speaker [F(2, 123) = 16.83, p < 0.001], but not with tone [F(4, 123) = 0.79, p = n.s.]. T-tests were conducted for each speaker then, the results of which are shown in table 16.

Table 16

T-test results for spectral tilt comparisons between [h] and [p^h]

Speaker	T-Test Results
C	t(34.90) = 1.47, p = n.s.
T	t(52.37) = 3.98, p < 0.001
K	t(50) = 3.57, p < 0.001

Figure 3 shows that [h] has higher spectral tilt than [p^h] for all three speakers but as table 16 shows, this is only significant for speakers T and K. These results suggest a greater degree of breathiness in [h] than in [p^h].

The final comparison is between [p^h] and [m]. In this case, an effect was discovered in ANOVA tests for speaker [F(2, 131) = 5.07, p < 0.01] and for the interaction between onset and speaker [F(2, 131) = 7.09, p < 0.01]. T-tests were conducted for each speaker since tone did not affect spectral tilt in this case [$F(4, 131) = 1.18$, p = n.s.]. The results are shown in table 17.

Table 17

T-test results for spectral tilt comparisons between [m] and [p^h]

Speaker	T-Test Results
C	$t(35.90) = 4.15$, $p < 0.001$
T	$t(60.26) = -0.51$, p = n.s.
K	$t(52.79) = 2.69$, $p < 0.01$

For speakers C and K, the spectral tilts were significantly different, with [p^h] having higher values than [m]. On the other hand, speaker T had slightly higher mean spectral tilt for [m], but this result was not statistically significant. This result is not consistent with the hypothesis that [p^h] and [m] are both non-laryngeal since a difference was detected. However, it is consistent with [p^h] being more breathy than [m], a possible result since aspiration involves spreading of the glottis, which could raise spectral tilt.

In conclusion, spectral tilt measurements were significantly different in all comparisons made. The comparisons between the oral stops found that [p] had the lowest spectral tilt, followed by [b], and then by [p^h], which had the highest spectral tilt. It was expected that both [p] and [b] would have lower spectral tilt than [p^h] under the hypothesis that the former two sounds are laryngealized, while the latter is not. The finding that [p] has lower spectral tilt than [b] indicates that [p] has a greater degree of glottal constriction than [b]. Comparisons with glottals found that [h] had higher spectral tilt than [ʔ], as expected since [h] is produced with spread glottis and [ʔ] with constricted glottis. Unexpectedly however, [ʔ] was found to have higher spectral tilt than [p], indicating a greater degree of constriction for [p] than [ʔ]. Finally, [h] was found to have higher spectral tilt than [p^h], which in turn had higher spectral tilt than [m]. These last three results indicate a hierarchy of breathiness: [h] is more breathy than [p^h], while [m] is not produced with any breathiness.

3.3 F0

With F0 as the dependent variable, a significant effect was found for onset type [$F(5, 389) = 58.49$, $p < 0.001$]. Additionally, tone affected F0 on a following vowel [$F(4, 389) = 213.85$, $p < 0.001$]. Speakers differed with respect to F0 at vowel onset as well [$F(2, 389) = 447.43$, $p < 0.001$]. Significant effects were also found for the interactions between onset and tone [$F(19, 389) = 1.84$, $p < 0.05$], onset and speaker [$F(10, 389) = 2.18$, $p < 0.05$], and tone and speaker [$F(8, 389) = 14.00$, $p < 0.001$].

The seven onset filter conditions, as summarized in table 3, were applied and ANOVA's were conducted in order to test for effects among the filtered sets of data. The results of these tests are given in table 18.

Table 18

ANOVA Results for F0 with Onset Type

Comparison	ANOVA Results
[p] vs. [pʰ]	F(1, 140) = 15.71, p < 0.001
[p] vs. [b]	F(1, 141) = 1.66, p = n.s.
[b] vs. [pʰ]	F(1, 127) = 30.57, p < 0.001
[ʔ] vs. [h]	F(1, 118) = 27.97, p < 0.001
[ʔ] vs. [p]	F(1, 136) = 89.72, p < 0.001
[h] vs. [pʰ]	F(1, 122) = 0.90, p = n.s.
[pʰ] vs. [m]	F(1, 130) = 86.26, p < 0.001

In five of the seven comparisons, onset type affected F0. Significant effects were discovered in both the [p]-[pʰ] and [b]-[pʰ] comparisons, but not in the [p]-[b] comparison. An effect was also discovered for the [ʔ]-[h] comparison, as expected. However, the [ʔ]-[p] comparison did yield a significant affect, counter to expectations. In the [h]-[pʰ] comparison, no significant effect was discovered, also counter to expectations. The final comparison, between [pʰ] and [m] yielded an effect, refuting the hypothesis that neither should affect F0.

T-tests were conducted for the five comparisons where significant effects due to onset type were discovered. First the [p]-[pʰ] comparison is considered. Tone [F(4, 140) = 96.90, p < 0.001] and speaker [F(2, 140) = 205.06, p < 0.001] both affected F0 along with the interaction between tone and speaker [F(8, 140) = 6.24, p < 0.001]. T-tests were conducted for each tone and for each speaker. The results are given in table 19.

Table 19

T-test results for F0 comparisons between [p] and [pʰ]

Tone	Speaker C	Speaker T	Speaker K
Falling	t(6.921) = –1.40, p = n.s.	t(13.99) = –3.19, p < 0.01	t(11.95) = –1.28, p = n.s.
High	(not enough data)	(not enough data)	(not enough data)
Mid	t(7.38) = –1.17, p = n.s.	t(8.70) = –1.00, p = n.s.	t(12.00) = 0.26, p = n.s.
Low	t(4.49) = –1.72, p = n.s.	t(5.84) = 0.22, p = n.s.	t(8.61) = –0.12, p = n.s.
Rising	t(5.96) = –2.31, p = n.s.	t(13.76) = –1.60, p = n.s.	t(10.82) = –0.60, p = n.s.

The only significant difference was discovered with falling tone for speaker T, where [pʰ] has higher F0 than [p], as can be seen in figure 6.

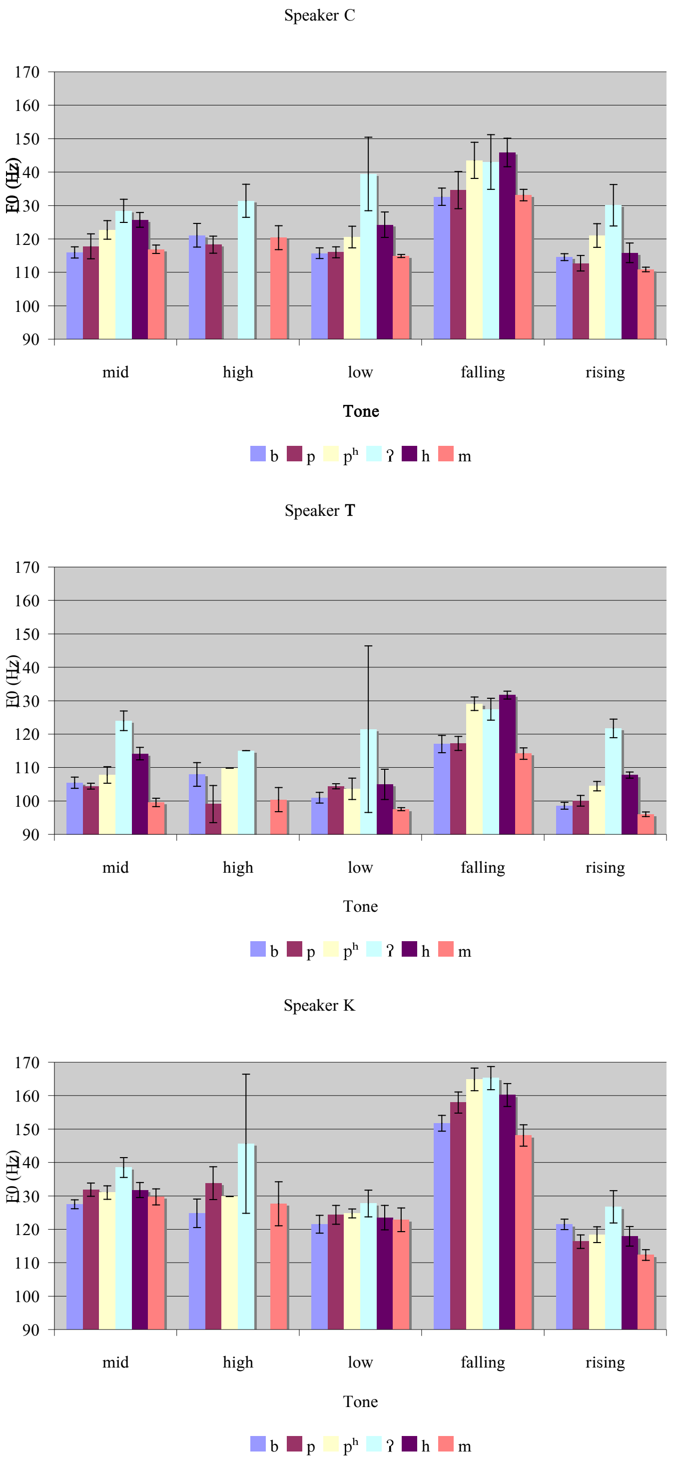

Fig. 6. Mean F0 measurements for speaker C (top), speaker T (middle) and speaker K (bottom), categorized by onset and tone.

This result is consistent with [p] being laryngealized and inducing creaky voice. However, since only one out of fifteen tests proved this result, this only amounts to weak evidence. A comparison of [pʰ] and [p] in figure 6 shows that F0 is higher for [pʰ] for speaker C in all tones. Speaker T shows the same result for all tones except low tone, where [pʰ] has a higher F0 than [p]. Speaker K shows less consistent results: While [pʰ] has higher F0 for falling and rising tone, the reverse is true for mid and high tone, and with low tone there is virtually no difference. The evidence that [p] is creakier than [pʰ] based on F0 is weak at best then.

Next, the [b]-[pʰ] comparison is considered. In addition to onset, significant effects were discovered in the ANOVA with tone [F(4, 127) = 115.77, p < 0.001], speaker [F(2, 127) = 205.26, p < 0.001], the interaction between onset and tone [F(4, 127) = 3.61, p < 0.01], and the interaction between tone and speaker [F(8, 127) = 5.08, p < 0.001]. T-tests were conducted for each tone and for each speaker. The results are reported in table 20.

Table 20
T-test results for F0 comparisons between [b] and [pʰ]

Tone	Speaker C	Speaker T	Speaker K
Falling	t(5.05) = –2.39, p = n.s.	t(13.10) = –2.83, p = 0.014 (n.s.)	t(10.76) = –2.73, p = 0.02 (n.s.)
High	(not enough data)	(not enough data)	(not enough data)
Mid	t(5.66) = –2.44, p = n.s.	t(12.79) = –0.59, p = n.s.	t(10.46) = –1.25, p = n.s.
Low	t(4.36) = –1.84, p = n.s.	t(5.52) = –0.77, p = n.s.	t(8.02) = –1.00, p = n.s.
Rising	t(4.43) = –1.99, p = n.s.	t(14.00) = –2.30, p = 0.037 (n.s.)	t(5.10) = 1.25, p = n.s.

None of the t-tests yielded a significant difference in F0 for [b] and [pʰ]. While the t-tests did not confirm a difference, inspection of the means in figure 6 shows that in all of the 12 categories, except one, [pʰ] has higher F0 than [b]. The lone exception is for speaker K, where [b] has higher F0 than [pʰ] for rising tone. While not significant, these results still suggest that [pʰ] has higher F0 than [b], as expected.

Next, consider the [ʔ]-[h] comparison. F0 was affected by tone [F(4, 119) = 5.07, p < 0.001] and speaker [F(2, 119) = 9.13, p < 0.001]. T-tests were conducted for each tone and for each speaker then. The results are given in table 21.

Table 21
T-test results for F0 comparisons between [ʔ] and [h]

Tone	Speaker C	Speaker T	Speaker K
Falling	t(5.36) = –0.41, p = n.s.	t(8.82) = –0.96, p = n.s.	t(10.00) = 1.00, p = n.s.
High	(not enough data)	(not enough data)	(not enough data)
Mid	t(5.86) = 0.77, p = n.s.	t(11.85) = –0.59, p = 0.049 (n.s.)	t(10.22) = 1.61, p = n.s.
Low	t(3.78) = 1.77, p = n.s.	t(2.43) = 1.26, p = n.s.	t(5.82) = 1.02, p = n.s.
Rising	t(5.70) = 2.31, p = n.s.	t(6.55) = 4.46, p < 0.01	t(9.08) = 1.45, p = n.s.

Only one significant difference was found: Speaker T has significantly higher F0 for [ʔ] than [h] for rising tone. In fact, means for F0 for [ʔ] are higher than [h] in all but two of the comparisons. The only exceptions are in falling tone for speakers C and T, where F0 is higher for [h]. Therefore, [ʔ] is associated with higher F0 than [h].

Next, consider the [ʔ]-[p] comparison. F0 was affected by tone [F(4, 136) = 31.44, p < 0.001] and speaker [F(2, 136) = 99.04, p < 0.001], as well as the interaction between tone and speaker [F(8, 136) = 5.84, p < 0.001]. T-tests were conducted for each speaker and tone then. The results are shown in table 22.

Table 22
T-test results for F0 comparisons between [ʔ] and [p]

Tone	Speaker C	Speaker T	Speaker K
Falling	t(6.38) = 1.08, p = n.s.	t(11.9) = 2.07, p = n.s.	t(10.94) = 1.42, p = n.s.
High	t(5.30) = 2.70, p = 0.04 (n.s.)	(not enough data)	t(3.03) = 1.03, p = n.s.
Mid	t(7.96) = 2.32, p = 0.049 (n.s.)	t(8.2) = 4.96, p < 0.01	t(10.45) = 1.60, p = n.s.
Low	t(3.16) = 2.95, p = n.s.	t(2.02) = 1.36, p = n.s.	t(8) = 0.80, p = n.s.
Rising	t(4.72) = 2.98, p = 0.03 (n.s.)	t(9.54) = 6.07, p < 0.001	t(7.15) = 1.87, p = n.s.

Two of the comparisons yielded significant differences between F0 for [ʔ] and [p]. Speaker T had higher F0 for [ʔ] than [p] for mid and rising tone. While not significant, the other means all showed the same pattern: [ʔ] had higher F0 than [p] in all cases. In general, the F0 measures were much higher for [ʔ] than would be expected if [ʔ] were to induce creaky voice in the onset of the following vowel. These results indicate that [ʔ] is not creaky, in accordance with the spectral tilt results.

The final comparison is between [pʰ] and [m]. In addition to onset, tone [F(4, 130) = 126.49, p < 0.001] and speaker [F(2, 130) = 257.97, p < 0.001] both affected F0. Additionally, the interaction between onset and tone [F(4, 130) = 3.83, p < 0.01], and the interaction between tone and speaker [F(8, 130) = 5.78, p < 0.001] also affected F0. T-tests were conducted for each tone and for each speaker then. The results are given in table 23.

Table 23
T-test results for F0 comparisons between [pʰ] and [m]

Tone	Speaker C	Speaker T	Speaker K
Falling	t(5.87) = 2.02, p = n.s.	t(11.57) = 4.74, p < 0.001	t(10.95) = 3.24, p < 0.01
High	(not enough data)	(not enough data)	(not enough data)
Mid	t(4.98) = 2.19, p =n.s.	t(10.39) = 2.32, p = 0.04 (n.s.)	t(10.72) = 0.38, p = n.s.
Low	t(5.90) = 1.81, p = n.s.	t(5.22) = 1.79, p = n.s.	t(5.94) = 0.53, p = n.s.
Rising	t(4.38) = 3.13, p = 0.03 (n.s.)	t(10.25) = 4.17, p < 0.01	t(9.72) = 1.98, p = n.s.

In three of the comparisons, a significant difference in F0 was discovered. Speakers T and K had higher F0 for [pʰ] than [m] in falling tone. Speaker T also had higher F0 for [pʰ] than [m] in rising tone. Although not statistically significant, the remaining comparisons all had higher F0 for [pʰ] than [m]. Inspection of figure 6 shows that [m] has lower mean F0 values than even [b] and [p] in some cases, suggesting that [m] is also lowering F0.

In conclusion, F0 was lowered in [b] and [p] relative to [pʰ]. This result is in accordance with the result for spectral tilt: [b] and [p] induce creaky phonation while [pʰ] does not. Counter to expectations, [ʔ] was found to *raise* pitch, suggesting it is not inducing creaky phonation in a following vowel. Finally, [m], like [b] and [p] lowers F0.

3.4 Conclusion

In summary, F0 and spectral tilt measurements both suggested that [b] and [p], but not [pʰ] are laryngealized, inducing creaky voice on a following vowel. Jitter measurements, on the other hand, were not significantly different among any of the oral stops. Jitter was higher for [ʔ] than [h] and for [pʰ] than [m], but otherwise did not distinguish between other onsets. F0 and spectral tilt results for [ʔ] suggested that it is not laryngealized in the same manner as [b] and [p] since [ʔ] induces very high spectral tilt and F0 at the onset of a following vowel. [h] yielded very high spectral tilt as well, indicating breathiness, but F0 was not lowered significantly. Finally, [m] lowered F0, but did not have a significantly lowered spectral tilt, and therefore it is not laryngealized.

4. Discussion and Concluding Remarks

Lowered spectral tilt and F0 both suggest that [b] and [p] are laryngealized, inducing creaky voice at the onset of a following vowel. Jitter did not distinguish among the oral stops though. Instead, [ʔ] was found to have higher jitter values, even though it was found to have *raised* F0 and spectral tilt. Jitter does not seem to correlate with creakiness then as it was expected to.

One possibility is that jitter is instead correlating with the tenseness of the vocal folds. This is possible since both significant differences in jitter ([ʔ] > [h] and [pʰ] > [m] for low, mid and rising tone) correlate with increased F0. As Halle & Stevens (1971) first pointed out, tense vocal folds cause an increase in F0, while slack vocal folds cause a decrease in F0. If [ʔ] is articulated with tense vocal folds, and if the status of [m] as an F0-depressor is due to its production with relatively slack vocal cords, then jitter might correlate with the tense/slack distinction rather than with glottal constriction. However, this does not easily explain the fact that jitter is actually *greater* for [m] than [pʰ] after falling tone. Falling tone does have a higher pitch target and so the vocal folds will be relatively tense, but this should apply equally to both [m] and [pʰ], raising F0 and hypothetically, jitter, in both. Why it would induce higher jitter in [m], but not [pʰ] is not clear.

A second issue is with the status of [ʔ], which showed F0-raising effects as well as relatively high spectral tilt and jitter. Inspection of the glottal stops produced by the three speakers showed that, in fact there was little or no creak associated with them. Instead, a clear glottal stop release can be seen in the spectrogram in figure 7 below. Importantly, there aren’t any irregular glottal pulses indicative of creakiness.

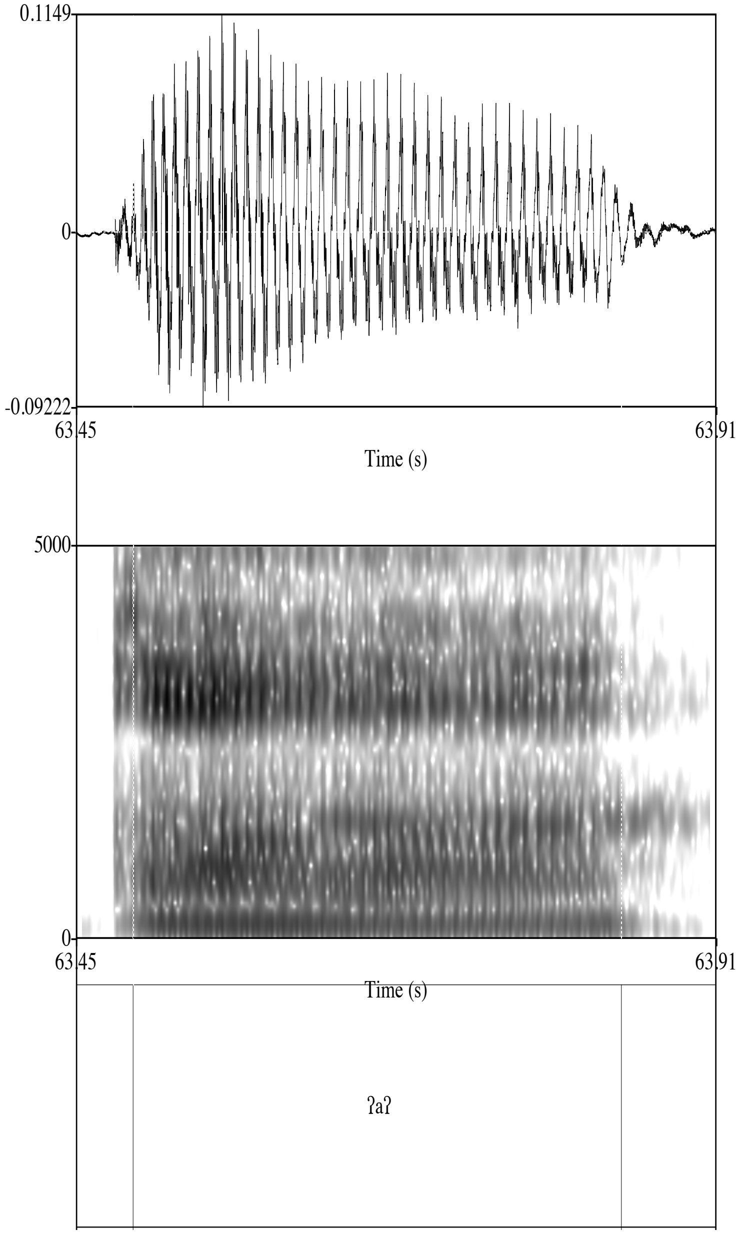

Fig. 7. Glottal Stop Onsets

The spectrogram in figure 7 has a sharp vertical boundary at the onset of the word, indicating the release of the glottal stop. [ʔ] in Thai is not associated with creakiness, but is rather associated with a glottal closure and release. This explains the results where F0 and spectral tilt are higher for [ʔ] than [p]. [p] is articulated with creakiness in the following vowel, but [ʔ] is not. [ʔ] must be articulated with a different kind of laryngeal constriction, one that is not associated with creakiness. The explanation offered previously is that glottal stops are produced with tense vocal cords, which explains the raised F0. Alternatively, Esling & Harris (2005) and Edmondson & Esling (2006) describe two modes of laryngealized voicing: creaky voice, which they note, is associated with lower F0, and what they describe as harsh voice, which they note is associated with higher F0. They further note that harsh voice is typically associated with constriction of the ventricular folds (located above the vocal folds), an articulation that they note is common in glottal stops. This suggests that Thai glottal stops employ harsh voice rather than creaky voice.

Another observation that requires comment was the finding that spectral tilt differences between [ʔ] and [p] were substantially less for low tone (see figure 5). The spectral tilt for [ʔ] is lower than with other tones. Possibly this is related to coarticulation between low tone (slack, creaky phonation) and [ʔ] (harsh, tense phonation), where [ʔ] is made less harsh or slack when followed by low tone. However, this should also be seen with rising tone, but it is not. Additionally, this would not explain the increase in spectral tilt seen in [p] with low tone. The net effect is that low tone is essentially pulling the spectral tilt to some neutral state. It is unclear why this happens though.

The comparison between [h] and [pʰ] was made with the phonological high-tone restriction in mind. Hypothetically, the ban on [h], but not [pʰ] preceding high tone would correlate with the fact that [h] would be breathy, whereas [pʰ] would not. Evidence for this was discovered in that [h] had higher spectral tilt than [pʰ]. However, the F0 measurements were not significantly different, whereas breathiness should lower F0. A similar finding where spectral tilt apparently distinguished between the onsets was found with [b] and [p]. While no difference was discovered between the F0 measurements, [p] was found to have lower spectral tilt than [b]. The resolution of the spectral tilt comparisons seems finer than the F0 comparisons then.

This difference in resolution actually has a principled explanation. F0 effects of onset consonants in non-tonal languages were found to be both larger and lasting over a longer duration by Hombert et al (1979). They suggest that this effect can be explained since in tone languages, F0 is a primary indicator of lexical contrast, whereas in non-tone languages, it is not. Therefore, tone language speakers actively minimize the phonetic effects that consonants have on F0. Kingston & Diehl (1994) note that speakers can actively control phonetic details in this manner. If this is the case in Thai (as Gandour, 1974 suggests it is), then the fact that F0 effects are less significant than spectral tilt is unsurprising. Spectral tilt is not involved (directly) in any contrast in Thai, and so it is not controlled to the same extent as F0. As such, it is a better indicator of laryngealization than F0 in a tone language such as Thai.

The picture that emerges is one where voiced and voiceless unaspirated stops are laryngealized in a different manner than glottal stops in Thai (creaky versus harsh). There is additional evidence that [h] is breathy, unlike [pʰ], and that [m] depresses F0, without any evidence of laryngealization. A phonological account of onset-high tone restrictions that refers directly to these phonetic findings can capitalize on this. Previous accounts that used the feature specification [–spread glottis] (Ruangjaroon, 2006; Lee, 2008) did so in order to group the voiced and voiceless unaspirated stops in a single class. However, both are laryngealized phonetically, suggesting that Thai involves [+constricted glottis], rather than [–spread glottis] as the active feature. The ban on [ʔ] and [h] with high tone would then be handled via some other mechanism.

Given the differences in the articulatory mechanisms for laryngealization between [ʔ] on the one hand, and [b] and [pʰ] on the other, might there be two separate phonological features distinguishing these modes of laryngealization (i.e. [creaky] and [harsh])? It is notable that although [ʔ] raises F0, it is banned with high tone. This suggests, instead that the Thai phonology groups laryngealized sounds together as a class, regardless of their phonetic effects on F0 or spectral tilt and imposes a ban on high tone. This of course means that the phonological restriction is not completely natural, and instead that an abstraction is being made over a restriction that *is* natural (that [b] and [p] should not occur with high tone) onto a restriction that is not natural (that [ʔ] should not occur with high tone). This opens up questions into the phonological representation of laryngeal articulations, including tone and laryngeal consonants that go beyond the reach of this paper, and so they must be left for future research.

References

Abramson, Arthur. 1962. 'The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments', International Journal of American Linguistics 28, 2, part II. (Also published by the Indiana Research Center in Anthropology, Folklore, and Linguistics: Bloomington, Indiana).

Boersma, Paul, and David Weenink. 2005. Praat (version 4.3.31). Amsterdam: University of Amsterdam. Online: http://www.fon.hum.uva.nl/praat.

Bradshaw, Mary. 1998. *A Cross-Linguistic Study of Consonant-Tone Interaction.* Ph.D. dissertation, Ohio State, Columbus, OH.

DiCanio, Christian. 2007. (Praat Script for Extracting Spectral Tilt). Berkeley: University of California, Berkeley. Online: http://www.linguistics.berkeley.edu/~dicanio/Get_Spectral_Tilt.praat.

Downing, Laura J., and Gick, Bryan. 2001. Voiceless Tone Depressors in Nambya and Botswana Kalang'a. In: *Berkeley Linguistics Society 27*, 65-80.

Edmondson, Jerold A., and Esling, John H. 2006. The valves of the throat and their functioning in tone, vocal register and stress: laryngoscopic case studies. *Phonology* 23:157-191.

Erickson, D. 1975. Phonetic implications for a historical account of tonogenesis in Thai. In J. G. Harris & J. R. Chamberlain (eds.), Studies in Tai Linguistics in Honor of W. J. Gedney. Bangkok, Central Institute of English Language Office of State Universities, 100-111.

Esling, John H., and Harris, Jimmy G. 2005. States of the glottis: an articulatory phonetic model based on laryngoscopic observations. In: *A figure of speech: a* 228 *Festschrift for John Laver*, eds. William J. Hardcastle and Janet Mackenzie Beck, Mahwah, NJ: Erlbaum, 347-383.

Gandour, Jack. 1974. *On the Representation of Tone in Siamese.* UCLA Working Papers in Phonetics **27**: 118-146.

Gandour, Jack & Ian Maddieson. 1976. *Measuring Larynx Movement in Standard Thai Using the Cricothyrometer.* Phonetica **33**: 241-267.

Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: a cross-linguistic overview. Journal of Phonetics **29,** 383-406.

Halle, M. and K. N. Stevens. 1971. A note on laryngeal features. Quarterly progress report (vol. 101, pp.198–212). Cambridge, MA: Research Laboratory of Electronics, MIT.

Harris, J.G. 1972. *Phonetic notes on some Siamese consonants.* In Harris and Noss Tai phonetics and phonology, pp. 8-22. Central Institute of English Language, Mahidol University, Bangkok.

Holmberg, E., Hillman, R., Perkell, J., Guiod, P. & Goldman, S. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice, *Journal of Speech and Hearing Research*, 38, 1212}1223.

Hombert, Jean-Marie, John Ohala and William G. Ewan. 1979. Phonetic explanations for the development of tones. Language **55**: 37-58.

Keating, Patricia and Christina Esposito. 2007. "Linguistic Voice Quality", UCLA Working Papers in Phonetics #105, pp. 85-91.

Kingston, J. & R. Diehl. 1994. Phonetic Knowledge. Language **70** (3): 419-454.

Kingston, John 2005. The phonetics of Athabaskan tonogenesis, in S. Hargus & K. Rice (eds.), Athabaskan Prosody, John Benjamins Press.

Ladefoged, Peter. 2003. Phonetic data analysis: an introduction to fieldwork and instrumental techniques. Blackwell Publishing, Malden MA.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford, Blackwell.

Laver, John. 1994. Principles of Phonetics. Cambridge University Press, Cambridge, United Kingdom.

Lee, Seunghun. 2008. Consonant-Tone Interaction in Optimality Theory. Ph.D. dissertation, Rutgers University. New Brunswick, NJ.

Maddieson, Ian. 1977. 'Tone Effects on Consonants', UCLA Working Papers in Phonetics 36: Studies on Tone.

Morén, Bruce and Elizabeth Zsiga. 2006. *The Lexical and Post-Lexical Phonology of Thai Tones.* Natural Language & Linguistic Theory **24**: 113–178.

Ruangjaroon, Sugunya. 2006. *Consonant-Tone Interaction in Thai: An OT Analysis.* Taiwan Journal of Linguistics **4.2**: 1-66.

Slayden, Glenn. 2009. "IPA Query Results via AJAX." thai-language.com.

Stevens, Kenneth N. 1977. Physics of Laryngeal Behavior and Larynx Modes. *Phonetica* 34:264-279.

Tang, Katrina. 2008. *The Phonology and Phonetics of Consonant-Tone Interaction.* Ph.D. dissertation, University of California Los Angeles.

Tumtavitikul, Apiluck. 1993. *Perhaps, the Tones are in the Consonants?* Mon-Khmer Studies 23**:** 11-41.