

Article

The Phonetics of Tone and Voice Quality Interactions in Sylheti

Amalesh Gope 

Department of English, Tezpur University, Tezpur 784028, Assam, India; amalesh@tezu.ac.in

Abstract: This paper examines the phonetic interactions of tone and voice qualities in Sylheti. Data from six native speakers are examined to understand the voice qualities of the vowels carrying contrastive tones. The results identify three spectral measures (viz., H1*–H2*, H1*–A2*, and H1*–A3*) and one noise measure (viz., CPP) as reliable indicators of modal (or in the continuum of *modal to tense*) vs. breathy (or, in the continuum of *breathy to lax*) phonation contrasts in the vowels carrying high and low tone, respectively. Finally, a statistical model is proposed that predicts consistent phonation contrasts across the total duration of the contrastive tones.

Keywords: Sylheti; tone; voice quality; tense phonation; lax phonation; spectral tilt

1. Introduction

Phonation or voice quality refers to the production of speech sounds by the vibration of vocal folds (Ladefoged 1971; Silverman et al. 1995; Gordon 2001; Wayland and Jongman 2002; Esposito and Khan 2020). This paper seeks to examine the phonetic contribution of phonation in the production of tonal contrasts in Sylheti. This eastern Indo-Aryan language exhibits two-way tonal contrasts (Gope 2016; Gope and Mahanta 2014).

Voice quality distinctions or phonation types among phonemes are exploited in many world languages to preserve lexical contrasts. Languages may employ voice quality distinctions on either consonant (obstruents) as in Hindi (Ohala 1993b; Dixit 1989; Dutta 2007), or Bangla (Khan 2010), or on vowels as in many Zapotec languages (Jones and Knudson 1977; Esposito 2010b). Only a handful of languages reported employing phonation contrasts on both consonants and vowels as a phonemic property to draw lexical contrasts. The list includes Gujarati (Khan 2012; Esposito et al. 2019), !Xóõ (Traill 1985), Ju |'hoansi (Miller 2007), Wa (Watkins 2002), and White Hmong (Esposito 2012; Esposito and Khan 2012) as reported in Esposito and Khan (2020).

Furthermore, phonation contrasts may also function either as prosodic indications (stress and focus in German (Mooshammer 2010)) or as added cues to identify tonal contrasts Green Mong (Andruski and Ratliff 2000) and Chichimec (Kelterer and Schuppler 2020), or the correlation between creaky voice and low dipping tone in Mandarin (Belotel-Grenié and Grenié 1994). Studies on tonal contrasts primarily focus on pitch level and movement, and define tone based on pitch contrasts (Yip 2002; Hyman 2010). As such, the phonation properties, for long, were not considered to be part of phonological representations of tones. However, the recent developments in phonation studies show that tone and phonation may co-exist in many languages. In fact, phonation contrasts that are independent of tone are reported to exist in a few languages. For example, three phonation contrasts are combined with three-level tones (plus contours) in Jalapa Mazatec (Garellek and Keating 2011).

The primary objective of this paper is to examine the phonetic interactions of the vowels carrying contrastive tones¹ (viz., high and low) in terms of f_0 variations and voice qualities in Sylheti. This paper further examines different acoustic components associated with different voice quality measurements for the male and female speakers to determine the timing of the voice quality contrasts in Sylheti with a qualitative and quantitative acoustic analysis. The following research questions are examined in this paper: (i) Does



Citation: Gope, Amalesh. 2021. The Phonetics of Tone and Voice Quality Interactions in Sylheti. *Languages* 6: 154. <https://doi.org/10.3390/languages6040154>

Academic Editors: Christina Esposito and Sameer Khan

Received: 7 May 2021

Accepted: 15 September 2021

Published: 23 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

the tonal contrasts in the Sylheti lead to different voice qualities in the vowels (carrying contrastive tones)? (ii) If so, how does the pitch and voice quality interact in the tonal space? (iii) How are such interactions maintained across gender (male and female) and vowel timecourse? (iv) Are the (assumed) voice quality contrasts realized in a similar way (in terms of acoustic measures) for both male and female speakers' data? Finally, this study proposes a statistical model that predicts consistent phonation contrasts across the total duration of the contrastive vowels produced by both female and male speakers.

The Language: Sylheti

Sylheti, also written as *Syloti* (ISO 639-3), is an (eastern) Indo-Aryan language (sub-group of the Indo-European language family) spoken by over 11 million people across the globe. The majority of the Sylheti native speakers reside in the Surma river valley of Sylhet Division that comprises four districts viz., Sylhet, Sunamgonj, Hobigonj, and Moulvibazar of north-east Bangladesh, and in the Barak River valley of the neighboring north-eastern Indian states of Assam and (north) Tripura. The Barak valley division of Sylheti speakers in India includes three districts in Assam, viz., Cachar, Hailakandi, and Karimgonj; a few districts such as Dharmanagar, Kailasahar, and Kumarghat in (north) Tripura, and the Jiribam district in Manipur. Many Sylheti speakers are also settled in the United Kingdom and several other diasporic communities in South Asia and worldwide. The data for the current study is recorded from the Sylheti speakers settled in Assam and (north) Tripura in India. This variety can be termed the *Cachar variety of Sylheti* (named after one of the districts of Assam).

Sylheti is a tonal language (Gope 2016; Gope and Mahanta 2014). The presence of tonal contrasts following the loss of aspirated series², both in voiced and voiceless segments, has been well documented in Gope (2016); and Gope and Mahanta (2016b). Studies on tonal contrasts generally associate a lower pitch (in the neighboring vowel) with the voiced obstruent. Their voiceless counterparts are expected to raise the pitch in the following vowel (Yip 2002; Hombert et al. 1979). However, such predictions do not seem to work in Sylheti strictly. Pieces of evidence derived through instrumental experiments along with the statistical quantifications show that Sylheti exhibits a high tone following the loss of aspiration (primarily in the voiced segments), and a low tone is observed elsewhere ([ḍán] 'paddy' [ḍ^h > ḍ] and [ḍàn] 'donate,' [bá ṭ] 'rice' [b^h > b], and [bà ṭ] 'arthritis') (Gope 2016, 2018; Gope and Mahanta 2015). In a recent study, Raychoudhury and Mahanta (2020) reported that the loss of aspiration in the voiceless segments occurring in (different) onset positions, viz., in the first or the second syllable of a disyllabic word, might trigger a three-way lexical contrast in Sylheti. The study claims that a high tone is surfaced following the loss of the feature (–voice, –asp.) occurring in the second syllable of a disyllabic word [xútá] 'room' [t^h > t]; while, the (onset) feature (–voice, +asp.) occurring in the first syllable of a disyllabic word resulted in a contrasting tone-lowering [xùtà] 'taunting,' [k^h > x]. A default mid-tone is observed where there is no such apparent loss [xūtā] 'stick' [k > x].³ The development of contrastive tones following the loss of the feature (+spread glottis, –voice)⁴ can be observed in Figure 1i,ii. The current paper examines the correlation between tonal contrasts (high and low) and (a possible) voice quality contrasts in Sylheti. It must be noted that the status of the high tone reported in this paper is developed due to the loss of aspiration in the voiced segments.

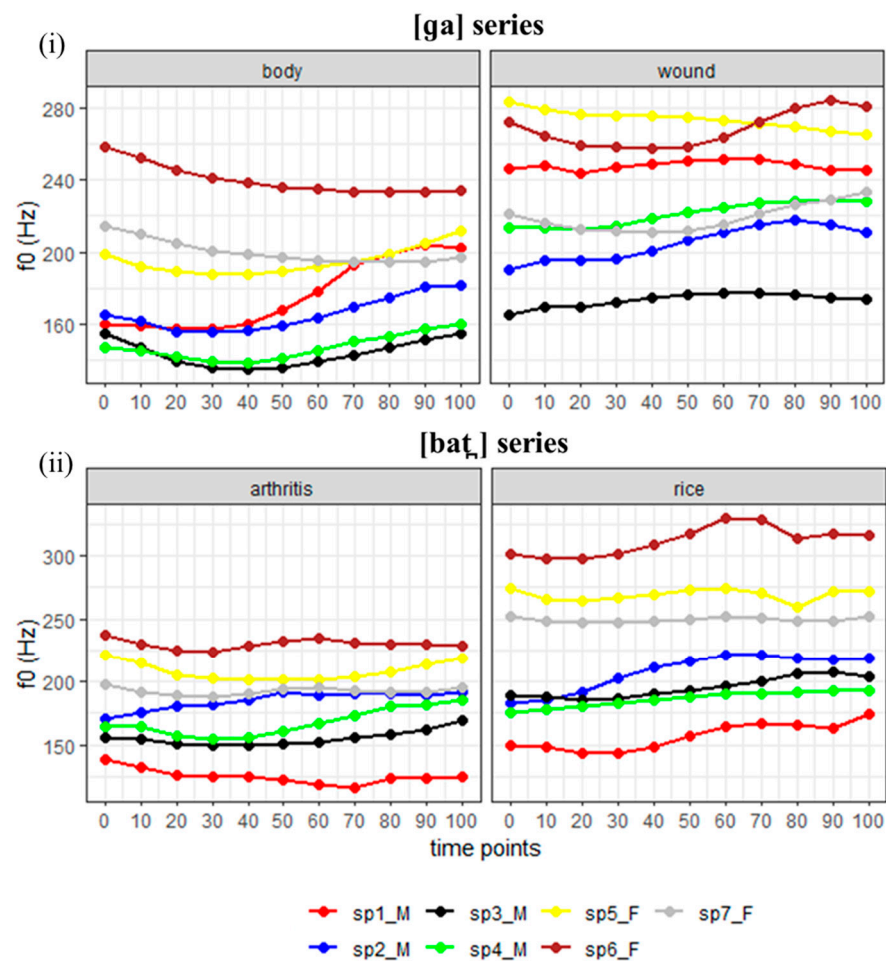


Figure 1. The loss of aspiration in [ga] ‘wound’ [g^h > g] (i) and [baʈ] ‘rice’ [b^h > b] (ii) led to a high tone in Sylheti. (n = 8) *per subject* for each word that has been shown in each panel (4 repetitions × 2 words).

2. Phonation Types

Ladefoged (1971) argued that degrees of phonation types could be determined depending on the variations of the glottal constriction continuum. In his model, Ladefoged (1971) proposed that the size of the glottis might range from voiceless (when the vocal folds are held furthest apart), through breathy voice (where the glottis is held more open), to regular (modal voicing), to creaky (produced with a constricted glottis), and lastly to glottal closure (when the vocal folds are held closest together, hence no vibration and without phonation). According to Esposito and Khan (2020), modal phonation stands out as the most common phonation type in the majority of the world’s languages that distinguish between one or both of the extremes of the continuum (i.e., voiceless sounds) and one center point (i.e., voiced sounds). However, many languages make distinctions within the voiced range of this continuum (viz., breathy, lax, tense, and creaky) in addition to modal phonation (Esposito and Khan 2020).

Since the glottal constriction is a continuum, the degrees of creakiness, breathiness, or even the modal voice might vary from more constricted to more open. Two vital intermediate phonations exist along this continuum, viz., *lax* and *tense*. In terms of Ladefoged’s (1971) continuum model, lax phonation is a phonation in the breathy-modal range, whereas tense phonation is a phonation in the creaky-modal range (Kuang and Keating 2013). Lax and tense phonations, therefore, refers to the low and high degree of muscular tensions, respectively. The four types of non-modal phonations, viz., breathy, lax, tense, and creaky, are examined in relation to modal phonation in the continuum.

Ladefoged (1971) further added that these phonation categories are the outcome of the controllable variations produced by exploiting the state of the glottis, either by individual idiosyncratic efforts or due to unintentional pathological conditions and are not absolute. Hence, these might differ somewhat across languages and even across speakers of the same language. For example, what counts as a breathy voice for one speaker might count as modal phonation for another in the same language.

Acoustic Correlates of Phonation

The primary acoustic display of waveforms and spectrograms are capable of drawing a primary distinction between a modal and non-modal phonation such as creaky (irregularly spaced glottal pulses and reduced intensity compared to modal) and/or breathy (lower f_0 relative to modal and by substantial aperiodic or noisy Energy) (Gordon 2001). However, the waveforms and the spectrograms do not show a consistent pattern that can be used to draw distinctions between different types of phonations, especially in the case of connected speech. To overcome this problem, many researchers proposed a variety of acoustic properties suitable for measuring various phonation types (Kirk et al. 1993; Iseli et al. 2007; Hillenbrand et al. 1994; Fischer-Jorgensen 1967). Among those, spectral balance and spectral tilt measurements have been the most popular and consistent. The difference between the amplitude of the first and second harmonics (H1–H2) is regarded as the most consistent and popular spectral balance measure. In a cross-linguistics phonation study on Gujarati, White Hmong, Jalapa Mazatec, and Southern Yi, Keating et al. (2010) observed that of all the acoustic measures, only H1–H2 was successful and consistent in all four languages in distinguishing phonation contrasts. H1–H2 has also been successful in distinguishing modal and (different types of) non-modal voice in many other languages such as Hmong (Huffman 1987; Gordon 2001); Mazatec (Silverman et al. 1995; Blankenship 1997), !Xóõ (Ladefoged 1983; Ladefoged et al. 1988), Khmer (Wayland and Jongman 2002), Jingpo, Hani, Eastern Yi, Wa (Maddieson and Hess 1986), and so on.

Spectral tilt measures, on the other hand, quantifies the amplitude between the first harmonic (H1) and the harmonics existing higher r formants such as (H1–A₁), (H1–A₂), and (H1–A₃). Spectral tilt measures have been successful in distinguishing phonation contrasts in languages, such as English (Stevens and Hanson 1995), Krathing Chong (Blankenship 2002), Takhian Thong Chong (DiCanio 2009), Gujarati (Khan 2012), Marathi (Berkson 2019), Phnom Penh Khmer (Kirby 2014), and so on.

The spectrum is weaker for a breathy voice in the higher frequencies than a modal voice; however, it is stronger for a creaky voice. On the other hand, during the production of a breathy voice, the glottis opening might create an intense audible noise that can distinguish it from other phonation types. In general, all the spectral balance and tilt measures are reported to be the highest for breathy phonation, intermediate for modal, and lowest (and often negative) for creaky phonation. Similarly, these measures are also higher in lax phonation than tense phonation in many Tibeto-Burman languages (Kuang and Keating 2014).

Furthermore, non-modal phonations such as breathy and lax can also be quantified by the presence of noise. Cepstral peak prominence (CPP) reflects the harmonics-to-noise ratio (Hillenbrand et al. 1994). A more significant cepstral peak indicates stronger harmonics above the floor of the spectrum. Irregular vibration associated with creaky and tense phonations and noise due to turbulent airflow in breathy and lax phonations generally associate lower CPP values in all the non-modal phonations than modal phonation. In languages such as Hmong, Mazatec, and Yi, phonation contrasts are identified through the CPP measure (Keating et al. 2010).

Gordon and Ladefoged (2001) noted that non-modal phonation types are generally associated with lowering fundamental frequency. For example, in a language such as Mam (England 1983) and many Northern Iroquoian languages such as Mohawk, Cayuga, and Oneida (Chafe 1977; Michelson 1988; Doherty 1993), the creaky phonation is associated with a lower f_0 (relative to modal phonation). This lowering effect of creaky voice, however,

is not universal across languages. [Hombert et al. \(1979\)](#) showed that the process of glottalization could be associated with high tone in the historical development of some of the Athabaskan languages, while the same glottalization process is associated with a low tone in closely related languages ([Leer 1979](#); [Gordon and Ladefoged 2001](#); [Kingston 2011](#)). On the other hand, breathy phonation is more consistently connected with a lower f0 in the majority of languages ([Gordon and Ladefoged 2001](#)).

An experiment has been conducted to examine various spectral components mentioned above to understand the voice-quality related properties of the vowels associated with contrastive tones in Sylheti. The following section discusses the experimental design, the methods adopted for the acoustic analysis, and the statistical tests and procedures employed for the current experiment

3. Methods

3.1. Speakers and Data

This study is based on 14⁵ monosyllabic words with [a] as the nucleus that is specified for contrastive tones (viz., high and low) in Sylheti (Table 1). Six⁶ native Sylheti speakers (4 M and 2 F), aged between 19 and 30 years, repeated the dataset at least three times. All the participants use Sylheti as their first language. Apart from Sylheti, all the participants had a working knowledge of Hindi and English. Subjects were given a token of remuneration for participating in the production experiment. The entire recording took place in a quiet room situated in the Dharmanagar district of (north) Tripura.

The target words were embedded in a fixed carrier sentence (*'I said X'* [ami X xoi.ar], X is the target word). The dataset also includes priming sentences⁷ for each word to trigger the actual tonal contrasts in the lexical items embedded in the fixed sentence frame.

Table 1. Dataset displaying the list of monosyllabic words considered in this work.

Sylheti Words	Gloss	Sylheti Words	Gloss
[bán]	'to pretend'	[bàn]	'to tie'
[bát]	'rice'	[bàt]	'arthritis'
[dák]	'drum'	[dàk]	'roaring of cloud'
[dán]	'paddy'	[dàn]	'donate'
[zál]	'hot/spicy'	[zàl]	'net'
[gá]	'wound'	[gà]	'body'
[xál]	'drain'	[xàl]	'skin'

3.2. Data Annotation and Acoustic Measures

Each recording was prepared and annotated in Praat ([Paul and Weenink 2008](#)), and a Praat TextGrid was created for each iteration. First, the target word was manually separated from the fixed sentence frame, and an individual .wav file was created for each target word. The target words were then segmented, and a TextGrid file containing one tier was created in which the target vowels were manually labeled. These soundtracks were then manually examined for a possible voice quality contrast.

The annotated sound files were analyzed using *VoiceSauce* ([Shue et al. 2009](#)) for each vowel portion. The acoustic components considered in this study include the spectral balance and spectral tilt measures, i.e., the Amplitude differences of various Harmonics and Formants viz., H1*–H2*, H1*–A1*, H1*–A2*, H1*–A3*, Fundamental Frequency (f0) calculated using the STRAIGHT algorithm ([Kawahara et al. 1998](#)), and cepstral peak prominence (CPP) and Energy (a measure that calculates the overall intensity). Each acoustic component was automatically measured at every millisecond of the target vowel portion and averaged across nine equal timepoints of the vowel's duration. The division of nine equal timepoints of the vowel duration will enable us to track the possible acoustic changes across the total duration of the target vowel. Apart from the values of nine

timepoints, I have also used the mean values of each measure to draw an initial assessment. A total of 18,144 data points (14 words \times 3 repetitions \times 6 subjects \times 9 timepoints \times 8 acoustic components \times 8 measures) were examined in this study.

It has to be noted that all the spectral balance and spectral tilt measures are expected to be highest for breathier phonation and lowest (or even negative) for creakier phonation. An intermediate value is expected for modal phonation. CPP or Energy, on the other hand, are expected to be lower for non-modal phonations compared to modal phonation. Furthermore, it is also important to understand that there is no specific range of values that can be considered to be modal, breathy, or creaky. These specifications are absolutely relative to each other.

4. Results

4.1. Fundamental Frequency (f_0), CPP, Energy, and Spectral Tilt Measures

Figure 2i,ii shows the mean⁸ f_0 of the vowels associated with high and low tones (Figure 2i) and the distribution of f_0 across nine timepoints by the female and male speakers (Figure 2ii). The analysis of a one-way ANOVA confirms that f_0 is significantly different for each tone types [F (1, 1708) = 321.9, $p \leq 0.001^*$]. On average, the high-tone vowels are observed to be 27 Hz higher than their low-tone counterparts. This difference is almost the same across all the nine timepoints in both female and male speakers' data. Please note that higher f_0 is associated with the words that contained (+spread glottis) historically. The diagrams in Figure 2 confirm that the loss of (+spread glottis) results in a high tone in Sylheti. The possible differences in voice quality on the vowels carrying contrastive tones are examined in the succeeding figures.

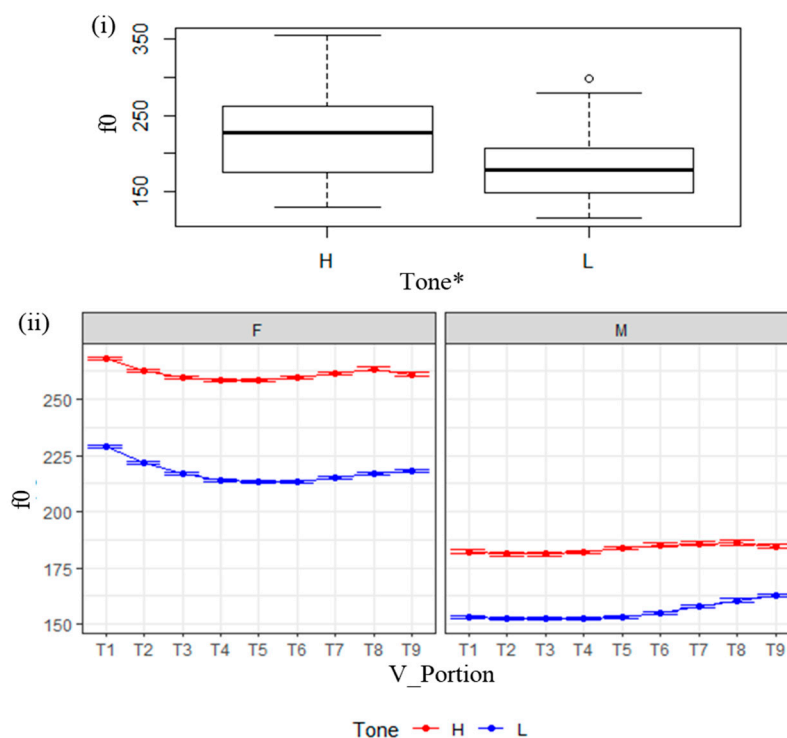


Figure 2. (i) Box plot representing the mean f_0 shows a significant difference in terms of tone types (viz., high and low) (indicated with an asterisk in the tone). (ii) The f_0 values averaged across nine timepoints (shown as T1, T2, . . . T9) show that the high tone is generated following the loss of the feature (+spread glottis) shown in Table 1. F = female, and M = male.

Figure 3i–iv shows the mean values of the acoustic measures considered in this study. All the box plots are drawn in R (Version 3.6.3) embedded with R studio (Version 1.2.1335, RStudio, Inc., Boston, MA, USA) using the *boxplot(y ~ x)* function where *y* is the dependent variable (viz., CPP, Energy H1*–H2*, H1*–A1*, H1*–A2*, and H1*–A3*), and *x* is the independent variable, i.e., vowels representing the tonal contrasts high (H) and low (L). All these parameters are observed to be significantly higher in the vowels carrying low tones (except for CPP). In CPP, however, the values are significantly higher in the vowels associated with a high tone.

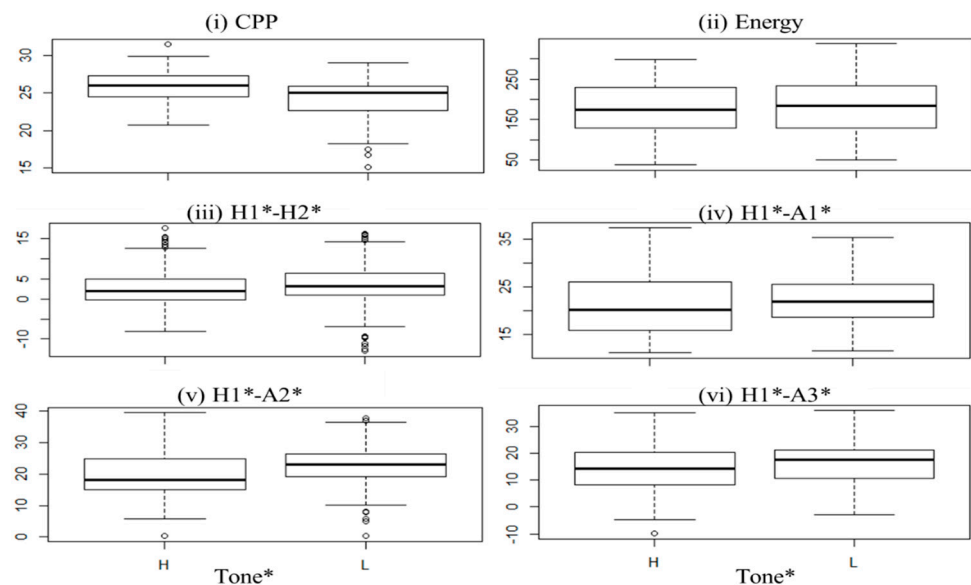


Figure 3. Box plots representing the mean values for each parameter: (i) CPP, (ii) Energy, (iii) H1*–H2*, (iv) H1*–A1*, (v) H1*–A2*, and (vi) H1*–A3*. Tone* indicates significant interactions in terms of Tone types.

A one-way ANOVA was conducted to examine the effect of dependent variables (the acoustic measures considered in this study) on the independent variable (tone). For this, the *avov(y ~ x)* function in R is used, where *y* is one of the dependent variables (viz., CPP, Energy H1*–H2*, H1*–A1*, H1*–A2*, and H1*–A3*), and *x* is the independent variable, i.e., the vowels that represent High (H) and Low (L). The ANOVA⁹ results confirm that the vowels carrying contrastive tones are specified with contrastive phonations in terms of all the acoustic measures (viz., CPP ($F[1, 1708] = 29.26, p \leq 0.001^*$), Energy ($F[1, 1708] = 4.922, p = 0.027^*$), H1*–H2* ($F[1, 1708] = 25.53, p \leq 0.001^*$), H1*–A1* ($F[1, 1708] = 5.128, p = 0.024^*$), H1*–A2* ($F[1, 1708] = 52.77, p \leq 0.001^*$), and H1*–A3* ($F[1, 1708] = 24.86, p \leq 0.001^*$)). The trend observed in Figure 3 indicates that the vowels associated with high tone are in the continuum of the modal to tense. In contrast, the vowels associated with low tone are in the continuum of lax to breathy.

To examine and understand if the differences observed while comparing the mean values of all these measures are reflected across the total duration of the vowels with high and low tone, and if all the acoustic measures yield similar level of significance in both female and male speakers' data, I conducted another round of one-way ANOVA (following the same method discussed above) for each timepoint on female and male speaker's data separately.

The results (Figures 4–6) suggest that female and male speakers employ different strategies to maintain the voice quality distinctions in the vowels carrying contrastive tones. Furthermore, the raw values of each measure also indicate that the voice-quality differences are primarily reflected in the middle (to end) portion of the duration of the vowel. The detailed results for each measure are summarized in Tables S1–S12 in the Supplementary File. The timepoint values associated with Energy did not draw a consistent result in both male and female speakers' data (even though the mean values indicated it to be a significant parameter in the comprehensive data); thus, this parameter is not considered any further in this study. The scatterplot in Figure 4 suggests that CPP (timepoints 3 to 9) is indeed a significant parameter predicting the voice-quality differences in male speakers' data. However, the female speakers' data does not show any difference in terms of CPP measures. The scatterplot in Figure 5 indicates that both H1*–H2* (timepoints 1 to 5) and H1*–A1* (timepoints 5 to 8) correctly predict the voice-quality differences in the female speakers' data only. No such phonation contrasts in H1*–H2* and H1*–A1* measures are observed in male speakers' data. Figure 6, on the other hand, confirms H1*–A2* as a most consistent measure (mostly the middle to end timepoints) that distinguishes the voice quality contrasts in both female and male speakers' data. H1*–A3* (timepoints 3 to 9) predicts the voice quality differences in female speakers' data only.

This paper's findings so far confirm that the vowels specified with high and low tones also maintain distinct voice qualities in Sylheti. Phonation contrasts in female speech are identified with acoustic measures such as H1*–H2*, H1*–A1*, H1*–A2*, and H1*–A3*, while the phonation condition contrasts in male speech is exhibited through CPP, and H1*–A2* measures (Figures 4–6).

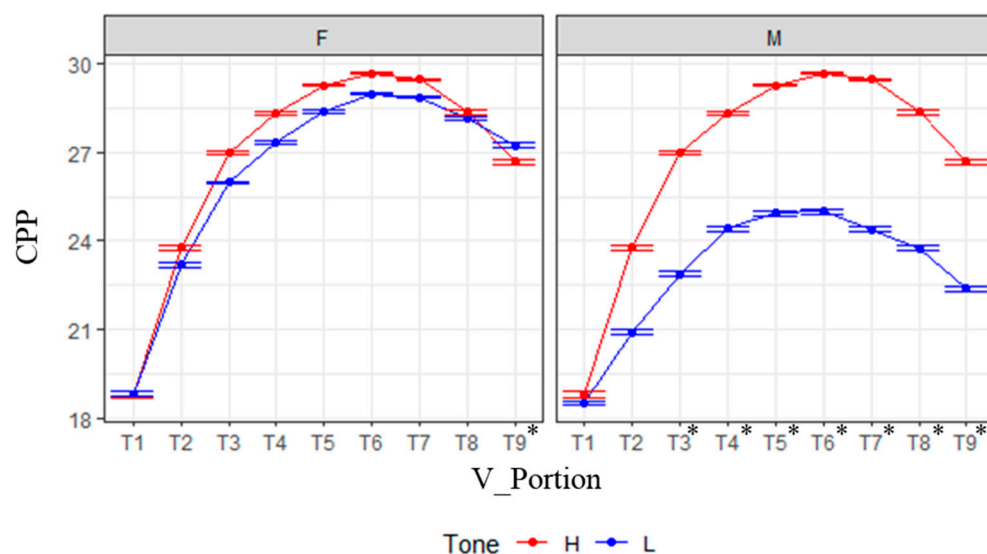


Figure 4. Scatter plots representing independent variables as a function of timepoints (T1 to T9) for CPP measure. The asterisk (*) at each timepoint indicates a significant difference in CPP values at a particular time point. F and M indicate female and male speakers' data, respectively.

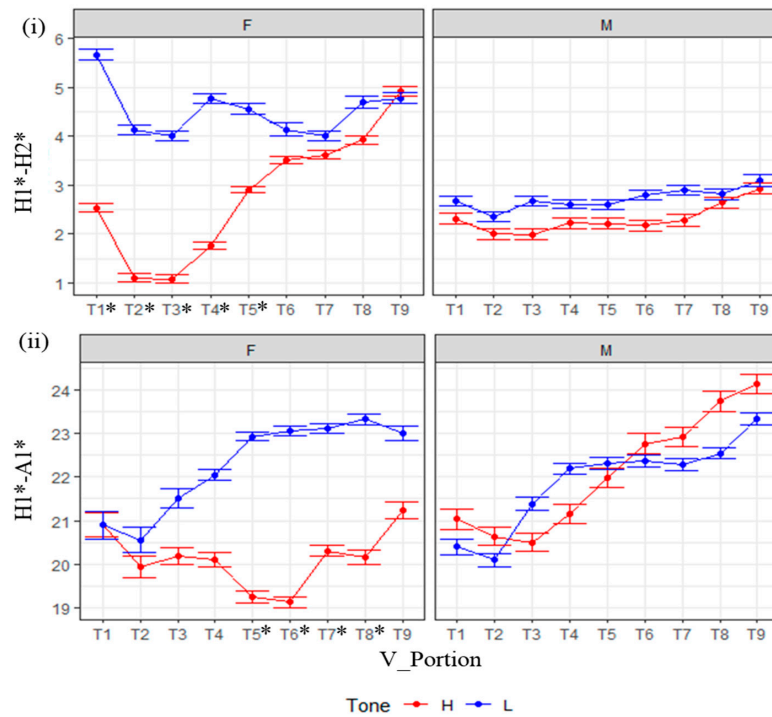


Figure 5. Scatter plots representing independent variables as a function of timepoints (T1 to T9) for (i) H1*-H2* and (ii) H1*-A1*. The asterisk (*) at each timepoint indicates significant interactions at a particular time point. F and M indicate female and male speakers’ data, respectively.

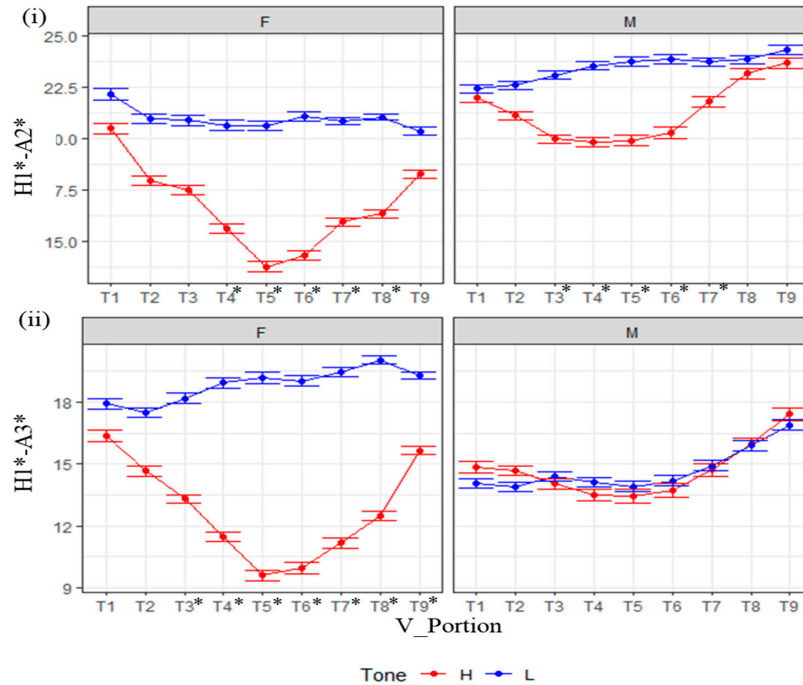


Figure 6. Scatter plots representing independent variables as a function of timepoints (T1 to T9) for (i) H1*-A2* and (ii) H1*-A3*. The asterisk (*) at each timepoint indicates significant interactions at a particular time point. F and M indicate female and male speakers’ data, respectively.

4.2. Statistical Model

4.2.1. Model Generation

Linear mixed-effects regression models were generated first with the values of dependent variables, viz., CPP, H1*-H2*, H1*-A1*, H1*-A2*, and H1*-A3* on the whole data set.

Two independent variables were used in this model, viz., time points (nine equal portions: T1, T2, T3, T4, T5, T6, T7, T8, T9), and tone types viz., high (H) and low (L). A combination of both the gender, viz., female (F) and male (M) speakers' data were also incorporated in this study to account for the gender-specific properties as observed in Figures 4–6. The models were generated for each dependent variable following the same procedure.

In this model, all the time points were converted to a “factor” with *as.factor()* function in R (Version: 1.2.1335, RStudio, Inc., Boston, MA, USA) to treat each level as a group rather than to treat these as a continuous variable. The *lm()* function was used to fit a basic linear model where the dependent variable (plotted in the *y*-axis) is assumed to vary linearly with the independent variable (plotted in the *x*-axis). The primary objective of this model is to explore if the dependent variable (i.e., the tonal contrasts leading to voice quality contrasts in the vowels carrying contrastive tones) significantly varies across all the time points concerning all the measures. The *x*-axis, therefore, was considered as the timepoints. The test was finalized using one of the dependent variables in the *y*-axis. The first assumed fitted model (*lm0*) is represented as $lm(y \sim x)$, where *y* = CPP and *x* = Timepoints. The fitted equation does not consider any interaction terms, and the model reflects a significant interaction in terms of CPP ($F(8, 1701) = 132.81, p < 0.001^*$ with Adj. $R^2 = 0.38$). The second model is generated by adding the terms, i.e., $lm1 = lm(y \sim x + Tone + Gender)$, $F(10, 1699) = 164.05, p < 0.001^*$ with Adj. $R^2 = 0.49$. It is to be noted that the *lm1* does not contain any interaction terms; however, the R-squared values were increased when the other parameters were taken into account. The next plausible model (*lm2*) included all the possible interaction terms- $lm2 = lm(y \sim x * Tone * Gender)$. In this model, the variables were specified as *x* Tone * Gender*. The model is expected to yield a new model in terms of *x*, *Tone*, *Gender*, and the respective interactions between them (*x*Tone*, *Tone*Gender*, and *x*Gender*). The generated model did not show much difference in the R-squared values ($F(18,1691) = 96.25, p < 0.001^*$ with Adj. $R^2 = 0.50$). In such a situation, we need to test whether all the interaction terms are significant using type II analysis of variance (ANOVA). The *anova()* function from the *car* package in R is used to get the results.

The generated model indicated that all the terms are significant except *x*Tone*. Hence, a revised model in $lm3 = lm(y \sim x +Tone * Gender)$ was generated further. The fitting parameters are $F(18, 1691) = 96.25, p < 0.001^*$ with Adj $R^2 = 0.51$. The *lm3* equation is used when the gender of the subjects is considered as one of the parameters. To elucidate the importance of gender and/or to find any effects of gender on the dependent variables, one can separate the male and the female data. If the gender is missing from the study, equation *lm3* will not work as there are no levels available (aka either M or F in the present case). To model in such a situation, one needs to discard the interaction term with gender. The equation of type $lm(y \sim x + tone)$ might work the best. However, it is dependent on the study and the type of data whether to model the raw data or transformed data. It is always important to confirm its validity with any of the appropriate statistical tests.

To ensure a better interpretation of the models for each dependent variable, the *emmeans* package in R was used to calculate *Estimated Marginal Means* (henceforth EMMs) and perform pairwise comparisons with Tukey's test. It is worth mentioning here that the use of *emmeans(lm3, specs = "Tone," by = "x")* package only will generate the same output of *emmeans* (i.e., *confidence intervals*, *estimate*, *t ratios*, and *p values*) at each *x* point. This indicates that the (estimated) difference between gender and timepoints will be neutralized for each level of tonal contrasts viz., high and low, the dependent variable is conditioned with. *Specs = "~ Tone + Gender * x"* is used to solve such an issue, and the calculations for the *emmeans* are performed on the log scale. The *emmeans* used in this study is (*lm3, specs = ~Tone + Gender * x, by = "x", trans = "log", type = "response"*). However, multiple interpretations can be drawn while dealing with the transformations in *emmeans* due to many possibilities making it a somewhat complex outcome. The key is to examine the outcome rationally. The summaries of the results are generated via a sequence of steps. For example, here, a reference grid is constructed for the transformed model. The predictions on the reference grid are averaged over the two levels of tone and two levels of gender to

obtain the EMMs reflected on the log scale. The standard errors and confidence intervals for these EMMs are further computed on the log scale. The *regrid(emm)* function in R is used to interpret the generated EMM results correctly. A careful observation of the *summary()* function will confirm that the transformation is no longer included in the structure. It is no longer relevant; the reference grid is on the same scale as the original data. However, the *confidence intervals*, *t ratios*, and *p values* are not identical for the transformed and non-transformed cases.

4.2.2. Outcomes of the Generated Model

The statistical output is tabulated for each dependent variable in the supplementary section, Tables S13–S24. The proposed model is capable of drawing consistent predictions across the nine timepoints and for different measures. Figures 7–9 confirm that the acoustic measures such as CPP, H1*–H2*, H1*–A2*, and H1*–A3* can draw significant differences across nine timepoints in both male and female speakers’ data. However, one of the spectral tilt parameters, H1*–A1*, did not appear to be a significant component in drawing the voice quality distinctions in male or female speakers’ data. The proposed model predicts higher CPP values associated with the vowels bearing high tone at each timepoint in both male and female speakers’ data. However, all the spectral balance and tilt measures confirm higher values associated with the vowels specified with low tone. These significant differences observed in terms of all the acoustic measures ensure that the tonal contrasts in Sylheti also maintain independent phonation contrasts.

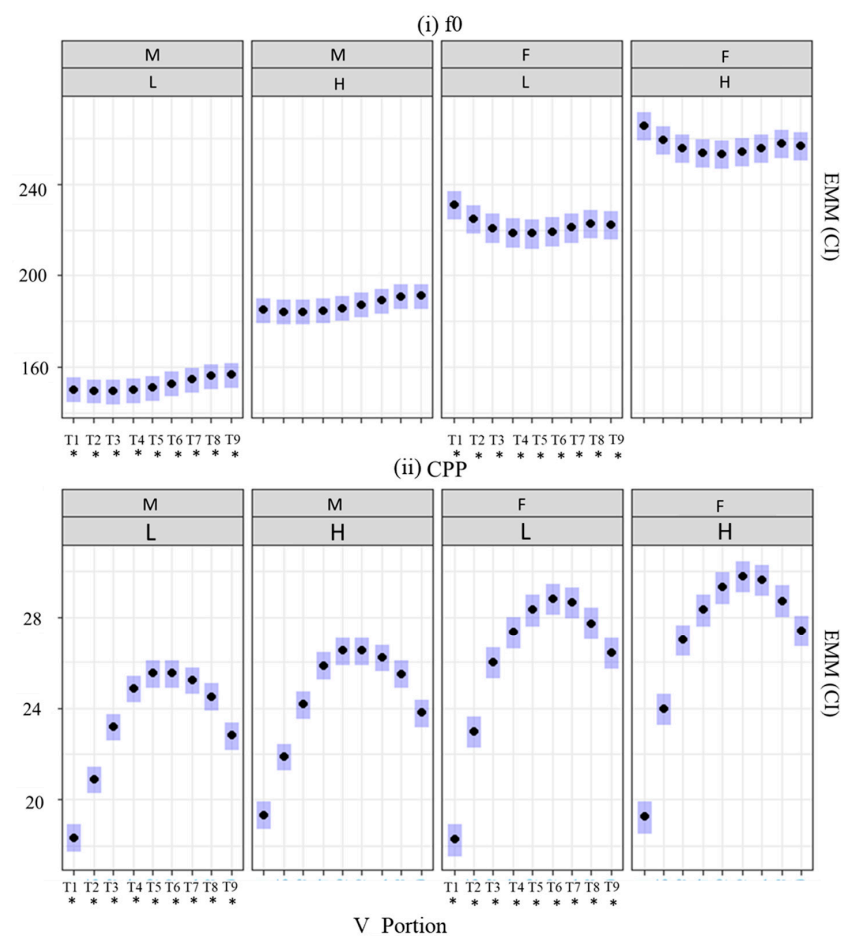


Figure 7. EMMs and confidence intervals (blue bars) of the nine timepoints of the vowels carrying contrastive tones (viz., low and high, indicated as L and H) derived from (i) f0 and (ii) CPP for male (M) and female (F) speakers separately. The asterisk (*) at each timepoint indicates the significance level of pairwise comparisons between the timepoint.

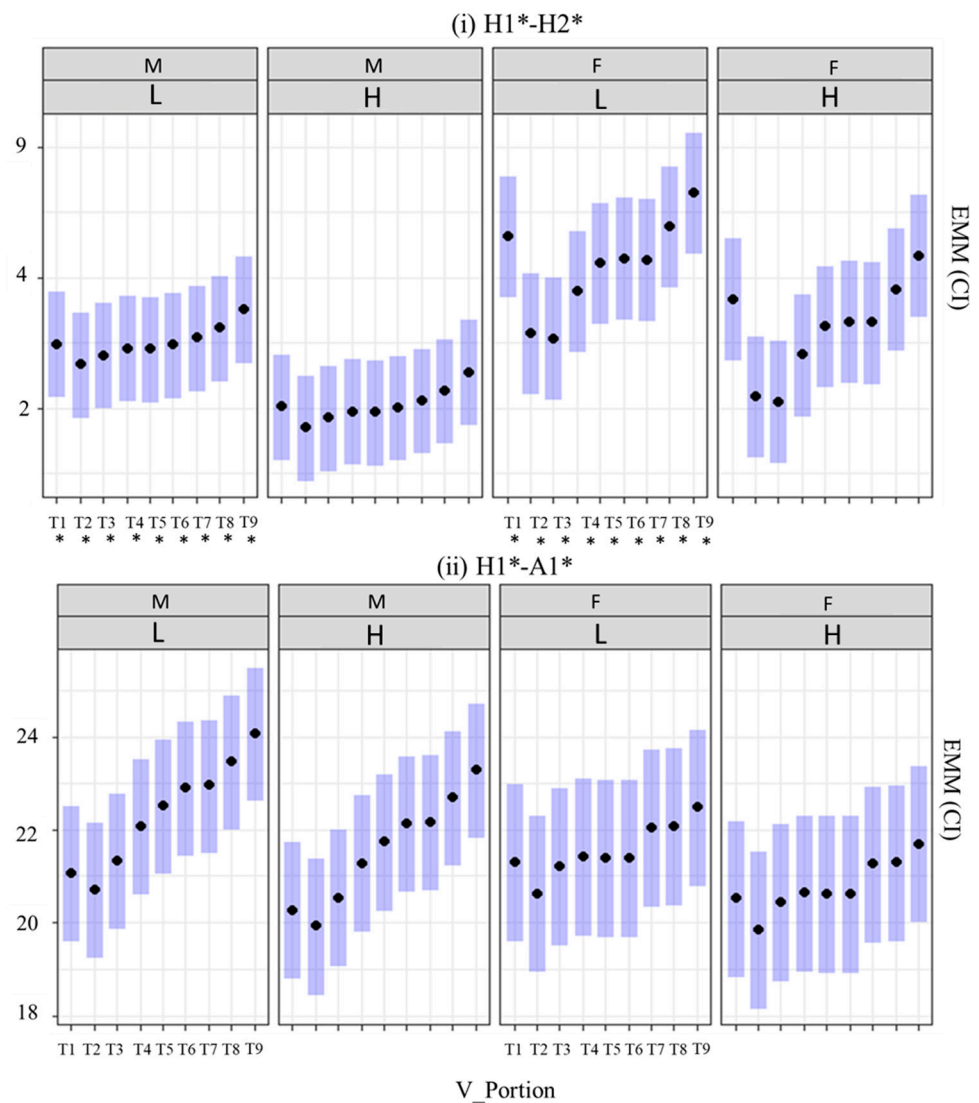


Figure 8. EMMs and confidence intervals (blue bars) of the nine timepoints of the vowels carrying contrastive tones (viz., low and high, indicated as L and H) derived from (i) H1*-H2* and (ii) H1*-A1* for male (M) and female (F) speakers. The asterisk (*) at each timepoint indicates the significance level of pairwise comparisons between the timepoint.

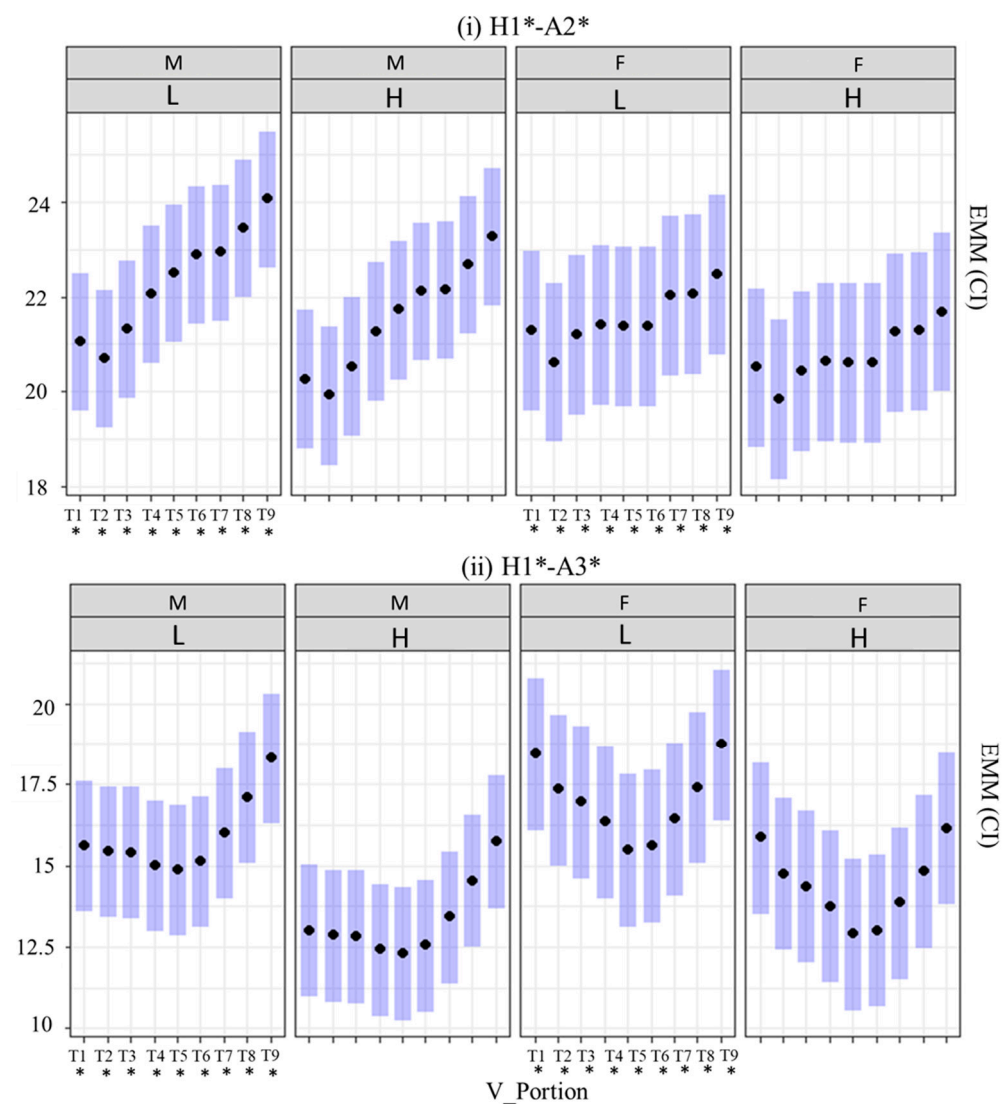


Figure 9. EMMs and confidence intervals (blue bars) of the nine timepoints of the vowels carrying contrastive tones (viz., low and high, indicated as L and H) derived from (i) H1*-A2* and (ii) H1*-A3* for male (M) and female (F) speakers separately. The asterisk (*) at each timepoint indicates the significance level of pairwise comparisons between the timepoint.

5. Discussion and Conclusions

The model proposed in this study addresses the inconsistency of the raw data for each acoustic measurement and specifies the relevant measures. The generated model predicts consistent voice quality differences in acoustic measures such as EMM, H1*-H2*, H1*-A2*, and H1*-A3* in both male and female speech. The detailed statistical results derived through the proposed statistical model are shown in T13–T24 in the Supplementary Section.

The results discussed so far answer all the research questions raised in this paper. Tonal contrasts in Sylheti display consistent distinctions in voice quality contrasts in male and female speech (question (i)). The high and low tones in Sylheti are observed to be inversely proportionate in terms of spectral balance and tilt measurements. EMM measure, on the other hand, shows higher values for the vowels carrying high tone in both male and female speech (question (ii)). However, the analysis of raw data indicates that the voice quality contrasts in male and female speech are maintained in terms of different acoustic measures (question (iii)). For example, CPP and H1*-A2* correctly predict the voice contrasts in male speech, whereas H1*-H2*, H1*-A1*, H1*-A2*, and H1*-A3* do so in female data. The raw data derived through various acoustic measures further confirms

that the phonation difference can best be observed (middle portions) of the tonal space in both male and female speech (question (iv)).

The findings discussed so far, nonetheless, open up more complicated questions. Data presented in this paper confirmed that the high tone in Sylheti is realized due to the loss of (+spread glottis) in the (+voice) segments (mostly).¹⁰ This means that high-toned vowels in Sylheti were historically associated with the voiced aspirated consonants that are likely to be breathy in earlier stages. As such, one might predict that those vowels should have been generally breathier even now. In this context, let us recall that the literature on phonation types generally associate greater and positive (spectral balance and tilt measures) values with breathy and lax vowels, an intermediate value for modal vowels, and less and often negative values for a tense and creaky vowel (Stevens and Hanson 1995; Blankenship 2002; DiCiano 2009; Wayland and Jongman 2002; Esposito 2010a, 2010b; Esposito and Khan 2020). All the acoustic measures considered in this study indicate a higher and positive value for the vowels associated with low tone. On the other hand, the high tone vowels display consistent low values for all the spectral balance and tilt measurements and higher CPP values. This apparent reversal makes us wonder what had led the (apparently breathy) vowels associated with high tone to be less breathy (or move towards modal continuum) at present?

One of the possible answers to this question could be to recall that the loss of the feature (+spread glottis) led to a higher pitch in Sylheti. Many researchers have discussed the relationship between tonogenesis and consonant types (Hombert 1978; Ladefoged 1971; Thurgood 2002; Gordon and Ladefoged 2001; Gordon 2001; Kingston 2011; and so forth). In this regard, Hombert (1978) argued that an intended feature of a sound might have intrinsic effects on neighboring sounds. Thus, in the historical development of a language, when a particular feature of a specific sound is lost, the speakers tend to reinterpret (or instead compensate) the lost feature (or the original intended feature) as an intrinsic feature in the neighboring sound. When the intrinsic features are reinterpreted as the main features and the original feature is lost, the result is likely to be the usual source of tonogenesis (Hombert 1978; Bhaskararao 1998).

To elucidate the mechanism of 'sound change,' Ohala (1993a, 1993b) proposed two phonetically motivated mechanisms- *hypo-correction* and *hyper-correction*. Sound changes related to hypo-correction such as the regressive assimilation (Ohala 1990a, 1990b) or tonogenesis, where tonal contrast surfaces due to prior and gradual loss of voicing contrast, occur when a listener fails to normalize the perturbations in the speech signal and ends up taking/judging the signal at face value. This process might lead to a different conceptualization of that speech signal by the listener from the intended (original) one. Hyper-corrective sound change (such as labialization, retroflexion, and glottalization), on the other hand, refers to contexts where a listener not only fails to normalize the actual deviation caused during the production of a particular speech sound but may also over-analyze and erroneously attributes an intended phonetic cue as contextual (Ohala 1990a, 1993a).

In this context, let us recall the process of tonal development in Punjabi due to the apparent loss of breathy voiced consonants. In Punjabi, the [C^hV] sequence (voiced aspirate + vowel) became voiceless unaspirated, leaving a low tone on the vowel. Notice, a two-way loss of the breathy voiced consonants is compensated in the neighboring vowels, resulting in a falling tone (Bhatia 1975; Bhaskararao 1998; Hombert et al. 1979). Other studies have also shown that breathy voiced consonants are stronger f₀ depressor (Ladefoged 1971; Kingston 2011). However, the depression of f₀ following the loss of breathy voiced consonant is hardly common, let alone a universal phenomenon.

To account for a similar kind of tonal phenomena, i.e., the emergence of a high tone following the (loss of the) feature (+spread glottis) in Nakhorn Sithammarat Thai, Kingston (2011) distinguished the features (spread glottis) and (constricted glottis). The higher tones in Yung-Chiang Kam have developed after [constricted glottis] rather than after (spread glottis) consonants. On the other hand, the reflexes after (spread glottis) consonants have

been realized with a higher pitch in Nakhorn Sithammarat Thai. These features are termed *constricted high* and *spread high* ‘splits’ (Kingston 2011). Since both (spread glottis) and (constricted glottis) consonants can either raise or lower the f_0 of the adjacent vowel, it is left with the native speakers of the individual language, which feature to retain. The development of high tone as a reflection of glottal constriction (or due to a process of glottalization) is reported in many Athabaskan languages (Kingston 2011; Gordon and Ladefoged 2001).

Under the purview of the above discussion, a two-fold evolution of Sylheti tonogenesis is proposed. First, the historical development of the loss of intended feature (+spread glottis) associated with breathy voiced obstruents is reinterpreted and readjusted with a perturbed f_0 in the following vowels to maintain the lexical distinctions among the likely homophonous words. In the second stage, the vowels following the aspirated obstruent consonants might have been compensated with a perturbed f_0 . In this stage, the native speakers of Sylheti possibly failed to normalize the coarticulatory effects such as the effects of (+spread glottis) on f_0 as an attribute to feature *hypo-correction*¹¹. Since (+spread glottis) quality of the consonants was reinterpreted on the adjacent vowels, the vowels might have acquired a property similar to modal (or in the continuum of modal to tense) to maintain the lexical distinction of (otherwise likely) homophonous pairs, with an increased f_0 . Once the conditioning environment (for example, the feature (+spread glottis)) was completely lost, the f_0 patterns were phonologized and perceived as a high tone by the native speakers.

The second important issue is categorizing the voice qualities associated with the contrastive tones observed in this study. The consistent low (and yet positive) values associated with the vowels carrying high tone generated through all the spectral balance and tilt measurements suggest that the high tone vowels are in the continuum of *modal to tense*. On the other hand, the significantly higher (and positive) spectral balance and tilt values associated with the vowels carrying low tones confirm that these vowels are in the continuum of *breathy to lax* phonation. Let us recall that the classifications of tense and creaky and lax and breathy are somewhat ambiguous. Gobl and Chasaide (2012) argued that tense and lax phonations are similar in producing creaky and breathy phonations. However, Gordon and Ladefoged (2001) argued that the tense and the lax phonations are often considered similar to modal phonation (than either creaky or breathy) even though they contrast instead of modal phonation in a few languages. Since all the acoustic measures considered in this study generate mostly positive yet significantly different values for both the vowels associated with high and low tones, this study concludes that the vowels associated with high tones are mostly modal (or in the continuum *modal to tense*). The vowels associated with low tones are breathy (or are in the continuum of *breathy to lax*).

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/languages6040154/s1>.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the non-involvement of any animal data.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data is available with the author and can be made available for valid reasons.

Conflicts of Interest: The author declares no conflict of interest.

Notes

- 1 The current paper is a revised and updated version of the paper ‘Correlation between Sylheti tone and Phonation’ presented in *Speech Prosody 2016* (Gope and Mahanta 2016a). The content of this paper is a rework that is extracted from my *Ph.D. thesis The Phonetics and Phonology of Sylheti Tonogenesis*. In this paper, I have primarily concentrated on the (mono-syllabic) data discussed in Gope and Mahanta (2016a). Most of these word-pair that are specified for high tone at present, contained the feature (–voice, +spread glottis). It is quite evident that, the loss of the feature (+spread glottis) in both the voiced and voiceless series led to the development of a high tone in Sylheti. In this paper, however, I did not include voiceless series (only one pair is included) primarily due to lack of sufficient number of tokens. See Gope (2016) for more discussion on how the loss of aspiration in both voiced and voiceless obstruents led to a high tone in Sylheti.
- 2 The loss of aspiration reported in this paper is a case of historical development and not an instance of the synchronic allophonic process.
- 3 Gope (2016) first claimed that the tonogenesis in Sylheti is indeed triggered by the loss of underlying aspiration in the obstruents; the process of fricativization and deaffrication of the obstruents was just a simultaneous process and did not have much role in the formation of tonal contrasts in this language. The study by Raychoudhury and Mahanta (2020) further confirmed this claim. Note that, most of the high tone words considered in this study contained the feature (–voice, +aspiration).
- 4 Aspirated consonants are defined as (+spread glottis) in this paper. See Ohala (1993a) for a detailed discussion on this feature.
- 5 Words of other shapes, including fillers, were also recorded, but not analyzed for the current study for the sake of increased quantitative comparability
- 6 Originally I recorded 9 native speakers (6 male, 3 female). However, data recorded from three older speakers were not analyzed here due to possible generational differences in tonogenesis.
- 7 For example, to initiate the meaning [bát] ‘rice’, an example sentence ‘I eat rice’ [àmì bát xáí. ár] is recorded first.
- 8 Mean f_0 is averaged across all the tokens produced by each subject.
- 9 In this study, a p -value of ≤ 0.05 is considered to be significant.
- 10 Note that, most of the high tone words considered in this study contained the feature (–voice, +aspiration).
- 11 See Ohala (1993a, 1993b) for a detailed discussion on this feature.

References

- Andruski, Jean A., and Martha Ratliff. 2000. Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association* 30: 37–62. [CrossRef]
- Belotel-Grenié, Agnès, and Michel Grenié. 1994. Phonation types analysis in Standard Chinese. Paper presented at 3rd International Conference on Spoken Language Processing, Yokohama, Japan, September 18–22.
- Berkson, Kelly Harper. 2019. Acoustic correlates of breathy sonorants in Marathi. *Journal of Phonetics* 73: 70–90. [CrossRef]
- Bhaskararao, Peri. 1998. Voiced Aspiration and Tonogenesis in Some South-Asian Languages. In *Symposium on Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*. Edited by Kaji Shigeki. Tokyo: ILCAA, pp. 337–51.
- Bhatia, Tej K. 1975. The Evolution of Tones in Punjabi. *Studies in the Linguistic Sciences* 5: 12–24.
- Blankenship, Barbara. 1997. The time course of breathiness and laryngealization in Vowels. Ph.D. dissertation, University of California, Los Angeles, CA, USA.
- Blankenship, Barbara. 2002. The Timing of Non-modal Phonation in Vowels. *Journal of Phonetics* 30: 163–91. [CrossRef]
- Chafe, Wallace. 1977. Accent and related phenomena in the Five Nations Iroquois languages. In *Studies in Stress and Accent*. Edited by Larry Hyman. Los Angeles: University of Southern California Press, pp. 169–81.
- DiCanio, Christian. 2009. The Phonetics of Register in Takhian Thong Chong. *Journal of International Phonetic Association* 39: 162–88. [CrossRef]
- Dixit, R. Prakash. 1989. Glottal Gestures in Hindi Plosives. *Journal of Phonetics* 17: 213–37. [CrossRef]
- Doherty, Brian. 1993. The acoustic-phonetic correlates of Cayuga word-stress. Ph.D. dissertation, Harvard University, Cambridge, MA, USA.
- Dutta, Indranil. 2007. Four-way stop contrasts in hindi: An acoustic study of voicing, fundamental frequency and spectral tilt. Ph.D. thesis, University of Illinois, Urbana, IL, USA.
- England, Nora C. 1983. *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.
- Esposito, Christina M. 2010a. The effects of linguistic experience on the perception of phonation. *Journal of Phonetics* 38: 306–16. [CrossRef]
- Esposito, Christina M. 2010b. Variation in Contrastive Phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association* 40: 181–98. [CrossRef]
- Esposito, Christina M. 2012. An Acoustic and Electroglottographic Study of White Hmong Phonation. *Journal of Phonetics* 40: 466–76. [CrossRef]
- Esposito, Christina M., and Sameer ud Dowla Khan. 2012. Contrastive Breathiness Across Consonants and Vowels: A comparative study of Gujarati and White Hmong. *Journal of the International Phonetic Association* 42: 123–43. [CrossRef]

- Esposito, Christina M., and Sameer ud Dowla Khan. 2020. The cross-linguistic patterns of phonation types. *Lang Linguist Compass* 14: e12392. [CrossRef]
- Esposito, Christina M., Kelly Harper Berkson, and Nelson Max. 2019. Distinguishing breathy consonants and vowels in Gujarati. *Journal of South Asian Languages and Linguistics* 6: 214–43. [CrossRef]
- Fischer-Jorgensen, Eli. 1967. Phonetic Analysis of Breathy (murmured) Vowels. *Indian Linguistics* 28: 71–139.
- Garellek, Marc, and Patricia A. Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41: 185–205. [CrossRef]
- Gobl, Christer, and Ailbhe Ní Chasaide. 2012. Voice source variation and its communicative functions. In *The Handbook of Phonetic Sciences*. Oxford: Wiley-Blackwell, pp. 378–423.
- Gope, Amalesh. 2016. The Phonetics and Phonology of Sylheti Tonogenesis. Ph.D. dissertation, IIT Guwahati, Assam, India.
- Gope, Amalesh. 2018. The Phoneme Inventory of Sylheti: Acoustic Evidences. *Journal of Advanced Linguistic Studies* 7: 7–37.
- Gope, Amalesh, and Shakuntala Mahanta. 2014. Lexical Tone in Sylheti. Paper presented at 4th International Symposium on Tonal Aspects of Languages, Nijmegen, The Netherlands, May 13–16; pp. 10–14.
- Gope, Amalesh, and Shakuntala Mahanta. 2015. An Acoustic Analysis of Sylheti Phonemes. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, UK, August 10–14.
- Gope, Amalesh, and Shakuntala Mahanta. 2016a. Correlation between Sylheti tone and phonation. Paper presented at Speech Prosody 2016, Boston, MA, USA, May 31–June 3.
- Gope, Amalesh, and Shakuntala Mahanta. 2016b. Perception of Lexical Tones in Sylheti. Paper presented at TAL-2016, Buffalo, NY, USA, May 24–27.
- Gordon, Mathew. 2001. Linguistic aspects of voice quality with special reference to Athabaskan. Paper presented at 2001 Athabaskan Languages Conference, Los Angeles, CA, USA, May 18–20.
- Gordon, Mathew, and Peter Ladefoged. 2001. Phonation Types: A Cross-Linguistic Overview. *Journal of Phonetics* 29: 383–406. [CrossRef]
- Hillenbrand, James, Ronald A. Cleveland, and Robert L. Erickson. 1994. Acoustic correlates of breathy voice quality. *Journal of Speech and Hearing Research* 37: 769–78. [CrossRef] [PubMed]
- Hombert, Jean-Marie. 1978. Consonant types, vowel height, and tone in Yoruba. In *Tone: A Linguistic Survey*. Edited by Victoria A. Fromkin. New York: Academic Press, pp. 77–111.
- Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. Phonetic Explanations for the Development of Tones. *Language* 55: 37–58. [CrossRef]
- Huffman, Marie. K. 1987. Measures of phonation type in Hmong. *Journal of the Acoustical Society of America* 81: 495–504. [CrossRef]
- Hyman, Larry M. 2010. Does tone have features. In *Tones and Features: Phonetic and Phonological Perspectives*. Berlin: Walter de Gruyter, pp. 50–80.
- Iseli, Markus, Yen-Lian Shue, and Abeer Alwan. 2007. Age, sex, and vowel dependencies of Acoustic measures related to the voice source. *Journal of the Acoustical Society of America* 121: 2283–95. [CrossRef]
- Jones, Ted E., and Lyle M. Knudson. 1977. Guelavia Zapotec phonemes. In *Studies in OtoManguanean Phonology*. Edited by William Merrifield. Dallas: SIL/University of Texas, Arlington, pp. 163–80.
- Kawahara, Hideki, Alain de Cheveigné, and Roy D. Patterson. 1998. An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: Revised TEMPO in the STRAIGHT-suite. Paper presented at 5th International Conference on Spoken Language Processing, Sydney Convention Centre, Sydney, Australia, November 30–December 4.
- Keating, Patricia, Christina M. Esposito, Marc Garellek, Sameer ud Dowla Khan, and Jianjing Kuang. 2010. Phonation contrasts across languages. *UCLA Working Papers in Phonetics* 108: 188–202.
- Kelterer, Anneliese, and Barbara Schuppler. 2020. Phonation type contrasts and tone in Chichimec. *Journal of Acoustic Society of America* 147: 3043–59. [CrossRef]
- Khan, Sameer ud Dowla. 2010. Bengali (Bangladeshi Standard). *Journal of the International Phonetic Association* 40: 221–25. [CrossRef]
- Khan, Sameer ud Dowla. 2012. The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* 40: 780–95. [CrossRef]
- Kingston, John. 2011. Tonogenesis. In *The Blackwell Companion to Phonology*. Edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume and Keren Rice. Malden: Blackwell, vol. IV.
- Kirby, James P. 2014. Incipient tonogenesis in Phnom Penh Khmer: Acoustic and perceptual studies. *Journal of Phonetics* 43: 69–85. [CrossRef]
- Kirk, Paul L., Jenny Ladefoged, and Peter Ladefoged. 1993. Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In *American Indian Linguistics and Ethnography in Honor of Lawrence C. Thompson*. Edited by Anthony Mattina and Timothy Montler. Ann Arbor: University of Michigan, pp. 435–50.
- Kuang, Jianjing, and Patricia Keating. 2013. Glottal articulations in tense vs. lax phonation contrasts. *Journal of the Acoustical Society of America* 134: 4069–69. [CrossRef]
- Kuang, Jianjing, and Patricia Keating. 2014. Vocal fold vibratory patterns in tense versus lax phonation contrasts. *Journal of Acoustical Society of America* 136: 2784–97. [CrossRef]
- Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press, pp. 1–123.
- Ladefoged, Peter. 1983. The linguistic use of different phonation types. In *Vocal Fold Physiology: Contemporary Research and Clinical Issues*. Edited by Diane Bless and James Abba. San Diego: College-Hill Press, pp. 351–60.

- Ladefoged, Peter, Ian Maddieson, and Michael Jackson. 1988. Investigating phonation types in different languages. In *Vocal Fold Physiology: Voice Production, Mechanisms and Functions*. Edited by O. Fujimura. New York: Raven Press, pp. 297–317.
- Leer, Jeef. 1979. *Proto-Athabaskan Verb Stem Variation, Part One: Phonology*. Research Papers 1. Fairbanks: Alaska Native Language Center, p. 100.
- Maddieson, Ian, and Susan Hess. 1986. 'Tense' and 'Lax' revisited: More on phonation type and pitch in minority languages in China. *UCLA Working Papers in Phonetics* 63: 103–9.
- Michelson, Karin. 1988. *A Comparative Study of Lake-Iroquoian Accent*. Dordrecht: Kluwer.
- Miller, Amanda L. 2007. Guttural vowels and guttural co-articulation in Juj'hoansi. *Journal of Phonetics* 35: 56–84. [CrossRef]
- Mooshammer, Christine. 2010. Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German. *Journal of the Acoustical Society of America* 127: 1047–58. [CrossRef]
- Ohala, John J. 1990a. Respiratory activity in speech. In *Speech Production and Speech Modelling*. Edited by W. J. Hardcastle and A. Marchal. Dordrecht: Kluwer, pp. 23–53.
- Ohala, John J. 1990b. The phonetics and phonology of aspects of assimilation. [And: A response to Pierrehumbert's commentary]. In *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Edited by J. Kingston and M. Beckman. Cambridge: Cambridge University Press, pp. 258–75, 280–82.
- Ohala, John J. 1993a. The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*. Edited by C. Jones. London: Longman Academics, pp. 237–78.
- Ohala, John J. 1993b. Sound change as nature's speech perception experiment. *Speech Communication* 13: 155–61. [CrossRef]
- Paul, Boersma, and David Weenink. 2008. Praat: Doing Phonetics by Computer (Version 5.3.04_win32) [Computer Program]. Available online: <http://www.praat.org/Bradley> (accessed on 22 December 2011).
- Raychoudhury, Priti, and Shakuntala Mahanta. 2020. The Three Way Tonal System of Sylheti. Paper presented at 10th International Conference on Speech Prosody, Tokyo, Japan, May 25–28; pp. 503–7.
- Shue, Yen-Liang, Patricia A. Keating, and Chad Vicenik. 2009. VoiceSauce: A program for voice analysis. *The Journal of the Acoustical Society America* 124: 2221. [CrossRef]
- Silverman, Daniel, Barbara Blankenship, Paul L. Kirk, and Peter Ladefoged. 1995. Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics* 37: 70–88.
- Stevens, Kenneth N., and Helen M. Hanson. 1995. Classification of glottal vibration from acoustic measurements. In *Vocal Fold Physiology: Voice Quality Control*. Edited by O. Fujimura and H. Minoru. San Diego: Singular Publishing Group.
- Thurgood, Graham. 2002. Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica* 19: 333–63. [CrossRef]
- Truill, Anthony. 1985. *Phonetic and Phonological Studies of !Xóõ Bushman*. Hamburg: Buske.
- Watkins, Justin. 2002. *The Phonetics of Wa: Experimental Phonetics, Phonology, Orthography and Sociolinguistics*. Canberra: Pacific Linguistics.
- Wayland, Ratre, and Allard Jongman. 2002. Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics* 31: 181–201. [CrossRef]
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.