

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/329365806>

The Phoneme Inventory of Sylheti: Acoustic Evidences *

Article · December 2018

CITATIONS

4

READS

720

1 author:



[Amalesh Gope](#)

Tezpur University

31 PUBLICATIONS 81 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Bio-Colloidal Drying Drops [View project](#)



Sylheti [View project](#)

The Phoneme Inventory of Sylheti: Acoustic Evidences*

AMALESH GOPE

Department of EFL, Tezpur University

ABSTRACT

The physical constraints of articulatory and/or auditory speech mechanism are vital to the sound change of human languages (Hombert 1977; Ohala 1974, 1993). These constraints often affect the way speech sounds are produced and perceived by listeners. In this paper, with the help of acoustic experiment(s) the phoneme inventory of a relatively less explored language i.e. Sylheti is examined and established. The phoneme inventory of this language is substantially reduced due to the loss of (underlying) breathiness contrast [+spread glottis]: both voiced and voiceless aspirated stops became unaspirated ([q^han > dan] “paddy,” [t^hala > tala] “plate”), and the voiceless labial and dorsal stops went one step further and spirantized to homorganic fricatives ([por] > [ɸɔr] “read,” [p^hul] > [ɸul] “flower,” ([kali] > [xali] “ink,” [k^hal] > [xal] “drain/channel”). The post-alveolar affricates (both aspirated and unaspirated) also spirantized to dental fricatives ([tʃa] > [sa] “tea,” [tʃ^huti] > [suti] “holiday,” [dʒal] > [zal] “net,” [dʒ^hal] > [zal] “spicy”). Among the voiceless stops, only the dental and retroflex stops remained stops. Two controlled production experiments have been conducted to establish the phoneme inventory of Sylheti. The first experiment is meant to explore the status of the underlying breathy voice property

* Part of this paper has been presented in the 18th International Congress of Phonetic Sciences, 10-14 August, 2015, Glasgow, Scotland, UK. Subsequently, a tiny version of this paper was published in the *Proceedings of ICPHS 2015*.

[+spread glottis] of Sylheti obstruents in terms of durational measurements (VOT). The result of one-way ANOVA did not show any significant interaction among the obstruents in terms of aspiration ($p > 0.05$, $[F(1, 359) = 0.095, p = 0.76]$). The second production experiment is conducted to understand and determine total number of Sylheti vowels. A one way ANOVA confirmed significance effect on vowel quality in terms of duration $[F(4, 600) = 57.77, p = 0.00]$. Further a significant interaction is also noticed between vowel types and the (first three) formants values (examined separately for male and female speakers): female speakers: $F1: p < 0.05 [F(4, 300) = 83.89, p = 0.000]$, $F2: p < 0.05 [F(4, 300) = 17.38, p = 0.000]$; and for female speakers: $F1: p < 0.05 [F(4, 300) = 233.37, p = 0.000]$, $F2 (p < 0.05 [F(4, 300) = 87.62, p = 0.000]$.

Keywords: Sylheti, [+spread glottis], lenition, obstruent, VOT, ANOVA, formants.

1. INTRODUCTION

Any human language system is considered as a system comprising of a set of [minimal] meaningful sound units termed as phonemes. Human languages of the world exploit different phonemes (consonants and vowels forming different minimal pairs) to distinguish the lexical meanings of words. Thus, the Standard Colloquial Bangla word [kal] “tomorrow” is different from [k^hal] “canal/drain” or [gal] “cheek” since the first consonant of each word [k], [k^h] and [g] are different. Similarly, the word [kal] “tomorrow” is also different from the word [kil] “a fist,” since their vowels [a] and [i] are different from each other. Whereas such minimal pairs of words are universally present in all human languages, the number of phonemes (consonants, vowels, semi-vowels and diphthongs) used to employ lexical differentiation varies from one language to the other and are subjected to change. The physical constraints of articulatory and/or auditory speech mechanism are vital to sound change of human language (Hombert 1977; Ohala 1974, 1993). Researchers (Hombert 1977; Hombert et al. 1979) also suggest that these

constraints may affect the way speech sounds are produced and perceived by listeners. A speaker's pronunciation may get distorted and may not be perceived as intended, thus creating space for inclusion and/or reduction of phoneme inventory. Several factors causing sound change over generations are found in all human speakers. This paper also predominantly examines such sound changes and explores the acoustic properties of the phonemes in a relatively less explored language – Sylheti which is generally regarded as a variety of (eastern) Bangla. Historical developments of this language exhibits considerable reduction of its phoneme inventory largely due to the loss of underlying breathiness contrasts and partially due to the phonological process of lenition. Two controlled production experiments have been conducted to establish the phoneme inventory of Sylheti. Results of these two acoustics experiments have been discussed in this chapter. The first experiment is meant to explore the status of the underlying breathy voice property [+spread glottis] of Sylheti obstruents in terms of durational measurements (VOT). The second production experiment is conducted to understand and determine total number of Sylheti vowels. The process of spirantization and deaffrication is discussed with the help of spectrographic evidence along with adequate examples.

The paper is organized in following manner: Section 2 provides a general background of Bangla and its dialects followed by the demographic and geographic description of Sylheti (Section 2.1), Section 3 describes the acoustic properties of consonants in Sylheti, Section 4 talks about the production experiment of Sylheti Vowels and Section 5 draws the conclusion of this paper.

GENERAL BACKGROUND: BANGLA AND ITS DIALECTS

Genetically Bangla is derived from Indo-Aryan (IA) or the Indic sub branch of the Indo-Iranian branch of the Indo-European (IE) language family (Chatterji 1926). The language is considered as one of the world's densely inhabited languages with more than 170 million first-language speakers across the world (Gordon 2005). Apart from the entire Bangladesh, the Bangla speaking

zone includes parts of Bihar, Orissa, Assam and the whole of West Bengal and Tripura within India. Substantial Bangla inhabited pockets are also found in the Middle-East, Europe and the U.S, north-western Burma (also known as Myanmar), parts of Nepal (in the regions such as Mechi Zone, Jhapa district, Kosi Zone, Morang and Sunsari districts and so on). It is the national language of Bangladesh. In India, Bangla is the second most-spoken language (behind Hindi and Urdu) and is considered as an official language of the state of West Bengal and co-official language of the state of Tripura and the union territory of Andaman and Nicobar Islands.

Naturally, the wide spread of this language over a vast region witnessed various historical developments and thus formed numerous dialects of Bangla. Many of these dialects appear to be unintelligible due to their wide spread over a large continuum. Grierson (1928) while classifying the speech varieties spoken in the province of Bengal divided Bengali into two major branches: the Eastern Branch and the Western Branch (ignoring the geographical and/or political boundary). The western branch can further be divided into three different types, viz. Central Bengali (Standard Bengali: includes the varieties of Kolkata, Haorah in India), Northern Bengali (includes the varieties spoken in East Malda, Koch-Bihar of West Bengal in India and Rajshahi, Dinajpur, Bogra and in Bangladesh), and Western Bengali (includes varieties such as Kharia, Thar, Mal Paharia, Sarki etc). The Eastern Branch on the other hand, could further be divided into four different groups; viz., Eastern Bengali (includes the varieties spoken in districts such as Dhaka, Mymensingh, Comilla, Sylhet, Hajong of Bangladesh and the Cachar region of Assam in India), East-Central Bengali (includes the varieties spoken in Jessore, Khulna, Faridpur districts of Bangladesh), South-Eastern Bengali (varieties include Noakhali, Chittagong, Chakma, Tangchanga etc) and Rajbanshi (varieties spoken in Rongpur districts of Bangladesh, and Siripuria, Jalpaiguri in the state of West Bengal and Goalpara of Assam in India).

Chatterji (1926) however, categorized the varieties of Bangla spoken in the region of greater Bengal into four major clusters: Radha, Varendra, Kamrupa and Vanga. The region of greater

Bengal does not consider the political boundary and thus comprises the varieties of Bangla spoken in entire Bangladesh and Indian states of West Bengal and Tripura, and parts of Assam, Bihar and Jharkhand. Chatterji (1926) clubbed Sylheti under the subcategory of Eastern and South-Eastern Vanga under the Vanga cluster (Figure 1).

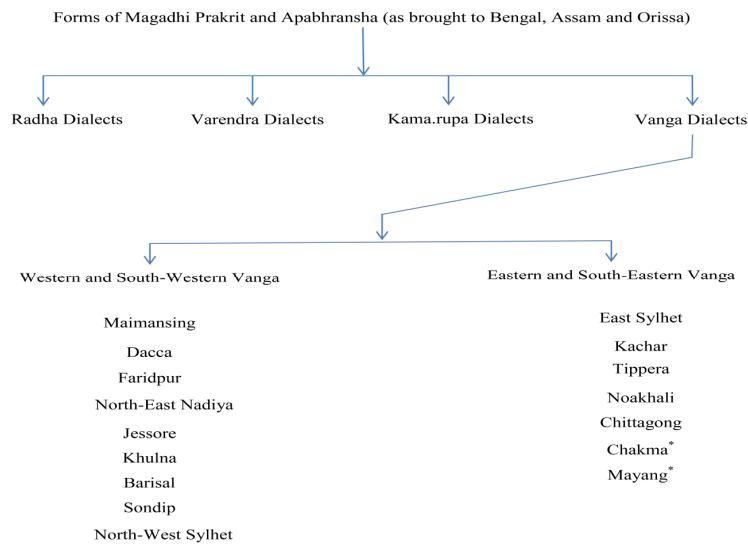


Figure 1. *Varieties of Bangla classified by Chatterji (1926)*

The languages with an asterisk mark (*) (viz., Chakma and Mayang) reported to be mixed up with the languages of Tibeto-Burman group of languages. While Chakma is considered to be a dialect of hill people of Chittagong (in Bangladesh) and parts of Tripura in India, Mayang or Bishnupriya is spoken by a subsection in Manipur (Chatterji 1926). All these dialects differ widely in all aspects of the grammar, especially in terms phoneme inventory, allophony, inflectional morphology and lexicon, thus making many dialects largely unintelligible. For this reason, several dialects such as Chakma, Hajong, Kharia Thar, Mal Paharia, Sylheti, Tangchangya, Noakhaliare often considered to be separate languages (Grierson 1928; Chatterji 1926; Gordon 2005).

2.1. *The language under study: Sylheti*

As shown above (Figure 1), Sylheti is generally regarded as one of the varieties of eastern Bangla. There are approximately 10,300,000 people using Sylheti as their first language (7,000,000 in Bangladesh) (Lewis et al. 2013). The name of this language has been derived from the place where it is being typically spoken; i.e., the Sylhet district (also known as Surma Valley including districts such as Sylhet, Habiganj, Maulvi Bazar, and Sunamganj) located in the north-eastern region of Bangladesh. In India, a considerable amount of Sylheti speakers inhabit the northeast state of Assam (the Barak Valley region comprising districts such as Cachar, Karimganj and Hailakandi) and Tripura (mostly the northern districts such as Dharmanagar, Kailasahar and Kumarghat). At present Sylheti exhibits two divisions; mostly due to the speakers' religious affinity. The Islamic followers (especially those from Bangladesh and Karimganj and Hailakandi districts of Assam in India) use a lot of borrowed words and phrases from Persian and Arabic (including many of the kinship terms). Sylheti used by the Hindu religion followers of Cachar district of Assam and North Tripura has somehow kept their original form intact. Data for this dissertation were collected mostly from the Dharmanagar district of north Tripura and partially from the Cachar district of Assam.

Interestingly, Sylheti had a distinct script in Syloti-Nagari (better known as Nagori) interrelated with Kaithi alphabet. This script appears to be very different than the Bangla alphabets. The exact origin of this alphabet of Sylheti is unknown and believed to be developed around the beginning of 14th century (Chatterji 1926). Books written in Nagori (manuscripts or printed) are known as *puthis*. The main purpose of this writing system was to record religious poetry.

The distinct phonetic and phonological property distinguishes this variety from Standard Colloquial Bangla (henceforth SCB). One of the properties that distinguish Sylheti from SCB or other regional varieties is the significant application of obstruent weakening involving de-aspiration (loss of underlying breathiness contrasts [+spread glottis]), and the phonological process of lenition involving spirantization and

deaffrication. Consequently, the consonant inventory (especially the obstruents) of Sylheti exhibit a major reduction and restructuring compared to that of SCB.

This paper discusses the phoneme inventory of Sylheti. To understand the status of the feature [+spread glottis], we conducted an acoustic experiment and measured the durational properties of Sylheti voiced stops using the concept of voice onset time (henceforth VOT), proposed by Liskar and Abramson (1964). I speculate that these voiced stops would differ only in terms of their place of articulation. The lenition process of spirantization and deaffrication is discussed with the help of spectrograms. Further, to provide a comprehensive account of Sylheti vowels, another acoustic experiment on Sylheti monophthongs has been conducted. Various statistical analyses were performed on different acoustic properties such as duration, first three formants of the target vowels to confirm the types of vowels and gender effects. Our findings confirmed that there are five (oral) vowels in Sylheti as compared to seven in SCB.

3. PHONEMES INVENTORY OF SYLHETI: CONSONANTS

Chatterji (1926) recorded twenty nine consonants and fourteen vowels of which seven are nasalised in SCB. The language maintains a four way contrasts of voicing and aspiration among the stops. Of the twenty nine consonants ten are aspirated stops [p^h], [b^h], [t^h], [d^h], [t̪^h], [d̪^h], [k^h], [g^h], [dʒ^h] and [tʃ^h]. The stop consonants also have their unaspirated counter parts. There are six sonorants consisting of three liquids [l, r, ɹ] and three nasals [m, n, ŋ]. As discussed above, Sylheti perceived drastic phonological changes resulting in large scale reduction in the overall number of phonemes and stops in particular. Assuming SCB as the base of roots, I will argue that all the aspirated stops in Sylheti were lost due to the phonological process of obstruent weakening. Das (1996) also showed a similar reduction process of obstruent weakening in Noakhali spoken in Southern Tripura.

Since the obstruents in Sylheti are the most affected ones, I decided to provide a detailed acoustic analysis of obstruent inventory. Findings from several field surveys on Sylheti suggest

the presence of only twelve stops in Sylheti. Let us consider the following examples:

3.1. *Deaspiration of voiced aspirated obstruents*

The entire set of voiced aspirated stops [b^h, d^h, d̪^h, g^h, dʒ^h] is deaspirated in all the positions of occurrences (viz., word initially, medially, as well as word finally) in Sylheti. The loss of the phonological feature [+spread glottis] could be seen in the following examples:

a. [b^h] → [b]

Initial	Gloss	Medial	Gloss	Final	Gloss
[bai]	“brother”	[na.bi]	“naval”	[lab]	“profit”
[bab]	“affection/pretend”	[ɔ.bab]	“poverty”	[lub]	“greed”

b. [d^h] → [d]

Initial	Gloss	Medial	Gloss	Final	Gloss
[dan]	“paddy”	[ga.d̪a]	“donkey”	[d̪ud̪]	“milk”
[dɔn]	“wealth”	[a.d̪a]	“ginger”	[ʃud̪]	“repayment”

c. [d̪^h] → [d̪] [the consonant [d̪^h] occurs only in the initial position of a word]

Initial	Gloss	Medial	Gloss	Final	Gloss
[d̪ax ¹ /	“drum”				
/d̪a.lu/	“sloping”				

d. /[g^h] → [g]

Initial	Gloss	Medial	Gloss	Final	Gloss
[gum]	“sleep”	[mɔ.ga]	“inauspicious star”	[bag]	“tiger”
[gɔr]	“house”	[rɔ.gu]	“name”	[mag]	“name of a Bengali month”

3.2. *Deaspiration of voiceless aspirated obstruents*

Similarly voiceless obstruents also display the loss of the feature [+spread glottis] in Sylheti. The entire set of voiceless aspirated obstruents [p^h, t^h, t̪^h, tʃ^h, k^h] lose the underlying breathiness

property and thus fall together with the already existing set of corresponding sounds with the feature specifications of [-son, +cons, -voice]. Interestingly, the underlying voiceless bilabial stop (both aspirated and unaspirated) [p] and [p^h] and the underlying voiceless velar stop (both aspirated and unaspirated) [k] and [k^h] are the most affected sounds. Not only these sounds lose their underlying breathiness property, these are further spirantized to [ɸ] and [x] respectively. The details of these patterns have been discussed in Section 3.4. Let us consider the following examples:

a. [t^h] → [t]

Initial	Gloss	Medial	Gloss	Final	Gloss
[t̥a.la]	“a plate”	[ma.t̥a]	“head”	[rɔ̥t̥]	“chariot”
[t̥a.ka]	“to stay”	[la.t̥i]	“ick”	[nɔ̥t̥]	“nose-ring”

b. [t^h] → [t]

Initial	Gloss	Medial	Gloss	Final	Gloss
[t̥ik ² /	“exact”	[la.t̥i]	“stick”	[mat̥]	“field”
[t̥a.x ³ ur]	“deity”	[x ⁴ a.t̥al]	“jackfruit”	[x ⁵ at̥]	“wooden piece”

One of the major goal of this paper is to examine how the loss of the feature [+spread glottis] among the Sylheti obstruents have affected the phonological and/or phonetic property of this language. The systematic historical development of this language resulted in a reduced and restructured phoneme inventory therefore requires to be investigated carefully. To examine the properties of stop consonants and to understand the nature of the (underlying) feature [+spread glottis] (such as the outcome of the loss of the underlying feature [+spread glottis] on the following vowels), I have conducted a production experiment and measured the voiced onset time (VOT) of all the voiced stops following the theory proposed by Lisker & Abramson 1976). Since the voiceless stop series, by and large, is the most effected ones (involving two-way loss viz., spirantization and deaspiration [p, p^h → ɸ] and [k, k^h → x]), and also partially because we did not observe any difference in terms of their VOT (voicing lag as proposed by Lisker & Abramson 1976) from our pilot study

between the pairs of underlying aspirated and non-aspirated series of the two voiceless stops [t̪] and [t], we decided to concentrate on the voiced series instead.

3.3. *Acoustic analysis of stop consonants*

To understand how the remnant feature [+spread glottis] may have affected consonants of Sylheti, I have measured the voice onset time of each of the voiced stops present in the language. Abramson (1977) argued that voice onset time (VOT) could be considered as the most reliable acoustic cue to be able to draw a distinction between the voicing categories (+voice versus -voice). In general, the presence or absence of glottal buzz during the oral closure is assumed to be the relevant indicator of voicing. For example, in a language like German, voiced stops in initial prevocalic and prestressed position exhibit a relatively short VOT (negative) and show no or little presence of aspiration, while voiceless stops have a relatively long VOT (positive) and are indicative of the presence of strong aspiration (Lisker 1986; Docherty 1992; Mikuteit 2009). With the help of data from 11 different languages, Lisker & Abramson (1964) claimed that the single dimension of VOT or duration measurement featuring stop consonants (those occurred word-initially) is sufficient enough to be able to distinguish the voicing contrasts in most of those languages.

This temporal characteristic of stop consonants reflects the complex timing of supralaryngeal – laryngeal coordination. VOT therefore, is defined in relation to the time difference between the release of closure for a stop consonant and the onset of vocal fold vibration. Voicing detected before the release or during stop occlusion is called the voicing lead, while voicing starting after the release is called voicing lag. VOT thus stands for the temporal relation between the onset of glottal pulsing and the release of the initial stop consonant, thus making the stop release as the measuring reference point (Lisker & Abramson 1964, 1967). A negative value is attributed to the VOT [+voice] if the onset of voicing is perceived before the release and a positive value is assigned if the voicing onset takes place after the release [-voice]. Thus it is generally assumed that, the voiceless

unaspirated stops is attributed the highest and positive VOT, their aspirated counter parts will have a relatively less and positive VOT, and the voiced unaspirated stops are associated with negative and relatively shorter VOT, and the voiced aspirated do have the lowest and negative VOT.

While the postulations derived from the VOT theory is capable of successfully distinguishing the voicing contrast and the contrast between voiceless aspirated stops with their unaspirated counterparts (thus, appears to be three dimensional), the theory is insufficient in accounting for the voiced aspirates (Lisker & Abramson 1964; Lombardi 1994; Dutta 2007; Mikuteit 2009). This constraint is surfaced due to the fact that the clue “voicing lead” (-VOT) employed for both voiced aspirated stops and voiced stops appears to be inadequate for making a distinction between these two stop types. To overcome this problem, Lisker & Abramson (1964) proposed to consider the low amplitude buzz combined with noise in the interval following the release as the vital factor to distinguish these two voiced categories. In my data of Sylheti voiced phonemes (with and without (underlying) breathiness property), I did not observe any low amplitude buzz and/or noise like signal.

Davis (1994), however, claimed that the lag time differences of “noise offset” measurements (measurement taken from the release burst to the onset of the second formant), of all four velar stops in Hindi ([k, k^h, g, g^h]) is capable of successfully distinguishing the voiced unaspirated stops from the voiced stops. In her studies she observed a significantly longer lag time for voiced stops than their voiced aspirated counterparts. On the other hand, many researchers such as Schiefer (1986), Lombardi (1994), Dixit (1989), Yadav (1984), Ladefoged (1971), Dutta (2007) debated on the inadequacy of the VOT theory especially for the voiced aspirated stops and suggested additional acoustic measurement techniques for successfully distinguishing the voiced aspirated stops from their unaspirated counterparts. While on one hand, Schiefer (1992) proposes the notion of “Voicing Lead Time” (to be measured depending on the period of voicing during closure), Ladefoged (1971) on the other hand, proposed the concept of independent mode of relationship between the

voiced aspirated stops and phonation types (such as breathy, modal and murmured) of the adjacent vowels. Dutta (2007) argued that the four way stop distinction (voicing [+/-], and aspiration [+/-]) in Hindi stops could well be analysed as a cumulative effect of several acoustic cues such as the effect of f_0 perturbation (of neighbouring vowel) following the stops. While analysing the stop categories of Hindi, he observed that both voiced aspirated stops and their unaspirated counterparts are stronger f_0 depressor⁶; voiced aspirated stops seem to be lowering the f_0 even further. Keeping in mind all the theoretical issues discussed above, I have conducted a production experiment on all the possible voiced stops in Sylheti as discussed above.

3.3.1. *Experimental procedure*

3.3.1.1. *Stimuli*

For the acoustic experiment, 18 different words representing each of the target voiced stops occurring word initially were carefully chosen (Table 1). As could be seen from the dataset given in the Table 1, the bilabial voiced stop [b] is represented word-initially in 4 different words (viz., [baṭ] “arthritis,” [ban] “tie,” [bala] “bracelet” and [bari] “home”). The contrastive pairs (viz., [baṭ] “rice,” [ban] “pretend,” [bala] “good” and [bari] “heavy”) representing the (underlying) aspirated voiced stop [b^h] had underlying breathiness property at some point of history, and appeared to be distinct phonemes. The vowel following the target stop was an [a] and form (near) minimal pairs. Thus, the dental voiced stop [ḍ] (along with the underlying aspirated counterpart) is represented in two words, while the retroflex voiced stop [ɖ] and the velar voiced stop [g] (along with their aspirated counterparts) are represented in 4 different words each. My major goal was to measure the VOT of each of the target stops and compare the intrinsic phonetic variations involving place of articulations and aspiration.

Table 2. *Dataset considered for VOT experiment*

Sylheti words	Gloss	Sylheti words with a history of underlying aspiration	Gloss
[baṭ]	‘arthritis’	[baṭ]	‘rice’
[ban]	‘tie’	[ban]	‘pretend’
[bala]	‘bracelet’	[bala]	‘good’
[bari]	‘home’	[bari]	‘heavy’
[ḡan]	‘donate’	[ḡan]	‘paddy’
[dax]	‘roaring of cloud’	[dax]	‘drum’
[dala]	‘tray’	[dala]	‘pour’
[ga]	‘body’	[ga]	‘wound’
[gai]	‘cow’	[gai]	‘stroke’

3.3.1.2. *Participants and recording procedure*

Eight native speakers of Sylheti (six male, two female) were recorded in a quiet environment in Dharmanagar district of North Tripura. Apart from Sylheti, the speakers were also fluent in Hindi and English. The age of the participants ranged between 18 and 42 years. None of speakers had any history of speech disorders.

Speech data was recorded with a Shure unidirectional head-worn microphone connected to a Tascam linear PCM recorder (ensuring a constant mike-to-mouth distance) via xlr jack. The material with target stops was displayed on a computer screen. The meaning of each word was written along with each individual word. Subjects were asked to pronounce each individual word first at normal utterance speech with preferably natural intonation. To avoid the effect of any neighbouring sounds, each word was recorded individually with a considerable amount of gap between two words. The author constantly monitored the whole recording procedure and ensured that each subjects understood the task correctly. Apart from the target stimuli mentioned in Table 1, an additional 20 words were also placed as fillers in the dataset. All the words (along with the target stimuli) were randomized and presented on three different lists, thus ensuring each stimuli was recorded three times each.

3.3.1.3. Acoustic measurement

As mentioned above I did not observe any noise like signal following the release in the interval that was proposed by Lisker & Abramson (1964) as the vital factor to distinguish the two voiced categories of aspirated and unaspirated (Figure 2, and Figure 3). Hence I decided to consider onset of the release till the offset of the voicing burst for the durational measurements of the voiced stop categories. I have used Praat 5.2 (Boersma & Weenick 2013) for all the acoustic measurements.

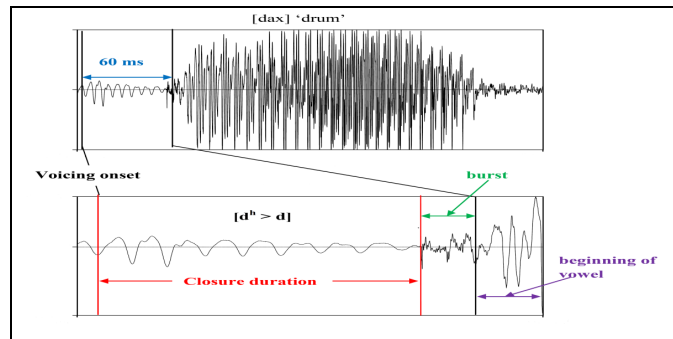


Figure 2. Waveform of Sylheti words [dax] “drum” [dʰ > d]; the enlarged portion of the target phoneme considered for VOT measurement is shown along with

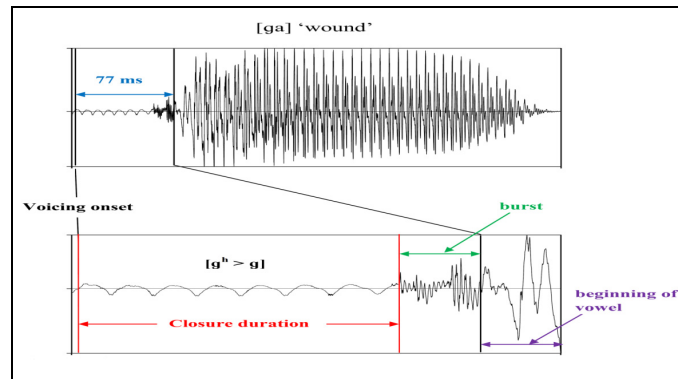


Figure 3. Waveform of Sylheti words [ga] “wound” [gʰ > g]; the enlarged portion of the target phoneme considered for VOT measurement is shown along with

3.3.2. Results and discussion

A one way ANOVA was performed on the durational values calculated for underlying aspirated and unaspirated voiced stops using SPSS software (version 20). Altogether 360 tokens were considered for acoustic and statistical analysis. Few tokens were left out due to either distortion problem arose from external noise, or if the token was disturbed due to individual speaker's starting problem. For the ANOVA test, voicing types were kept as categorical factor (independent factor, with two levels viz.; underlying aspiration and unaspiration) and duration values as variables. As expected the underlying voiced aspirated stops did not show any difference with their unaspirated counterparts (Table 2). The duration values of individual voiced stop are shown in the Figure 4.

Table 2. *Statistical analysis of variance among Sylheti voiced categories*

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	14.314	1	14.314	.095	.757
Within Groups	53665.895	358	149.905		
Total	53680.209	359			

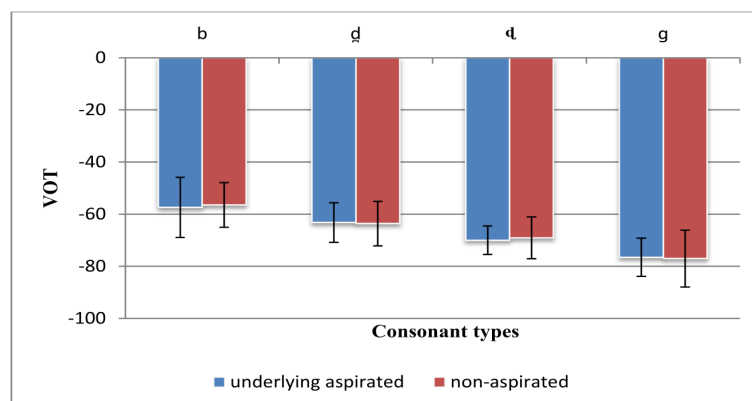


Figure 4. *Average duration of the voiced stop consonants with standard deviation as error bars*

To observe the interaction among the voiced stop categories a subsequent post-hoc Tukey test was also conducted on the data gathered from all the speakers (Table 3). The findings of post-hoc Tukey test revealed that Sylheti stop categories significantly differ from each other in terms of their place of articulation (POA) except for the pair [d̥] and [d] which could be due to less number of tokens considered for the statistical analysis (the consonant [d̥] has only 22 tokens). As assumed no significant interaction was observed between individual pair of the unaspirated and (underlying) aspirated stops. No significant interaction was also observed between [b] and [d̥], however, the underlying aspirated [d̥] and [b] significantly differ from each other where the former was found to be significantly shorter than the later. Similarly, underlying aspirated [b] and unaspirated dental [d̥] and underlying aspirate [d̥^h] is found to be significantly different from each other.

Table 3. *Significant matrix for duration of Sylheti stop consonants are presented; the asterisk [*] symbol indicates the significant while [!] represents the non-significant pairs*

Consonant Types	N	b	b ^h	d̥	d̥ ^h	ɖ	ɖ ^h	g
b ^h	80	!						
d̥	22	!	*					
d̥ ^h	22	*	*	!				
ɖ	36	*	*	!	!			
ɖ ^h	34	*	*	!	!	!		
g	43	*	*	*	*	*	*	
g ^h	43	*	*	*	*	*	*	!

3.4. *The process of lenition*

The lenition process of spirantization and deaffrication (along with deaspiration) further reduced and restructured the phoneme inventory of Sylheti. With the help of waveforms, these changes are discussed below.

3.4.1. *Spirantization*

In Sylheti, the underlying voiceless labial and dorsal stops not only lost their underlying breathiness contrasts but went one step further and spirantized to homorganic fricatives due to the phonological process of consonant weakening. Consider the following examples:

a. [p] → [ɸ]

Initial	Gloss	Medial	Gloss	Final	Gloss
[ɸan]	“betel”	[ru.ɸa]	“silver”	[faɸ]	“snake”
[ɸɔr]	“read”	[aɸ.ɽa]	“week”	[baɸ]	“father”

b. [p^h] → [ɸ]

Initial	Gloss	Medial	Gloss	Final	Gloss
[ɸul]	“flower”	[tu.ɸan]	“storm”	[maɸ]	“forgive”
[ɸɔl]	“fruit”	[bɔr.ɸi]	“plum”	[laɸ]	“leap/jump”

The transformation of these stops to fricatives has been shown in Figure 5 and 6. The enlarged portion of the target phoneme is also shown.

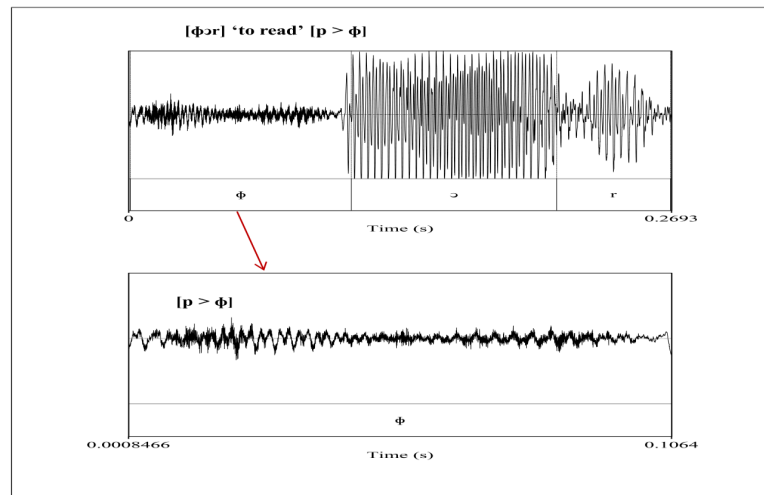


Figure 5. Waveform display of the phoneme [ɸ] as captured in the word [ɸɔr] “read”

c. [k] → [x]

Initial	Gloss	Medial	Gloss	Final	Gloss
[xa.li]	“ink”	[bɛ.xa]	“curve”	[xax]	“crow”
[xo.la]	“banana”	[dʒ.xa]	“calf”	[bux]	“chest”

d. [k^h] → [x]

Initial	Gloss	Medial	Gloss	Final	Gloss
[xa.li]	“empty”	[ma.xon]	“butter”	[mux]	“mouth”
[xup.ri]	“small hut”	[re.xa]	“line”	[dux]	“pain”

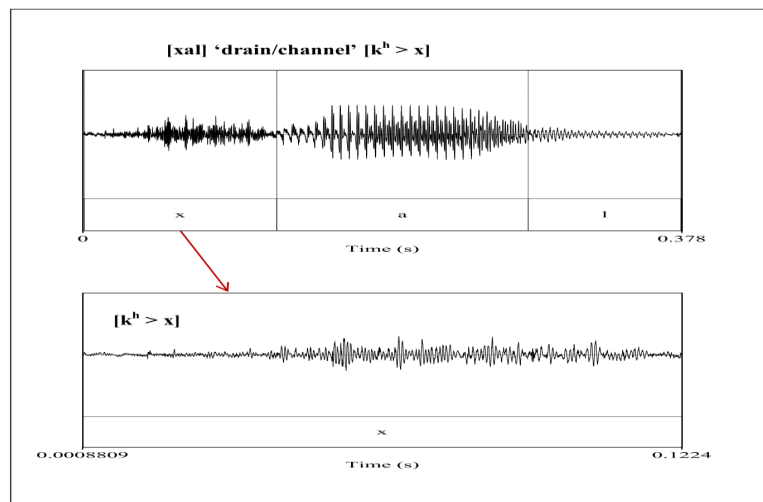


Figure 6. *Waveform display of the phoneme [x] as captured in the word [xal] “drain/channel”. The enlarged portion of the target phoneme is also shown.*

3.4.2. Deaffrication

The alveolar affricates (both aspirated and unaspirated, viz. [tʃ], [tʃ^h], [dʒ] and [dʒ^h]) also exhibit the process of deaffrication and change to alveolar fricatives [s] and [z] respectively. The transformation of these affricates to fricatives has been shown in Figure 7 and 8. The enlarged portion of the target phoneme is also shown. Let us consider the following examples:

a. [tʃ] → [s]

Initial	Gloss	Medial	Gloss	Final	Gloss
[sa]	“tea”	[xa.sa]	“raw”	[xas]	“glass”
[sal]	“rice”	[ʃe.sa]	“owl”	[ʃas]	“five”

b. [tʃ^h] → [s]

Initial	Gloss	Medial	Gloss	Final	Gloss
[sa.ɡi]	“she goat”	[ma.si]	“fly”	[mas]	“fish”
[saɟ]	“roof”	[bis.na]	“bed”	[ɡas]	“tree”

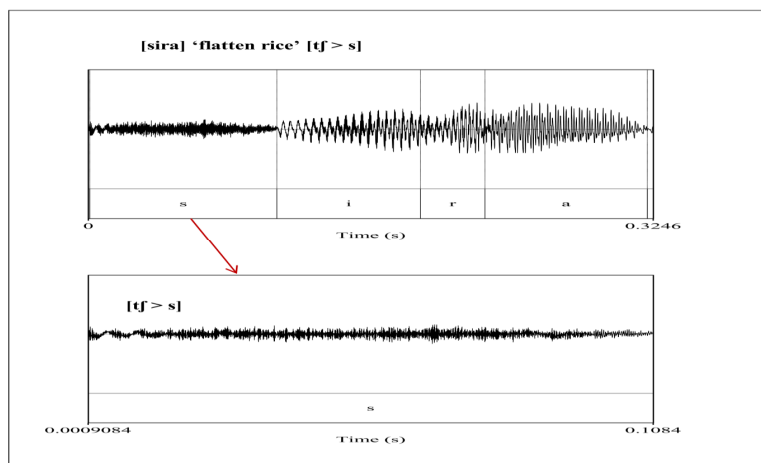


Figure 7. Waveform display of the phoneme [s] as captured in the word [sira] “flatten rice”. The enlarged portion of the target phoneme is also shown.

c. [dʒ] → [z]

Initial	Gloss	Medial	Gloss	Final	Gloss
[zal]	“net”	[baz.na]	“instrument”	[baz]	“thunderbolt”
[zaɟ]	“caste”	[ba.zar]	“market”	[faz]	“make-up”

d. [dʒ^h] → [z]

Initial	Gloss	Medial	Gloss	Final	Gloss
[zal]	“hot”	[ma.zi]	“boatman”	N/A	
[zap.ʃa]	“blur”	[ma.za.ri]	“medium”	N/A	

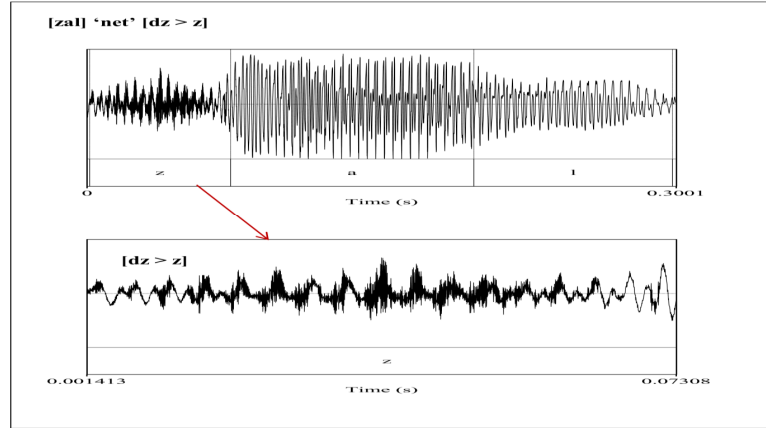


Figure 8. Waveform display of the phoneme [z] as captured in the word [zal] “net”. The enlarged portion of the target phoneme is also shown

I have shown the way consonants especially the stops and affricates in Sylheti exhibit a drastic change and thus resulted in reducing the total number of consonant phonemes. Table 4 represents the consonant inventories of Sylheti.

Table 4. Consonant inventory of Sylheti

Manner of Articulation	Bilabial	Dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Stop/Plosive	b	t̪	ɖ	t̠	ɟ	g	
Nasal	m	n̪				ŋ	
Fricative	ɸ		s	z	f	x	h
Affricate							
Approximant			r				
Lateral			l				

4. VOWELS IN SYLHETI

From my initial fieldwork on Sylheti, I got the impression that not only the consonant inventory but also the number of vowels in Sylheti have got reduced and restructured compared to that of SCB. There are seven oral vowels in SCB (viz., [i], [e], [æ], [a], [u], [o] and [ɔ]) (Chatterji 1926; Bhattacharya 1910). Recordings

from the pilot studies gave me an impression that the half-open front vowel [æ] and half-closed back vowel [o] have been merged with [e] and [u] respectively and thus reduced the number of vowels to five ([i], [ɛ], [a], [ɔ] and [u]). To validate the number of vowels and to provide a detailed acoustic analysis of the vowels I have conducted another production experiment to understand the vowel system of Sylheti. The methodology of the acoustic experiment and the findings are discussed in the subsequent sections.

4.1. Methodology: Stimuli, speakers and recording procedure

After consulting two of my primary informants, a dataset containing all the possible vowels were prepared. Four different consonantal contexts such as [bVl], [sVɸ], [xVl] and [zVl] (V being the target vowel) were prepared as stimuli. All the target vowels (embedded in different consonantal contexts) were recorded using a fixed carrier sentence such as [amicVcxɔiar] ‘‘I cVc said,’’ where cVc is one of the four different consonantal environment and V is the target vowel (see Table 5). Two of the target words (comprising the target vowel) were disyllabic (viz., [bala] ‘‘ornament,’’ and [bulbuli] ‘‘name of a bird’’), however, I considered only the first vowel of those two words for the purpose of acoustic and statistical analysis.

Table 5. Dataset prepared for acoustic analysis of the vowels

Sylheti words	Gloss	Sylheti words	Gloss
[bil]	‘pond’	[kil]	‘fist’
[bɛl]	‘wood-apple’	[xɛ]	‘who’
[bala]	‘bracelet’	[xal]	‘skin’
[bɔl]	‘ball’	[xɔl]	‘tube-well’
[bulbuli]	‘name of a bird’	[kul]	‘lap’
[zil]	‘lake’	[siɸ]	‘press’
[zɛl]	‘prison’	[sɛɸ]	‘spit’
[zala]	‘body’	[saɸ]	‘pressure’
[zɔl]	‘water’	[sɔɸ]	‘a type of snack’
[zul]	‘gravy’	[suɸ]	‘quiet’

All those words (as shown in Table 5) were randomized and additional thirty words were also used as fillers. Subjects were asked to produce those words (embedded in the fixed sentence frame) as naturally as possible. The list was prepared in such a way that each of the target stimuli (along with the fillers) embedded in the fixed sentence frame were repeated thrice.

Seven (5 male, 2 female) out of the eight speakers who participated in the first production experiment (on VOT) participated in the current experiment. Apart from these seven native speakers three more female speaker belonging to the same age group from the Cachar district of Assam, India were also included in this experiment, thus making the total number of speakers to ten (five male, five female). Recordings were done in the same way as mentioned in the section 3.3.1.2. A total of 600 vowel tokens were considered for acoustic and statistical analysis.

4.2. Acoustic and statistical analysis of Sylheti vowels

The start and the end times of each tokens (of the target vowels) were manually labelled using Praat 5.2 (Boersma & Weenick 2013). Labels were placed at zero crossings at the onset consonant till the end of the vowel by ensuring the surrounding speech sounds were not audible in the remaining signal. The frequencies of the first three formants (F1, F2 and F3) of each monophthong were calculated at vowel midpoint to avoid possible onset consonantal effect. A Praat script was written to measure the acoustic components viz. duration and formant values of the vowels. Further, the Hertz values were transformed to Mel using Praat's inbuilt function HertzToMel. The extracted formant values calculated in Mel (in dB) were further normalized for speaker variations using the Lobanov normalization method in NORM (Thomas & Kendell 2007). As speculated, the results confirm the presence of only 5 vowels (monophthongs) in Sylheti (see Figure 10). To confirm the vowel type differences and gender effects, I have conducted one-way Analysis of Variance (ANOVA) tests keeping vowel types/gender as factor and duration and formant values (F1, F2, F3 calculated in Mel) as dependent variables. All the statistical tests were conducted using SPSS (version 20).

4.3. Results and discussion

4.3.1. Duration

Figure 9 shows the average duration of each vowel drawn separately for male and female speakers. Result suggests that on average vowel durations do differ from each other. A one way ANNOVA conducted on the data from all the speakers confirmed significant effect on vowel quality in terms of duration [$F(4, 600) = 57.77, p = 0.00$]. The high vowels [u] (Mean duration = 81.04 milliseconds, $N = 120$) and [i] (Mean duration = 89.77 milliseconds, $N = 120$) appear to be the shortest.

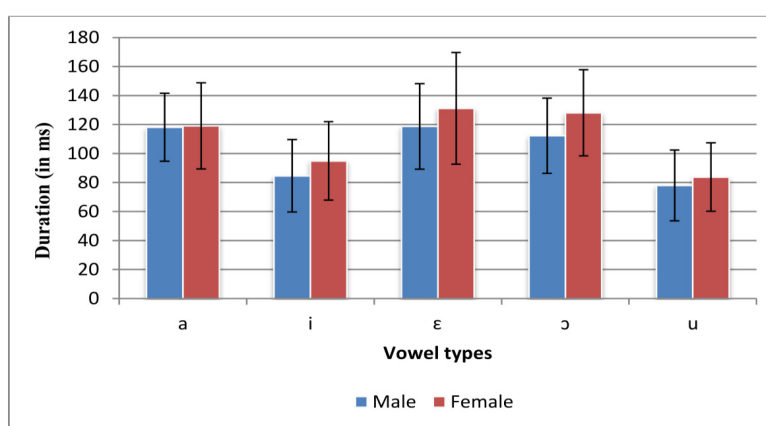


Figure 9. Average duration of Sylheti vowels with standard deviation as error bars

A successive post-hoc Tukey test (Table 6) was also conducted to observe the interaction between individual vowel pairs. The results revealed that the high vowels [i] and [u] seem to differ (significantly shorter) from the remaining three vowels in terms of duration. In all the other cases, vowel durations were not found to be significantly factor (which is fine since the remaining 3 vowels are mid to low vowels). This pattern was observed for both male and female speakers' data; i.e., out of all the 5 vowels, only the high vowels [i and u] were found to be systematically shorter than the remaining three vowels. Significant pairs are (marked with (*), and (!) is used to represent the non-significant pairs).

Table 6. *Significance matrix for duration of all the vowels in Sylheti*

Vowel Types	N	a	i	ɛ	ɔ	u
i	120	*				
ɛ	120	!	*			
ɔ	120	!	*	!		
u	120	*	!	*	*	*

4.3.2. *Formant frequencies*

As mentioned above, the first three formant values of each vowel were calculated at the mid-point and the values were subsequently converted in Mel. Since formant frequencies usually vary across male and female genders (Peterson & Barney 1952), values for the first two formants (F1 and F2) were normalized using Lobanov's (1971) normalization procedure and were plotted on an F1 – F2 plane using NORM (Thomas & Kendell 2007) (Figure 10). Non-normalized F1, F2 and F3 formant frequencies (calculated in Mel) of Sylheti vowels are shown in Table 7 and 8 for male and female speakers respectively.

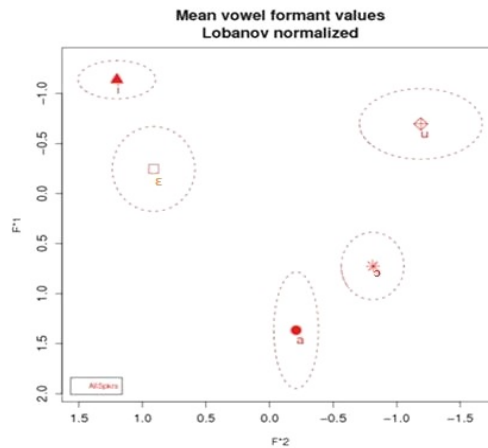


Figure 10. *Vowel diagram showing average Lobanov normalized formant frequencies of the first two formants with one standard deviation ellipses. N= 120 per vowel*

Table 7. *Non-normalized formant values (measured in Mel) of Sylheti male speakers with standard deviations (SD), N=60 for each vowels*

Male	F1 (SD)	F2 (SD)	F3 (SD)
a	531.27 (31.95)	804.62 (49.15)	991.93 (33.23)
i	368.58 (70)	816.06 (36.17)	990.78 (48.47)
ɛ	406.06 (21.56)	819 (45)	1004.49 (24.64)
ɔ	427.6 (35.07)	769.51 (37.75)	972.13 (42.75)
u	366.9 (54.58)	704.81 (39.17)	968.61 (40.79)

Table 8. *Non-normalized formant values (measured in Mel) of Sylheti female speakers with standard deviations (SD), N=60 for each vowels*

Female	F1 (SD)	F2 (SD)	F3 (SD)
a	525.63 (3156)	761.77 (55.79)	973.01 (72.14)
i	334.34 (62.48)	855.38 (34.1)	1029.83 (32.71)
ɛ	419.1 (27.01)	854.02 (63.41)	1021.22 (23.35)
ɔ	433.44 (30.64)	675.94 (70.82)	984.33 (45.17)
u	338.8 (47.70)	653.86 (53.53)	980.18 (47.7)

Figure 10, and the formant values shown in Table 7 and 8 (for male and female speakers respectively) demonstrate the distinctiveness of Sylheti monophthongs. Since Sylheti vowels do not host [±round] distinctions, the third formant values were not considered for statistical analysis. To see how these observed vowels interact in terms of their formant values, I have conducted one way ANOVA tests on F1, and F2 values of each vowel (using vowel types as factor and formant frequencies (F1, F2) as dependant variables). For the statistical tests, I decided to use the raw (non-normalized) data. It was done so as it is assumed that Lobanov's normalization procedure may introduce artefacts when used for measurements based on vowel system that might differ in the overall size and shape (Disner 1980; Clopper et al. 2005; Adank et al. 2007). Therefore, Lobanov's normalization procedure was used to draw the vowel diagram only, and all the statistical tests were done separately for male and female speakers using the non-normalized Mel data. The results of the

ANOVA test confirmed a strong interaction between vowel types based on their formant values (for both the gender). For all the formants analysed in this study, I have noticed a significant interaction among the vowels (for male speakers: F1: $p < 0.05$ [($F(4, 300) = 83.89, p = 0.000$), F2: $p < 0.05$ [($F(4, 300) = 17.38, p = 0.000$)]); and for female speakers: F1: $p < 0.05$ [($F(4, 300) = 233.37, p = 0.000$), F2 ($p < 0.05$ [($F(4, 300) = 87.62, p = 0.000$)]).

A subsequent post-hoc Tukey test was conducted to observe systematic interaction among the vowel pairs based on formant frequencies. The results revealed that in case of F1, all the vowel pairs differ significantly from each other except for the pairs [i] and [u] and [ɛ] and [ɔ]. The similar pattern was observed for both the gender which suggest that in terms of vowel height, Sylheti vowels significantly differ from each other (except for the pairs [i] and [u] (which are expected to be of similar height, and hence did not differ), and [ɛ] and [ɔ].

In case of F2, I noticed consistent patterns of interaction in female data. Except the pair of [i] and [ɛ] all the remaining vowels do seem to differ from each other significantly. Since, F2 denotes the backness property of a vowel; I conclude that all the vowels are significantly different in terms of their backness property. However, the male data did not show a consistent pattern. Analysis of male data revealed that the vowel [a] did not differ significantly from the other vowels except [u]. As it was also observed in case of female data the pair of [i] and [e] also did not show any significant interaction (which seems to be obvious) in terms of the backness property of these vowels.

The details of significant interaction between the vowel types and their formant values are provided in Table 9 and 10 ([*] indicates a significant interaction, whereas [!] denotes non-significant pair.

Table 9. *Significance matrix for formant frequencies of Sylheti male speakers*

Male	N	F1				F2			
		a	i	ɛ	ɔ	a	i	ɛ	ɔ
i	60	*				!			
ɛ	60	*	*			!	!		
ɔ	60	*	*	!		!	*	*	
u	60	*	!	*	*	*	*	*	*

Table 10. *Significance matrix for formant frequencies of Sylheti male speakers*

Female	N	F1				F2			
		a	i	ɛ	ɔ	a	i	ɛ	ɔ
i	60	*				*			
ɛ	60	*	*			*	!		
ɔ	60	*	*	!		*	*	*	
u	60	*	!	*	*	*	*	*	!

In order to examine the perceptual distance between the vowel pairs in Sylheti, I calculated the Euclidean distances that uses the formula in (1).

$$(1) \quad D = \sqrt{(F1_i - F1_j)^2 + (F2_i - F2_j)^2}$$

(D is distance being calculated and “i” and “j” stand for two different vowels), F1 = first formant, F2 = second formant.

Using the above mentioned formula, I have calculated the Euclidean distances between vowel pairs in Sylheti vowel inventory for male and female speakers separately (Table 11 and 12). As expected, the biggest differences were observed among the peripheral vowels [i], [u], and [a]. The pattern was found to be similar in both male and female speakers’ data. The pair of [i] and [ɛ] showed the smallest difference for both male and female speakers. The average Euclidean distance and overall vowel space is found to be less for male speakers (108 Mels) as compared to female speakers (168 Mels).

Table 11. *Euclidean distance between vowel pairs of Sylheti male speakers*

Male	a	i	ɛ	ɔ	u
a	0				
i	163.09	0			
ɛ	126.03	37.60	0		
ɔ	109.45	75.17	53.97	0	
u	192.30	111.26	120.72	88.72	0

Table 12. *Euclidean distance between vowel pairs of Sylheti male speakers*

Female	a	i	ɛ	ɔ	u
a	0				
i	212.97	0			
ɛ	140.92	84.77	0		
ɔ	125.96	204.99	178.66	0	
u	215.75	201.57	215.67	97.182	0

5. VOWELS IN SYLHETI

In this paper with the help of acoustic and statistical analysis I have established the phoneme inventory of Sylheti. The durational measurements considered for analysing Sylheti stops confirmed that the voiced stops in Sylheti do differ only in terms of place of articulation. The underlying aspirated series did not show any significant difference both in terms of acoustic signal and VOT measurements. I did not observe any noise following the burst of underlying voiced stops (as shown in Figure 2 and 3) which prompted me to consider the onset of the voicing start till the burst point. As expected, no statistical difference was observed between the underlying voice aspirated and their non-aspirated counterparts. Through acoustic waveforms, I have also discussed the process of consonant weakening involving spirantization and deaffrication. An acoustic analysis of the vowels has also been conducted and vowel inventories of Sylheti

have been presented. I conclude that the systematic deletion of underlying aspiration [+spread glottis] and the process of lenition are certain to affect the overall phonological property of this language.

NOTES

1. Notice the word-final voiceless stop [k] is also spirantized to [x] ([d^hak > d^hax]).
2. [k] is an allophone of [x] which occurs only when preceded by high vowels [i] and [u] (x → k/[+high, -consonantal]).
3. Spiratization [k] → [x].
4. [k] → [x].
5. [k] → [x].
6. Contrary to the findings of Dutta (2007) Gope (2016), Gope & Mahanta (2014) argued that the loss of the underlying breathiness contrast [+spread glottis] among the Sylheti obstruents indeed raised the f_0 of the neighboring vowels, creating a two way tonal contrasts in Sylheti. Thus Sylheti presents a special case of tonogenesis where a high tone is generated (largely) due to loss of underlying breathiness property of stop consonants (as observed in both voiced and voiceless stops). Following the findings of Dutta (2007), and the findings of VOT analysis discussed in this paper, we will argue that voiced stops in Sylheti do contrast only in terms of their place of articulation.

REFERENCES

- Abramson, A. S. 1962. The vowels and tones of standard Thai: Acoustical measurements and experiments. *International Journal of American Linguistics*, 28/2, part II.
- Anisuzzaman 1988. *Byabharik Bangla UchcharanAbhidhan* [A Dictionary of Standard Bengali Pronunciation]. Dhaka: National Institute of Mass Communication.
- Bhattacharya, K. 1910. *Bengali Phonetic Reader*. Mysore: CIIL.
- Boersma, P. & Weenink, D. 2008. Praat: Doing phonetics by computer. (Version 5.3.04_win32) [Computer program]. 22 Dec 2011. Available online: <<http://www.praat.org/Bradley1982>>.
- Chatterji, S. K. 1926. *The Origin and Development of the Bengali Language*. London: George Allen and Unwin.

- Das, S. 1996. Consonant Weakening in English and Bangla: A Lexical Phonological Investigation. M. Phil Dissertation, EFLU (Formerly CIEFL), Hyderabad.
- Disner, S. 1980. Evaluation of vowels normalization procedures. *Journal of the Acoustical Society of America*, 67, 253-261.
- Dixit, R. P. 1989. Glottal gestures in Hindi plosives. *Journal of Phonetics*, 17, 213-237.
- Dutta, I. 2009. *Acoustics of Stop Consonants in Hindi: Voicing, Fundamental Frequency and Spectral Intensity*. Saarbrücken: VDM Verlag.
- Gope, A. 2016. The Phonetics and Phonology of Sylheti Tonogenesis. PhD Dissertation, IIT Guwahati, India.
- . & Irene, N. V. 2010. Linguistic analysis of West Tripura variety of Bangla: A descriptive study. In C. Sivashanmugam et al. (Eds.), *New Perspectives in Linguistics* (pp. 198-205). Coimbatore: Department of Linguistic, Bharathiar University.
- . & Mahanta, S. 2014. Lexical tone in Sylheti. In *Fourth International Symposium on Tonal Aspects of Languages* (pp. 10-14).
- . & Mahanta, S. 2015. An acoustic analysis of Sylheti phonemes. *The XVIII ICPHS*, 10-14.
- . & Mahanta, S. 2016a. Perception of lexical tones in Sylheti. *Tonal Aspects of Languages*, 142-146.
- . & Mahanta, S. 2016b. Correlation between Sylheti tone and phonation. *Proceedings of the Speech Prosody*.
- Gordon, M. 1996. The phonetic structures of Hupa. *UCLA Working Papers in Phonetics*, 93, 164-187.
- Grierson, G. A. 1903. *The Linguistic Survey of India*. Vol. V. *Indo-Aryan Family. Eastern Group*. Part I. *Specimens of the Bengali and Assamese Languages*. Calcutta: Office of the Superintendent of Government Printing, India.
- Iseli, M., Shue, Y. L. & Alwan, A. 2007. Age, sex and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America*, 121, 2283-2295.
- Kingston, J. 1985. The Phonetics and Phonology of the Timing of Oral and Glottal Event. PhD dissertation, UCLA.
- Ladefoged, P. 1996. *Elements of Acoustic Phonetics*. 2nd Ed. Chicago & London :The University of Chicago Press.
- . & Maddieson, I. 1996. *The Sounds of the World's Languages*. Malden, MA: Blackwell.
- . 1971. *Preliminaries to Linguistic Phonetics*. Chicago, IL: University of Chicago Press.

- Ohala, M. 1983. *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass.
- Qayyum, A. & Razia, S. 2007. *Prachin o Madhyajuger Bangla BhasharAbhidhan* [Dictionary of Old and Medieval Bengali Language]. Dhaka: Bangla Academy.
- Rose, P. J. 1987. Considerations on the normalization of the fundamental frequency of linguistic tone. *Speech Communication*, 10, 229-247.
- Sapir, E. 1925. Sound patterns in language. *Language*, 1, 37-51.
- Shahidullah, Md. 1965. *BangladesherAnchalikBhasharAbhidhan* [A Lexicon of Bangladeshi Dialects]. Dhaka: Bangla Academy.
- Thomas, E. & Kendall, T. 2007. *NORM*: The vowel normalization and plotting suite. Available online: <<http://ncslaap.lib.ncsu.edu/tools/norm/>>.

AMALESH GOPE

DEPARTMENT OF EFL,
TEZPUR UNIVERSITY,
NAPAAM, SONITPUR,
ASSAM 784028.

E-MAIL: <AMALESH@TEZU.ERNET.IN>