ACOUSTIC ANALYSIS OF BANGLA CONSONANTS

Firoj Alam, S.M. Murtoza Habib, Mumit Khan
Center for Research on Bangla Language Processing, BRAC University
{firojalam, habibmurtoza, mumit}@bracu.ac.bd

ABSTRACT

This paper describes the acoustic characteristics of Bangla consonants, obtained by analyzing the recordings of male and female voices. First, the duration of each phoneme was identified by averaging both the male and female voice data; then, formant were measured and formant comparison was made for controversial phonemes, which also served to resolve the controversies in the existing phoneme inventories; and finally, a consonant phoneme inventory was designed.

Index Terms — Phoneme inventory, Speech Synthesis, Speech Recognition

1. INTRODUCTION

The goal of this paper is to determine the total number of consonant phonemes and their acoustic properties in Bangla language. This analysis is an essential component in linguistic of a language and in the diphone concatenation technique where proper duration and prosodic characteristics are needed to synthesize natural sounding speech. With nearly 200 million native speakers, Bangla (exonym: Bengali) is one of the most widely spoken languages of the world (it is ranked between four[1] and seven[2] based on the number of speakers). There have been quite a few articulatory investigations on Bangla phonemes during the last several decades. These analyses have resulted in a set of phoneme inventories, each with several phonemes that are controversial. One of the great motivations of this work is to solve these controversies. This study is focused on acoustic evidences of Bangla consonant phonemes based on the acoustic cues of Pickett [7], durational characteristics and comparison of controversial phonemes. Acoustic evidence is the perfect cue to identify aspiration, voicing, duration and other important features in a phoneme. The durational characteristics, closure, release burst, turbulent, gliding and the formant transitions in the onset and offset were measured for each class of consonant phonemes. Articulatory investigation like minimal pair testing did not consider in this study because it does not help in controversial phonemes in Bangla. A list of dictionary words containing all possible phonemes was selected for recording. The carrier sentences were designed with iC[3]i and aCa patterns. These sentences were then recorded by both professional and non-professional male and female speakers. The recorded data was analyzed with the help of Praat [8].

A brief literature review is given in section 2, followed by a description of the methodology in section 3. The analytical results are presented and discussed in section 4. A summary and conclusions of the study are given in section 5.

2. LITERATURE REVIEW

There have been several studies in the past, mostly based on articulatory phonetics, of the articulatory and acoustic properties of Bangla consonants. It is showed in [1] that Bangla have 20 stops (ক/k/, খ/k^h/, গ/g/, ঘ/g^h/, চ/c/, ছ/c^h/, জ/ɟ/, ঝ/ɟ^h/, ট/ʈ/, ঠ/ʈ^h/, ড/ɖ/, ঢ/ɖ^h/, ত/t̪/, থ/t̪^{h}/, দ/d̪/, ধ/d̪^{h}/, প/p/, ফ/p^h/, ব/b/, ভ/b^h/), 7 fricatives (শ/ʃ/, স/s/, ষ/ʃ^h/, য, ব’/w/, হ/h/, ঃ/ʔ/), 4 nasals (ঙ,ং/ŋ/, ন/n/, ণ, ম/m/), 1 lateral (ল/l/), 1 trill (র/r/), 2 flaps (ড়/ɾ/, ঢ়/ɾ^h/), 1 glide (য়/y/) as a total of 36 consonants. Incidentally, Abdul Hai [1] claimed that the sound produced by three letters শ[ʃ], স[ʃ], ষ[ʃ] is represented by a single phoneme /ʃ/. Daniul Huq [2] showed 21 stops (ক/k/, খ/k^h/, গ/g/, ঘ/g^h/, চ/c/, ছ/c^h/, জ,য/ɟ/, ঝ/ɟ^h/, ট/ʈ/, ঠ/ʈ^h/, ড/ɖ/, ঢ/ɖ^h/, ত/t̪/, থ/t̪^{h}/, দ/d̪/, ধ/d̪^{h}/,প/p/, ফ/p^h/, ফ’/f/, ব/b/, ভ/b^h/), 5 fricatives (ব’/b/, স/s/, জ’/z/, শ/ʃ/, হ/h/), 3 nasals (ঙ,ং/ŋ/, ন/n/, ম/m/), 1 lateral (ল/l/), 1 trill (র/r/), 2 flaps (ড়/ɽ/, ঢ়/ɽ^h/), 2 glides (ব’/w/, য়’/y/) (total of 35 consonants). It is explained in [6] that Bangla have 20 stops (ক/k/, খ/k^h/, গ/g/, ঘ/g^h/, চ/c/, ছ/c^h/, জ/ɟ/, ঝ/ɟ^h/, ট/ʈ/, ঠ/ʈ^h/, ড/d/, ঢ/d^h/, ত/t̪/, থ/t̪^{h}/, দ/d̪/, ধ/d̪^{h}/,প/p/, ফ/p^h/,

[1] http://www2.ignatius.edu/faculty/turner/languages.htm, Last accessed December 26, 2007.

[2] http://en.wikipedia.org/wiki/List_of_languages_by_total_speakers, Last accessed December 26, 2007.

[3] C - Consonant

ব/b/, ভ/b^h/), 4 nasals (ঙ,ং/ŋ/, ন/n/, ণ[n][4], ম/m/), 4 fricatives (স[s], ষ[ʃ], শ[ʃ], হ/h/)[5], 1 lateral (ল/l/), 2 flaps (র/r/, ড়/ɾ/) as a total of 31 consonants. According to [11] Bangla have 20 stops (ক/k/, খ/k^h/, গ/g/, ঘ/g^h/, চ/c/, ছ/c^h/, জ/ɟ/, ঝ/ɟ^h/, ট/t/, ঠ/t^h/, ড/d/, ঢ/d^h/, ত/t̪/, থ/t̪^{h}/, দ/d̪/, ধ/d̪^{h}/, প/p/, ফ/p^h/, ব/b/, ভ/b^h/), 3 nasals (ঙ,ং/ŋ/, ন/n/, ম/m/), 3 fricatives (শ/ʃ/, স/s/, হ/h/), 1 lateral (ল/l/), 2 flaps (ড়/ɽ/, ঢ়/ɽ^h/), 1 trill (র/r/), 2 glides (ব/w/, য়/y/) (total of 32 consonants). According to Wikipedia [10], Bangla has 29 consonants with 20 stops (ক/k/, খ/k^h/, গ/g/, ঘ/g^h/, চ/ʧ/, ছ/ʧ^h/, জ/ʤ/, ঝ/ʤ^h/, ট/ʈ/, ঠ/ʈ^h/, ড/ɖ/, ঢ/ɖ^h/, ত/t̪/, থ/t̪^{h}/, দ/d̪/, ধ/d̪^{h}/, প/p/, ফ/p^h/, ব/b/, ভ/b^h/), 3 nasals (ঙ/ŋ/, ন/n/, ম/m/), 3 fricative (শ/ʃ/, স/s/, হ/h/), 3 liquids (ল/l/, র/r/, ড়/ɽ/). Hossain et. al. [3] used acoustic analysis to study voiced and voiceless classification for labeling and recognition task. Another study [5] showed that Bangla has 16 stops (ক/k/, খ/k^h/, গ/g/, ঘ/g^h/, ট/t/, ঠ/t^h/, ড/d/, ঢ/d^h/, ত/t̪/, থ/t̪^{h}/, দ/d̪/, ধ/d̪^{h}/, প/p/, ফ/p^h/, ব/b/, ভ/b^h/), 4 affricates চ/ʧ/, ছ/ʧ^h/, য,জ/ʤ/, ঝ/ʤ^h/), 3 fricatives (শ,স,ষ/ʃ/, শ,স/s/, হ/h/), 3 nasals (ঙ,ং/ŋ/, ন,ণ/n/, ম/m/), 1 trill (র/r/), 2 flaps (ড়/ɾ/, ঢ়/ɾ^h/), 1 glide (য়/y/) as a total of 30 consonants. From these studies, it is observed that there is some controversy with the phoneme pairs ন/n/-ণ /n/, য/ɟ/-জ/ɟ/, and ড়/ɾ/-ঢ়/ɾ/ among different linguists. Manzur Morshed [6] mentioned the distinction of ন/n/-ণ/n/ whereas others showed these two as a single phoneme ন,ণ/n/. Abdul hai [1] mentioned the existence of য/ɟ/-জ/ɟ/ whereas others showed these two as a single phoneme জ,য/ɟ/. Wikipedia [10] showed ড়, ঢ়/ɽ/ as a single phoneme whereas others showed them as two separate phonemes (ড়/ɾ/, ঢ়/ɾ^h/). It is claimed that the following three phonemes ফ'/f/ [2], ব'/w/, ঃ/ʔ/ [1] are present in Bangla. The phoneme ফ'/f/ [2] is omitted from this study because there is no such word in Standard Bangla containing the said phonemes. We have considered the phoneme ব'/w/ as part of a diphthong, to be discussed in another paper and the phoneme ঃ/ʔ/ as the allophonic variation of হ/h/ [11, pp-105]. The consonant phoneme list used in this study is basically a union of the published inventories – ক/k/, খ/k^h/, গ/g/, ঘ/g^h/, চ/c/, ছ/c^h/, জ/ɟ/, য/ɟ/, ঝ/ɟ^h/, ট/t/, ঠ/t^h/, ড/d/, ঢ/d^h/, ত/t̪/, থ/t̪^{h}/, দ/d̪/, ধ/d̪^{h}/, প/p/, ফ/p^h/, ব/b/, ভ/b^h/, শ/ʃ/, ষ/ʃ/ স/ʃ/, স/s/, হ/h/, র/r/, ড়/ɾ/, ঢ়/ɾ^h/, ল/l/, ঙ, ং/ŋ/, ন/n/, ণ/n/ ম/m/, য়/y/ – with a total of 35 possible phonemes.

[4] Phonemic transcription, as author did not represent IPA in his literature.

[5] Authors identified as স[s] Dental, ষ[ʃ] Alveolar and শ[ʃ] Palatal. No IPA representation available in literature. We used to make it understandable by users.

3. METHODOLOGY

This study seeks to find the answer of the following three questions in order to develop a Bangla phoneme inventory: (1) to find the comparison between controversial phoneme pairs; (2) to determine the durational values; and (3) to determine the formant characteristics of the consonants in spoken Bangla utterances for the phoneme inventory. A list of dictionary words embedded in carrier utterances was chosen for the analysis. Different patterns were selected for identifying the list of words. This data was recorded by a number of speakers and then analyzed by the Praat software.

3.1. Recording material

The list of words selected for this investigation consists of all possible phonemes with the following two patterns: vCv [iCi] and vCv [aCa], embedded in carrier words to form utterances. The reason for recording consonants in a carrier utterance is so that the context remains the same. So a total of 35x2 (35 possible phonemes x 2 patterns) utterances were selected for recording of the following form.

1. aCa pattern

আমরা কাজ পাই -> ক

amra kaɟ pai -> /k/

1stP.Pl work get.pres

[We get work.]

2. iCi pattern

আমি কিছু পাই -> ক

ami kic^{h}u pai -> /k/

1st.Sg some get.pres

[I get something.]

The first character of the second word is the target phoneme in both patterns. The same carrier words were used for all target phonemes. Pickett [7, pp-87] summarizes from Umeda (1977) that the duration of consonants are also affected by the syllable stress, emphasis and position of the consonant in a word. So the utterances were selected in such a way such that the prosodic variation (such as stress, tone, emphasis and vocal effort) and feature dependent segment duration do not have any effect on the target phoneme. Also, the manner of articulation was considered when these utterances were collected, as the manner of articulation is the usual first basis for segmentation or duration calculation. All words are phonetically symmetrical i.e. the grapheme to phoneme correspondence is regular, an assertion that was confirmed by linguists.

3.2. Speaker selection

Both professional and non-professional male and female speakers were selected by considering different ages, heights and the speakers' locality in Bangladesh.

Unfortunately, we were unable to include any speaker from the Indian State of West Bengal in this analysis. Four male and four female speakers, with equal numbers of professional vs. non-professional male speakers, were selected. The professional speakers' ages ranged from 52 to 54 and non-professional speakers' ages ranged from 25 to 29. Each speaker was given flash cards containing the utterances, and was asked to record each utterance in straight tone/pitch level and without assigning any stress in a word.

3.3. Recording

The recording of the utterances was done using the Nundo speech processing software. A professional recording studio was chosen to record the utterances. The equipment consisted of an integrated Tascam TM-D4000 Digital-Mixer, a high fidelity noise free Audiotechnica microphone and two high quality speakers. The recorded waveform files were used for acoustic analysis with Praat version 4.6.27. The speakers were asked to keep a distance of 10-12 inches from the microphone. The speech data was digitized at 44100 Hz at 24-bit resolution and stored as wave format. After each recording, the moderator checked for any wrong pronunciation during the recording, and if so, the affected utterances were re-recorded.

3.4. Analysis

Total 35x2X8 = 560 (35 possible phonemes x 2 patterns x 8 speakers) segment were analyzed in this study. Scarborough R. [9] showed a set of segmentation criteria (i.e beginning and ending position of each classes of phoneme), in her lecture which was considered when the duration was calculated using Praat. The duration for each phoneme (both closure and leg VOT-Voice onset time) was computed from the recorded voice. The start position of the plosive is the end of preceding vowel and the end position of the plosive is the beginning of the voicing of the following vowel. The frication and aspiration after the release burst of plosives was considered a part of the consonants. The average value was then calculated for each phoneme. Durations for both male and female speakers were compared. Onset and offset formant movement of consonant phonemes was observed and noted to identify place of articulation. Acoustic cues of each place and manner [7, pp-120, 140] were used for each consonant phoneme. According to [7], acoustic cues of all phonemes were classified by place and manner of articulation. Then the comparisons between the controversial phonemes were identified as is explained in the result section.

Praat settings: The Praat spectrogram and formant settings were maintained for both male and female speakers. In spectrogram settings the window length is 0.005 second, the window shape Gaussian and the view range up to 5000 hz. The Fourier analysis method is used in spectrogram analysis settings. In formant settings, we used maximum formant 5500 Hz for five formants with window length 0.025 second. Pitch range was used 75 Hz to 500 Hz which covers both male and female speakers.

4. RESULTS

The durations, controversial phoneme comparison and consonant phoneme inventory were computed and identified in this analysis. The output of the analysis is presented and discussed in this section. The durational characteristics of Bangla consonant phonemes are shown in Table 1 and Table 2. Table 3 shows the consonant phoneme inventory of Bangla. According to the different formant value we classified and identified the place and manner feature of consonant phonemes which is explained in section 4.3. The following controversial phonemes শ/ʃ/, ষ/ʃ/ স/ʃ/, স/s/, ড়/ɾ/, ঢ়/ɾʰ/, ন/n/, ণ/n/, জ/ɟ/, য/ɟ/ solved in two steps. Formants value were observed for these phoneme শ/ʃ/, ষ/ʃ/ স/ʃ/, স/s/ and it is identified that the phoneme শ,স/s/ (fricative) has strong frequency above 4500 Hz and frication of the palatal phoneme শ,ষ,স/ʃ/ start at 1600 Hz and strong high frequency above 3000 Hz. So the two phonemes were eliminated in this step. In second step, formant comparison were made for these phoneme ড়/ɾ/, ঢ়/ɾʰ/, ন/n/, ণ/n/, জ/ɟ/, য/ɟ/ explained in section 4.2 and then we identified three phonemes among the six phonemes. So finally we concluded 30 phonemes from the 35 phonemes with their acoustic values.

4.1. Durational characteristics

In Table 1, column 1 shows the plosive phonemes and the rest of the column shows the average duration of male and female, total average of both male and female and standard deviation of both male and female. Closure and VOT were calculated separately in each case. It was observed that the duration of the aspirated sound is longer than the un-aspirated one and VOT (due to aspiration) is longer than closure. The duration of glide, trill, flap identification was the most difficult part during analysis as start and end segment could not be identified easily. Table 2 shows the durational value of fricative, nasal, trill, flap and glide.

Phoneme	Male				Female				Total			
	Closure (Avg)	VOT (Avg)	Closure (std)	VOT (std)	Closure (Avg)	VOT (Avg)	Closure (std)	VOT (std)	Closure (Avg)	VOT (Avg)	Closure (std)	VOT (std)
ক /k/	105.11	35.13	27.83	9.08	109.53	29.38	79.45	7.62	107.32	32.26	57.55	8.62
খ /k^h/	86.51	104.47	23.97	30.37	83.54	90.22	20.63	15.06	84.93	96.87	21.48	23.72
গ /g/	82.86	30.70	14.47	8.23	84.56	25.91	29.49	7.66	83.71	28.31	22.45	8.07
ঘ /g^h/	77.98	123.59	18.61	56.90	74.64	103.85	29.57	11.47	76.31	113.72	23.93	40.94
চ /c/	93.55	74.32	20.84	22.83	78.30	52.33	19.22	9.23	85.93	63.33	20.91	20.29
ছ /c^h/	72.88	134.90	14.47	61.74	70.33	102.35	11.02	30.10	71.61	118.63	12.50	49.84
য,জ /ɟ/	67.50	52.49	19.61	35.08	68.10	36.68	13.70	10.73	67.80	44.58	16.64	26.75
ঝ /$ɟ^h$/	64.83	130.71	8.25	63.26	57.17	111.91	10.23	19.15	61.00	121.31	9.81	46.18
ট /ʈ/	130.22	17.19	30.56	5.20	96.56	11.28	12.03	4.41	113.39	14.24	28.38	5.57
ঠ /$ʈ^h$/	102.34	91.95	17.57	25.13	83.97	62.75	12.85	14.87	93.15	77.35	17.64	25.01
ড /ɖ/	88.58	18.58	14.27	6.24	83.49	9.90	10.27	2.92	86.03	14.24	12.29	6.49
ঢ /$ɖ^h$/	86.16	98.72	24.89	54.66	71.81	80.19	9.79	32.90	78.98	89.45	19.72	44.62
ত /t̪/	131.75	19.92	54.33	7.12	120.35	13.11	12.13	5.84	126.05	16.52	38.48	7.21
থ /$t̪^h$/	139.98	103.95	67.52	22.37	89.10	77.55	15.93	17.34	114.54	90.75	54.19	23.66
দ /d̪/	100.35	12.69	23.53	8.07	86.47	10.24	11.90	2.91	93.41	11.46	19.38	6.00
ধ /$d̪^h$/	88.82	103.06	30.48	38.56	76.33	62.94	13.54	33.84	82.57	83.00	23.68	40.71
প /p/	129.30	19.18	34.66	10.25	98.98	9.54	12.23	3.43	114.14	14.36	29.59	8.90
ফ /p^h/	101.99	88.31	36.14	41.70	89.21	58.15	34.58	39.40	96.03	74.24	34.77	42.16
ব /b/	94.22	12.70	15.54	4.54	89.41	8.73	11.38	2.88	91.81	10.72	13.39	4.20
ভ /b^h/	81.85	95.35	18.14	42.73	76.61	58.70	10.22	34.07	79.23	77.02	14.47	41.86

Table 1**:** Average duration of Stops in millisecond

Phoneme	Male (Avg)	Female (Avg)	Male (Std)	Female (Std)	Total (Avg)	Total (Std)
শ,ষ,স/ʃ/	197.02	153.26	60.48	15.28	175.14	48.59
শ,স/s/	166.82	127.70	18.86	16.31	147.26	26.52
ম /m/	108.49	92.56	23.88	11.54	100.52	19.34
ঙ,ং /ŋ/	87.50	109.92	14.49	47.41	98.71	34.60
ণ, ন /n/	65.81	72.32	18.36	38.25	69.07	30.49
র /r/	62.35	67.07	9.30	13.63	64.71	11.09
ল /l/	108.30	93.10	19.36	10.92	100.70	16.66
ড়, ঢ় /ɾ/	48.21	50.93	23.23	14.64	49.57	18.80
য় /j/	81.64	69.77	30.46	20.09	75.70	24.71
হ,ঃ/h/	156.17	85.43	149.87	87.49	120.80	55.29

Table 2: Average duration of fricative, nasals, trill, lateral, flap and approximant in millisecond

4.2. Controversial comparison

There is a controversy in terms of the sound produced between the following pairs of letters: “ন”/n/-“ণ”/n/; “য”/ɟ/-“জ”/ɟ/; “ড়”/ɾ/-“ঢ়”/ɾ/. Formants were measured in whole durations for each pair of sounds by Praat formant listing. Praat uses 25ms window length, five formants and maximum 5500 Hz in formant settings. Each pair is then compared as shown in the graph. The same procedure is applied for every pair. The graph shown here is the calculated value of one male voice. Other speaker comparison values are similar, which has not been shown. The spectrographic analysis of letter “ন”/n/ in দিনার /d̪inar/ and letter “ণ”/n/ in কণা /kɔna/ shows that both are the same phoneme /n/ as shown in Figure 1. The average duration of “ন”/n/ is 69.69 (msec) and “ণ”/n/ is 68.44 (msec). Different

linguists claimed that the two letters "য"/ɟ/ and "জ"/ɟ/ represent two phonemes in Bangla. But in our analysis we identified that both are same phoneme (য, জ /ɟ/) as is shown in Figure 2. The average duration of "য"/ɟ/ is 67.33 (msec) and "জ"/ɟ/ is 68.27 (msec). The average duration of "ড়"/ɾ/ is 50.90 (msec) and "ঢ়"/ɾ/ is 48.29 (msec). The formant graph of the sound produced by the letters "ড়"/ɾ/ in ষাড়/ʃaɾ/, and "ঢ়"/ɾ/ in আষাঢ়ে/aʃaɾe/ is shown in Figure 3. It shows some differences in F2 formants. More data may be required for exhaustive analysis. For now we consider these as a single phoneme /ɾ/.

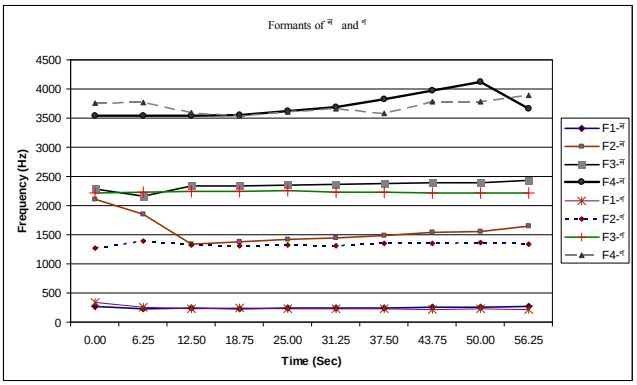

Figure 1: Comparison of sound produced by the lettter ন and ণ

Another finding of our analysis is that the sound produced by the letters "শ", "ষ" and "স" is the phoneme /ʃ/. But the letter "শ" and "স" also produce another sound which is identified as phoneme /s/.

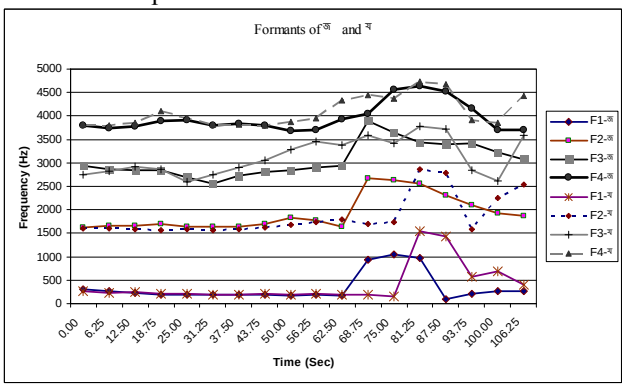

Figure 2: Comparison of sound produced by the lettter জ and য

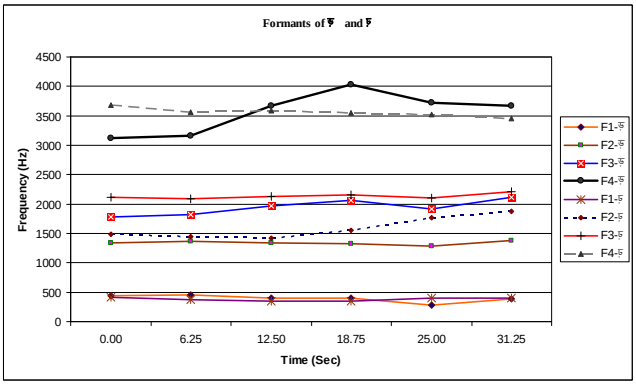

Figure 3: Comparison of sound produced by the lettter ড় and ঢ়

4.3. Consonants phoneme inventory

This study identified 30 consonant phonemes in Bangla as shown in table 3, including 20 plosives, 3 nasals, 1 tril, 1 flap, 3 fricatives, 1 lateral and 1 approximent. The plosive sound in each place composed of two voiced versus two voiceless and two aspirated versus two in-aspirated. Each aspirated sound is represented by superscript (^{h})

Bilabial: The average formant value of the following vowel of bilabial plosive shows that, F2 is upward. Another significant acoustic cue is the weak and diffuse spectrum in the release burst. The four bilabial plosives composed of two voiced (ব/b/, ভ/b^h/) versus two voiceless (প/p/, ফ/p^h/) and one bilabial nasal (ম/m/). The bilabial nasal also has an upward F2, strong frequency up to about 600Hz and very weak mid-formants.

Dental: The four dental plosives (ত/t̪/, দ/d̪/, থ/t̪^{h}/, ধ/d̪^{h}/) have downward F2 of the following vowel and have high frequency energy in release burst.

Alveolar: There are four alveolar plosives (ট/t/, ঠ /t^h/, ড/d/, ঢ /d^h/) of Bangla which have a downward F2 of the following vowel. Bangla has alveolar nasal, trill, flap, fricative and lateral. The spectrogram of the alveolar nasal phoneme shows that, the F2 of the following vowel is downward. The phoneme trill র/r/ has high intensity on F1, which gradually decreases in F2, F3 and F4, low freq strong up to 835 Hz, mid freq stronger than nasals and following vowel F2 is upward. The phoneme ড়, ঢ়/ɾ/ (flap) follow vowel transition on F1, F2 and F3. The phoneme ল/l/ (lateral) have high frequency at 400 Hz, 1350 Hz, and 3300 Hz, the intensity of the formants gradually decrease after the first formants. The phoneme শ,স/s/ (fricative) has strong frequency above 4500 Hz.

Place / Manner		Bilabial		Dental		Alveolar		Post- Alveolar		Palatal	Velar		Glottal
Stops	voiceless	প/p/	ফ/p^h/	ত/t̪/	থ/$t̪^h$/	ট/t/	ঠ /t^h/	চ/c/	ছ/c^h/		ক/k/	খ/k^h/	
	voiced	ব/b/	ভ/b^h/	দ/d̪/	ধ/$d̪^h$/	ড/d/	ঢ /d^h/	য, জ/ɟ/	ঝ/$ɟ^h$/		গ/g/	ঘ/g^h/	
Nasals		ম/m/				ন,ণ/n/					ঙ,ং/ŋ/		
Trill						র/r/							
Flap						ড়, ঢ়/ɾ/						
Fricatives						শ,স/s/				শ,ষ,স/ʃ/			হ,ঃ/h/
Lateral						ল/l/							
Approximant										য়/j/			

Table 3: Consonants phoneme inventory

Post-alveolar: It is observed that the spectrogram of the post-alveolar plosives (চ/c/, ছ/c^h/, য, জ/ɟ/, ঝ/$ɟ^h$/) shows downward F2 transition and generally stronger high frequency energy after release burst.

Palatal: Frication of the palatal phoneme শ,ষ,স/ʃ/ start at 1600 Hz. Strong high freq above 3000 Hz. The phoneme য়/j/ has different spectrographic patterns in different context.

Velar: The spectrogram of the velar phonemes (4 plosives ক/k/, খ/k^h/, গ/g/, ঘ/g^h/ and 1 nasal ঙ,ং/ŋ/) shows the divergent F2 and F3 of the preceding and following vowel. The phoneme ঙ,ং/ŋ/ have high intensity at 300 Hz, 1300 Hz, 2300 Hz and 3800 Hz. Release transient on F1 and F2 of the following vowel.

Glottal: The fricative হ/h/ phoneme have the strongest resonance at about 1100 Hz and go upward.

5. CONCLUSION

Here we described the duration of each consonant phoneme and we also identified the place and manner of articulation of consonants in the phoneme inventory. We also conclude that Bangla consonant phoneme inventory consist of 30 phonemes. This phoneme inventory can be used in all Bangla linguistic components and to develop speech application in Bangla. It may also help in diphone database for speech synthesis, speech recognition as well as speech processing such as speech-to-speech translation.

6. ACKNOWLEDGMENTS

This work has been supported in part by the PAN Localization Project (www.panl10n.net), grant from the International Development Research Center (IDRC), Ottawa, Canada, administrated through Center for Research in Urdu Language Processing (CRULP), National University of Computer and Emerging Sciences, Pakistan. We would also like to thank Dr Sarmad Hussain (NUCES), Sameer Ud Daula (UCLA), Naira Khan (Dhaka University) and BRAC University students who helped by providing their speech for analysis.

7. REFERENCES

[1] Abdul Hai. *Dhvani Vijnan O Bangla Dhvani-Tattwa*, 10th Reprint, 2007, Mullick Brothers, Dhaka, pp-12-35, 1967.

[2] Daniul Huq. *Bhasha Bigganer Katha (Facts about Linguistics),* Mowla Brothers, Dhaka, pp-81-93, 2002.

[3] Hossain A., Nahid N., Khan N. N., Gomes D. C., Mugab S. M.. *Automatic Silence / Unvoiced / Voiced Classification of Bangla Velar Phonemes: New Approach.* 8th ICCIT, Dhaka, 2005.

[4] Ladefoged P. *A course in phonetics.* 4th edition 2002, Thomson Asia Pte Ltd. Singapore. pp-161-170. 191-196, 2002

[5] Mahbubul Haque. *Bangla Bhashar Bekaron and Rachonariti (Grammar and Essay of Bangla Language)*, 10th reprint 2007, Boi Prokashony, Dhaka, pp-231-233, 6th Edition 2004.

[6] Manzur Morshed. *Adhunik Bhasatatto (Modern Linguistics)*, Mowla Brothers, Dhaka, pp-219-237, 3rd Edition 2001.

[7] Pickett, J.M. *Acoustics of Speech Communication, The: Fundamentals, Speech Perception Theory, and Technology*, Allyn & Bacon, 1998.

[8] Praat. 2007. www.fon.hum.uva.nl/praat/. Version - 4.6.27.

[9] Scarborough R. *Segmentation and Segment Durations*. http://www.stanford.edu/class/linguist205/index_files/Handout%203%20-%20Segmentation.pdf, last accessed December 26' 2007.

[10] Wikipedia, *Bengali phonology* http://en.wikipedia.org/wiki/Bengali_phonolog, last accessed December 26' 2007.

[11] Zeenat Imtiaz Ali. *Dhanibijnaner Bhumika (Introduction to Linguistics)*, Mowla Brothers, Dhaka, 2001.