DEVELOPMENT OF ANNOTATED BANGLA SPEECH CORPORA

Firoj Alam, S.M. Murtoza Habib, Dil Afroza Sultana, Mumit Khan
Center for Research on Bangla Language Processing, BRAC University
{firojalam, habibmurtoza}@bracu.ac.bd, upama.008@gmail.com, mumit@bracu.ac.bd

ABSTRACT

This paper describes the development procedure of three different Bangla read speech corpora which can be used for phonetic research and developing speech applications. Several criteria were maintained in the corpora development process that includes considering the phonetic and prosodic features during text selection. On the other hand, a specification was maintained in the recording phase as the speaking style is a vital part in speech applications. We also concentrated on proper text normalization, pronunciation, aligning, and labeling. The labeling was done manually – in the present endeavor sentence level labeling (annotation) was completed by maintaining a specification so that it could be expanded in future.

Index Terms **—**Speech corpora, Phonetic research, Speech processing

1. INTRODUCTION

The goal of this paper is to present the development of Bangla annotated read speech corpora which is essential for all kinds of speech processing work starting from acoustic analysis to the development of speech synthesis and speech recognition. These corpora were composed from three different corpora and those were developed for three different purposes.

1. "Corpus for acoustic analysis" was developed for acoustic analysis of Bangla phoneme inventory.
2. "Diphone corpus" was developed for diphone concatenation based speech synthesis.
3. "Continuous speech corpus" was developed for intonation model and unit selection based speech synthesis.

Though, these corpora were developed for different purpose however the use of this resource is innumerable i.e. speech recognition, speaker identification, and spoken information extraction. This resource is also an essential component in linguistic analysis of a language. Compared to other languages, very little work has been done in Bangla. CDAC [1] has developed speech corpora, but there is no published account about the details of their corpora. Since the current Bangla speech synthesis system [2] lacks the naturalness due to the intonation model, hence developing an intonation model from these corpora was one of the primary goals. It is also hoped that, a good unit selection based speech synthesis can be developed from these corpora which may have a naturalistic sound.

A brief literature review is given in section 2, followed by a description of the development in section 3, and section 4 presents the corpus annotation and analysis. Conclusion and future remarks of the study are given in section 5.

2. LITERATURE REVIEW

Several studies show significant improvement on designing, developing and annotating of the corpus. The bases of these are [3][4][5][6][7]. Though depending on the purpose of the corpus, different text pattern have chosen, but the developing process remains the same. Other than the designing and development procedure, significant work [8] [9] has proven the high performance of corpus based synthesizers. This signifies that, developing a phonetically and prosodically rich corpora can lead us to develop better speech applications.

3. DEVELOPMENT PROCEDURE

The development of corpora was done in three steps [3]; such as text selection, speaker selection and voice recording. According to [4] and [5], the following characteristics have to be considered during the development of the corpora:

- Area of speech corpora: Speech synthesis, phonetic research and speech recognition.
- Spoken content: Two approaches considered such as domain and phonological distribution.
- Professional recording studio: This is necessary for a clear acoustic signal from which it is possible to get clear acoustic information.
- Speaking style: Continuous read speech.
- Manual segmentation: Though this leads to significant amount of effort but it also affirm the accuracy of the labeling.
- Recording setup: Supervised onsite recording.

3.1. Text selection

Corpus for acoustic analysis: There are two categories of text in this corpus, one is for vowel and another is for consonants. For consonants phoneme investigation, the list of words selected consists of all possible phonemes with the following two patterns: vCv [iCi] and vCv [aCa], embedded in carrier words to form utterances. To maintain the same context we have embedded the consonants in a carrier utterance. So a total of 35x2 = 70 (35 [12] possible phonemes x 2 patterns) utterances were selected to record the data of the following form.

1. aCa pattern

আমরা কাজ পাই -> ক

amra kaɟ pai -> /k/

1st P.Pl work get.pres

[We get work.]

2. iCi pattern

আমি কিছু পাই -> ক

ami kicʰu pai -> /k/

1st.Sg some get.pres

[I get something.]

For vowel phoneme investigation, three different patterns were selected with the nearest number of phoneme segment in each pattern. Each pattern carried two to three syllables. The main intuition of selecting these patterns was duration calculation and formant measurement of vowels. These patterns are:

1. cV.Cv.cvc where V is the target vowel and C is either voiced or voiceless plosive.
2. cV.v.cvc where V is the target vowel
3. cV?V.Cvc where V?V indicates diphthong and C is either voiced or voiceless plosive.

The tricky part was the data collection comprising these patterns. For each pattern we have selected four words from the dictionary [10] to make a balance of the recording data. For the first pattern, the C of second syllable is voiced plosive in two words and voiceless in two words. The reason behind this is that the vowel before voiced is longer than the voiceless plosive [11]. So we will get average duration in both cases. For the second pattern, we were unable to find any word from the dictionary [10]. Then we changed the pattern to v.v.cv.cvc, as the main intention of this pattern was to calculate the duration of the two consecutive vowels appearing in two syllables. However, with this new pattern we found only two words. For the third pattern, four words were selected for each target diphthong. But due to word limitation of this pattern, the first consonant of the second syllable was chosen arbitrarily. In some cases we got only two words then we repeated these two to make it four which was a criterion to make a balance of all phonemes in this analysis. After that, another two carrier words were embedded to form sentences. For example,

এখন গবেষক বলো

ekʰon gɔbeʃok bɔlo

Now researcher say.pres

[Now say researcher]

The middle word is our target word. The vowel of the first syllable of the target word is the target phoneme. The list of words selected for this investigation consists of all possible vowel phonemes with the above patterns, embedded in carrier words to form the utterances. A total of 192 utterances were designed for recording with the following form:

1. 14x4 (14 possible phonemes x 4 words) = 56 (pattern cV.Cv.cvc)
2. 1x4 (1 phoneme x 4 words) = 4 (pattern cV.v.cvc)
3. 33x4 (33 possible phonemes x 4 words) = 132 (pattern cV?V.Cvc)

Total = 192 words

The utterances were selected in such a way so 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 listed words were phonetically defined if required, an assertion that was confirmed by linguists. It is proclaim that this corpus has 100% phoneme coverage.

Diphone corpus: According to [12] and [13] Bangla language has 30 consonants and 35 vowels (monophthong, diphthong) phonemes. In general, the number of diphone in a language is the square of the number of phones. Since Bangla language consists of 65 phones, so the number of diphones are (65X65) 4225. In addition, silence to phones are (1X65) 65 and phones to silence are (65X1) 65. So the total number of diphones is 4335. These diphones were embedded with carrier sentences. Though there have been various techniques to embed diphone with carrier sentences, here nonsense words were used to form carrier sentences [14]. It has 100% coverage of phone and diphone

Continuous speech corpus: Language is evolving; everyday new words appear in newspapers, magazines and blogs, which have different spoken variety. So we decided to use the spoken variety of texts. Then, texts were collected from various domains as shown in table 1. The text corpus contains 1,06,860 tokens, ~10K sentences. Some text was encoded in ASCII which was later converted into Unicode using the CRBLP Converter [15]. Then, the spelling and conversion errors were manually corrected. Table 2 shows the token and sentence count of the three corpora.

Category	Tokens	Source
Megazine (weekly)	31296	1
Novel (Beji-Weasel)	30504	2
Legal document (Child)	1909	3
History (Dhaka, Bangladesh, Language movement, 7th March)	10795	4
Blog (interview)	2347	5
Novel (Rupaly Dip)	12160	6
Editorial (prothom-alo)	3963	7
Constitute of Bangladesh	3278	8
News -Prothom alo	10608	9
Total	106860	

Table 1: Different domains of the corpus

Name	Tokens	Token type	Sentences
Corpus for acoustic analysis	602	203	262
Diphone corpus	12,938	2318	4,335
Continuous speech corpus	1,06,860	17,797	10,895

Table 2: Token and sentence count of the three corpora.

3.2. Speaker selection

Corpus for acoustic analysis: 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. The education of all speakers is above bachelor degree.

Diphone and Continuous speech corpus: A professional voice talent of a 29 years old male native Bengali speaker was hired for recording.

3.3. Recording

The recording of the utterances was done using the Nundo speech processing software. A professional voice 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 multimedia speaker systems. The voice talents were asked to keep a distance of 10-12 inches from the microphone. Optionally a pop filter was used between the speaker and the microphone to reduce the force of air puffs from bilabial plosive and other strongly released stops. The speech data was digitized at a sample rate 44.1 kHz, sample width 24-bit resolution and stored as wave format. After each recording, the moderator checked for any misleading pronunciation during the recording, and if so, the affected utterances were re-recorded.

There were a few challenges in the recording. First, speakers were asked to keep the speaking style consistent. Second, speakers were supervised to keep the same tone in the recording. Since speaking styles varies in different sessions a monitoring were required to maintain the consistency. To keep the consistency of the speaking style, in addition to [3] the following specifications were maintained:

1. Recording were done in the same time slot in every session i.e 9.00 am to 1.00 pm.
2. A 5 minutes break was maintained after each 10 minutes recording.
3. Consistent volume of sound.
4. Normal intonation was maintained without any emotion.
5. Accurate pronunciation.
6. Pre-recorded voice with appropriate speaking style was used as a reference. In each session, speaker was asked to adjust his speaking style according to the reference voice.

4. CORPUS ANNOTATION AND ANALYSIS

4.1. Annotation

There were a few challenges in annotation. The "Corpus for acoustic analysis" and "diphone corpus" was pre-modified. There was no non-standard word (NSW) [16]. That is why no text-normalization was required in those corpora. However, the challenges came in "Continuous speech corpus". A text-normalization tool [17] was required to normalize the text. In case of ambiguous token, the accuracy of the tool is 87% which motivated us to perform a manual check. After the manual checking, phonetic transcription was done using CRBLP pronunciation lexicon [18]. The CRBLP pronunciation lexicon contains all the lexicon entries that are available in a continuous speech corpus. In phonetic form IPA was used. It [18] also proclaims the 100% accuracy of pronunciation form. Later, a script was used to split the corpus into sentences based on punctuation marks such as ?, । and !. Each sentence was assigned a sentence id with orthographic and phonetic form and the same id was used in wave file.

The un-cleaned recorded data was around 24 hours and it has a lot of repetition of the utterances. So in annotation, the recorded wave was cleaned manually using wavlab which tends to reduce the recorded data to 13 hours 32 minutes. Then, it is labeled (annotated) based on id using praat [19]. Praat provides a textgrid file which contains labels (in our case it is wave id) along with start and end time for each label. A separate praat script was written to split the whole wave into individual wave based on id with start and end time. As praat does not support unicode, so id is used instead of text in labeling. Fig 1, 2 and 3 shows the examples of orthographic and phonetic form of the corpora. Fig 4 shows the labeling using praat.

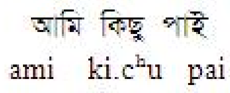


Figure 1: Orthographic and phonetic of "Corpus for acoustic analysis"

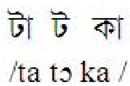


Figure 2: Orthographic and phonetic of "Diphone corpus"

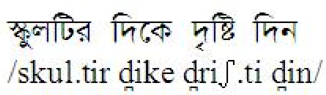


Figure 3: Orthographic and phonetic of "Continuous speech corpus"

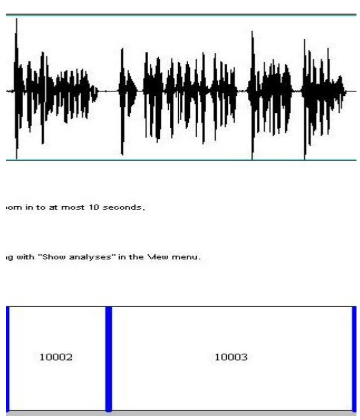
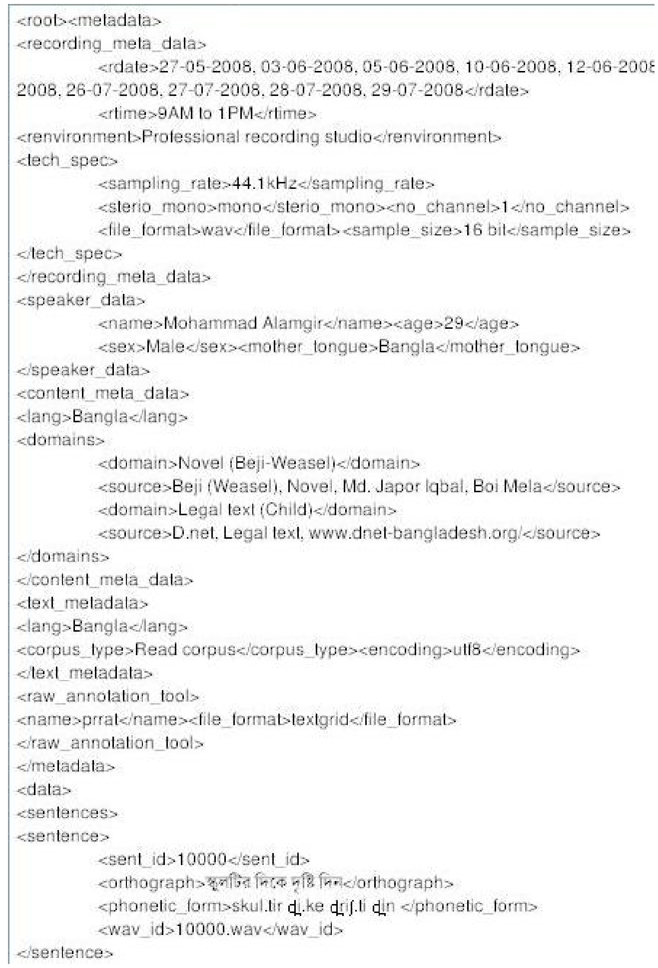


Figure 4: Labeling using praat

4.2. Corpus structure

The structure of the corpus was constructed in a hierarchical organization using the XML standard. The file contains meta data followed by data. The metadata contains recording protocol, speaker profile, text, annotation and spoken content. The data contains sentences with id, orthographic form, phonetic form and wave id. The structure is elicited in figure 5.



```
<root><metadata>
<recording_meta_data>
        <rdate>27-05-2008, 03-06-2008, 05-06-2008, 10-06-2008, 12-06-2008
2008, 26-07-2008, 27-07-2008, 28-07-2008, 29-07-2008</rdate>
        <rtime>9AM to 1PM</rtime>
<renvironment>Professional recording studio</renvironment>
<tech_spec>
        <sampling_rate>44.1kHz</sampling_rate>
        <sterio_mono>mono</sterio_mono><no_channel>1</no_channel>
        <file_format>wav</file_format><sample_size>16 bit</sample_size>
</tech_spec>
</recording_meta_data>
<speaker_data>
        <name>Mohammad Alamgir</name><age>29</age>
        <sex>Male</sex><mother_tongue>Bangla</mother_tongue>
</speaker_data>
<content_meta_data>
<lang>Bangla</lang>
<domains>
        <domain>Novel (Beji-Weasel)</domain>
        <source>Beji (Weasel), Novel, Md. Japor Iqbal, Boi Mela</source>
        <domain>Legal text (Child)</domain>
        <source>D.net, Legal text, www.dnet-bangladesh.org/</source>
</domains>
</content_meta_data>
<text_metadata>
<lang>Bangla</lang>
<corpus_type>Read corpus</corpus_type><encoding>utf8</encoding>
</text_metadata>
<raw_annotation_tool>
<name>prrat</name><file_format>textgrid</file_format>
</raw_annotation_tool>
</metadata>
<data>
<sentences>
<sentence>
        <sent_id>10000</sent_id>
        <orthograph>স্কুলটির দিকে দৃষ্টি দিন</orthograph>
        <phonetic_form>skul.tir d̪i.ke d̪riʃ.ti d̪in </phonetic_form>
        <wav_id>10000.wav</wav_id>
</sentence>
```


Figure 5: XML Structure of corpus

4.3. Analysis on “Continuous speech corpus”

A small analysis was done on "Continuous speech corpus" to evaluate the corpus. For this reason a statistical analysis has been conducted. Table 3 shows the phoneme, bi-phone and triphone coverage in the corpus. Fig 5 shows the frequency coverage of syllable, phone and biphone[1] in speech corpus. According to the table 3, this corpus lacks 4 phonemes. Besides that it has only 18.11% coverage of diphones and 5.93% coverage of the triphones. The phonemes are four diphthongs (aja, ua, ue, uo). To the best of our knowledge there is no published account about the frequency analysis of Bangla phoneme inventory & phonotactic constrain. This basically limits us to evaluate the phonetic coverage of this corpus. However, we can assert that this corpus has a domain variety. Moreover, this analysis raises a few research questions such as the following:

---

[1] The word biphone and diphone are used interchangeably.

1. Whether this 18.11% diphone coverage will cover all phonetic space in Bangla or not.
2. Does it make any problem if we omit the diphthongs that are not available in this corpus when designing phonetically balanced corpus?
3. Analysis of biphone set that should not belong to the language e.g. phonotactic constrain.

Pattern type	Possible (unique)	Total in the corpus	Coverage
phone	65	61	93.84%
biphone	4,225	765	18.11%
triphone	2,74,625	16,301	5.93%

Table 3: Phone and biphone coverage in "continuous speech corpus".

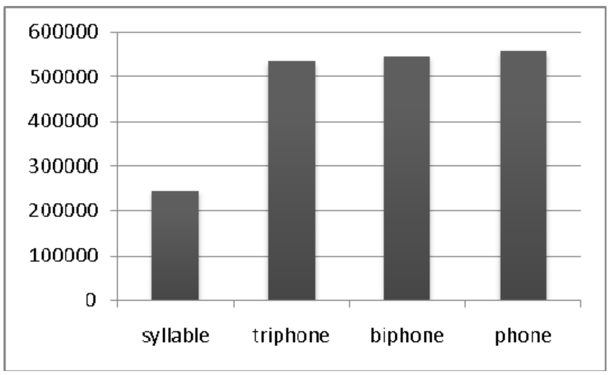


Fig 5: Frequency analysis

S/N	Syllable pattern	Frequency	Percentage
1.	cv	144668	59.0219
2.	cvc	67156	27.39842
3.	v	15460	6.307398
4.	vc	7592	3.097397
5.	ccv	4882	1.991767
6.	cvcc	2395	0.977116
7.	ccvc	2202	0.898376
8.	cvv	344	0.140346
9.	vcc	234	0.095468
10.	ccvcc	62	0.025295
11.	cccv	47	0.019175
12.	cccvc	25	0.0102
13.	cvccc	18	0.007344
14.	vv	18	0.007344
15.	ccvccc	4	0.001632
16.	cvvc	1	0.000408
17.	vccc	1	0.000408

Table 4: Syllable pattern and their frequency in the corpus

Table 4 shows different syllable patterns with their frequency available in the corpus. It is observed that, among these patterns some of the patterns formed from loan words. For example the patterns ccvcc, cvccc , cccvc and ccvccc are appeared in English loan words. Table 5, 6 and 7 shows a fragment of frequency analysis of phone, biphone and triphone.

Phone	Frequency	Percentage	Phone	Frequency	Percentage
a	59907	10.755431	ei	1897	0.340579
e	49200	8.833145	p^h	1671	0.300004
o	48754	8.753072	d	1333	0.239321
r	44180	7.931877	ai	1276	0.229087
i	36337	6.523780	t^h	1255	0.225317
n	32209	5.782658	oi	1094	0.196412
k	26974	4.842790	ã	1035	0.185819
ɔ	25885	4.647276	g^h	888	0.159427
t̪	22855	4.103283	ɟ^h	501	0.089947
ʃ	22126	3.972402	ui	484	0.086895
b	20154	3.618358	d^h	371	0.066608
l	16126	2.895189	ou	328	0.058888
m	14514	2.605778	iu	232	0.041652
u	12559	2.254786	eo	223	0.040036
d̪	12490	2.242398	õ	210	0.037702
p	12119	2.175790	ũ	194	0.034830
ɟ	11416	2.049577	ẽ	163	0.029264
j	9860	1.770220	eu	152	0.027289
t	8068	1.448492	ɔo	146	0.026212
h	7985	1.433591	ĩ	86	0.015440
g	6647	1.193372	ɔ̃	77	0.013824
c^h	6466	1.160876	au	55	0.009874
k^h	5467	0.981520	oa	39	0.007002
c	5254	0.943279	æ̃	38	0.006822
æ	5043	0.905397	io	31	0.005566
s	4246	0.762308	ie	11	0.001975
t̪^{h}	4054	0.727837	oe	6	0.001077
b^h	3508	0.629810	ia	2	0.000359
d̪^{h}	3389	0.608446	ea	1	0.000180
ŋ	3315	0.595160	æa	1	0.000180
ɾ	2086	0.374511			

Table 5: Frequency analysis of phoneme in the corpus

Biphone	Frequency	Percentage
a_r	11327	2.07416984
o_n	8205	1.50247758
e_r	7936	1.45321902
o_r	6518	1.19355867
a_n	5979	1.09485843
r_a	5625	1.0300349
r_o	5511	1.00915953
r_e	5264	0.96392955
ʃ_ɔ	5122	0.93792689
t̪_o	5025	0.92016451
n_i	4964	0.90899436
ɔ_r	4805	0.8798787
k_o	4781	0.87548389
n_a	4631	0.84801629
r_i	4591	0.8406916
k_a	4286	0.78484082
t̪_a	4241	0.77660054
a_k	4173	0.76414856
e_n	4170	0.76359921
n_e	3988	0.73027186

Table 6: Frequency analysis of biphone in the corpus

Triphone	Frequency	Percentage
k_o_r	3142	0.58706477
o_r_e	2049	0.382843957
p_r_o	1694	0.316514233
k_a_r	1687	0.315206323
b_o_l	1635	0.30549042
i_j_e	1564	0.292224475
d̪_e_r	1524	0.284750703
k_ɔ_r	1426	0.266439962
a_d̪_e	1386	0.258966191
t̪_a_r	1359	0.253921395
e_cʰ_e	1346	0.251492419
o_r_i	1339	0.250184509
n_e_r	1313	0.245326557
a_r_e	1243	0.232247457
o_n_e	1236	0.230939547

Table 7: Frequency analysis of triphone in the corpus

5. CONCLUSION AND FUTURE REMARKS

Here we described the development procedure of Bangla annotated read speech corpora and some statistics of analysis. Corpus building is a continuous process which includes annotation for prosody prediction and annotation in different levels such as word, syllable, biphone and phone level for phonetic research. This is not only required for phonetic research but also in speech applications. Future work includes the following:

5.1. Intonation model and unit selection voice

As mentioned earlier there is an existing Bangla speech synthesis system, which lacks the intonation model. There were two reasons to develop “Continuous speech corpus”. One concern was to develop an intonation model. The other reason was to develop a unit selection based speech synthesis system. Though there is no unique approach to design intonation, but the following two approaches consider a set standard to produce intonation for synthetic speech from corpus. One is to produce synthetic speech by concatenating waveform directly which is called unit selection based synthesis. In this process no signal processing is required so it preserves the quality of the original signal. The other approach is to generate intonation from ToBI label which could be generated from the corpus. Our research team is working on both unit selection based synthesis and concentrating on generating ToBI label from corpus.

5.2. Acoustic analysis

An extensive acoustic analysis [12], [13] has done on Bangla phonemes using “corpus for acoustic analysis”. However, a significant amount of work need to be done on prosody such as syllable, stress, F0 and accent. So this type of clean and high quality speech corpus will help in acoustic analysis and speech applications.

5.3. Speech technology applications

The multiple uses of these corpora are innumerable. Generating a phonetically balanced corpus is another step which could be done from these corpora. A phonetically balanced corpus is especially important for speech synthesis and speech recognition. Besides these applications, one can use this resource to do research on speaker identification, emotion extraction and spoken information extraction.

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), Naira Khan (Dhaka University) and BRAC University students who helped by providing their speech.

7. REFERENCES

[1] CDAC Bangla Speech Corpora, http://www.cdackolkata.in/html/txttospeeh/corpora/corpora_main/MainB.html, last accessed: December 10, 2009.

[2] Bangla Text to Speech Synthesis, http://crblp.bracu.ac.bd/demo/tts, last accessed: December 10, 2009.

[3] W. Zhu, W. Zhang, Q. Shi, F. Chen, "Corpus Building for Data-Driven TTS System," IEEE TTS Workshop, Santa Monica, 2002.

[4] Williams, Briony, "Levels of Annotation for a Welsh Speech Database for Phonetic Research", ISCA SIG SALTMIL, May-1998.

[5] D. Gibbon, R. Moore and R. Winski (eds.), Handbook of Standards and Resources for Spoken Language System. Berlin: Mouton de Gruyter.

[6] A. Nagy, P. Pesti, G. Németh, T. Bőhm: Design Issues of a Corpus-Based Speech Synthesizer, Hungarian Journal on Communications, 2005/6. special issue, pp. 18-24, Budapest, Hungary, 2005

[7] Florian Schiel, Christoph Draxler, Angela Baumann, Tania Ellbogen, Alexander Steffen, The Production of Speech Corpora, Version 2.5 : June 1, 2004

[8] Sinsuke Sakai, Building Probabilistic Corpus-based Speech Synthesis Systems from the Blizzard Challenge 2006 Speech Databases, Academic Center for Computing and Media Studies, Kyoto University

[9] Kishore S. Prahallad, Alan W Black, Rohit Kumar, Rajeev Sangal, Experiments with Unit Selection Speech Databases for Indian Language, Proc. National Seminar on Language Technology Tools: Implementations of Telugu, Hyderabad , India.

[10] Bangla Academy, *Bangla Academy Byabaharik Bangla Abhidhan*, 6th reprint 2005. Bangla Academy, Dhaka, 1992.

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

[12] Acoustic Analysis of Bangla Consonants, Firoj Alam , S. M. Murtoza Habib and Professor Mumit Khan, Proc. Spoken Language Technologies for Under-resourced language (SLTU'08), Vietnam, May 5-7, 2008, page 108-113.

[13] Firoj Alam, S.M. Murtoza Habib, Mumit Khan, *Research Report on Acoustic Analysis of Bangla Vowel Inventory*, Center for Research on Bangla Language Processing, BRAC University, 2008.

[14] Festvox, http://www.festvox.org/, last accessed: December 10, 2009.

[15] CRBLPConverter, http://crblp.bracu.ac.bd/converter.php, last accessed December 26' 2009.

[16] Sproat R., Black A., Chen S., Kumar S., Ostendorf M., & Richards C, "Normalization of Non-Standard Words: WS'99 Final Report", CLSP Summer Workshop, Johns Hopkins University, 1999, Retrieved (June, 1, 2008). Available: www.clsp.jhu.edu/ws99/projects/normal

[17] Firoj Alam, S. M. Murtoza Habib and Mumit Khan, Text Normalization System for Bangla, Accepted for present on the Conference on Language and Technology 2009 (CLT09), NUCES, Lahore, Pakistan, January 22-24, 2009.

[18] CRBLP pronunciation lexicon, http://crblp.bracu.ac.bd/demo/PL/, last accessed December 26' 2009

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

8. APPENDIX

[1] Shaptahik, article, http://www.shaptahik.com/arch_index.php?arc_day_id=2

[2] Beji (Weasel), Novel, Md. Japor Iqbal, Boi Mela

[3] D.net, Legal text, www.dnet-bangladesh.org/

[4] Bangladesh history: http://bn.wikipedia.org/wiki/ঢাকা, http://muktijuddho.wikia.com/wiki/বীর_শ্রেষ্ঠ, http://bn.wikipedia.org/wiki/বাংলাদেশ, http://bn.wikipedia.org/wiki/ভাষা_আন্দোলন, "History of 7 March, 1971", http://susanta.wordpress.com/2008/03/07/আবহাওয়া-নির্মল/, Wikipedia, , Wikipedia, http://bn.wikipedia.org/wiki/

[5] Somewhereinblog, Interview, http://www.somewhereinblog.net/blog/omipialblog/28803654

[6] Rupaly Dip, Humaun Ahmed.

[7] Prothom-alo, Editorial, http://www.prothom-alo.com/archive/headlines_mcat.php?dt=2008-07-24&issue_id=993&mid=Mw==

[8] Govt. of Bangladesh, Constitution of Bangladesh, April, 2008

[9] News, CRBLP-Prothomalo corpus.