A (Very) Brief Comparison of Bengali and English Phonological Systems

Man Ho Wong
m.wong@pitt.edu

LING 1878 Fall 2021
Department of Linguistics
University of Pittsburgh, PA, USA.

Revised: February 22, 2023

Introduction

Bengali (*Bangla-bhasha*) is an Indo-European language widely spoken in Bangladesh and neighboring areas in India. A language with a rich history of music and literature, Bengali is the mother language of many influential figures in South Asian arts and culture, such as Rabindranath Tagore and Kazi Nazrul Islam. The language also has a special place in the modern history of linguistic human rights: The language movement fighting for the official recognition of Bengali in former East Bengal during the 1950s has inspired the creation of International Mother Language Day by UNESCO to promote awareness of linguistic diversity (United Nations, 2021).

Bengali is a diglossic language. While the modern literary form, *Cholito-bhasha* (based on the Nadia dialect), is essentially the same across different Bengali communities, the colloquial forms of Bengali vary a lot in phonology that they can be grouped into six major dialects (Behrman, 2021). In this paper, I will describe the basic phonetics and phonology of the standard forms of Bengali (the Bengali dialect spoken in Dhaka, Bangladesh, and the West Bengali dialect spoken in the West Bengal region of India) and compare them with the phonological system of General American English. I will first analyze the consonant system, the vowel system, and the syllable structure. Bengali's notable prosody patterns will be discussed briefly in the last section.

To prevent misinterpretation of the speech analysis in this paper, I should remind the readers that the voice onset time of Bengali consonants, the F1/F2 frequencies of Bengali vowels, and the pitch contour were measured from a speaker of a West Bengali dialect. Although there are differences in phonology between different standard forms of Bengali, this should not affect the comparative analysis of Bengali and English.

Methods

Participant

A participant was recruited for the speech analysis in this study. She is a friend of mine who volunteered to record the audio used in this study. The participant is a native speaker of a West Bengali dialect, who is a female born in Varanasi, India, in the 1980s and lived in India until moving to the US in the 2010s. The participant completed all of her education in India and has a graduate degree.

Audio recording procedures

The audio recording was conducted remotely. The purposes of the study were briefly explained to the participant before the recording. The participant was instructed to record each of the words on the word list (**Table 2** for consonants; **Table 4** for vowels) in a quiet environment with minimal background noise and echo. For individual words, the English translation of a Bengali word was read before the production of the Bengali word, so that each Bengali word would have a similar environment (English was chosen here as suggested by the participant because the author is unfamiliar with Bengali and English helps the author to identify the Bengali words in the sound files). For pitch analysis, a short paragraph of Bengali text (see Sample text at the end of the paper) was recorded. A declarative sentence in the middle of the text was extracted for pitch analysis. The recording was done with a smartphone held in hand by the participant herself.

Speech Analysis

The audio files recorded in M4A format were converted to WAV format for analysis. The computer software Praat (Boersma, 2001) was used to annotate and measure the speech sounds in the files. Standard parameter settings in the software were used. Boundaries of phonemes were identified visually by inspecting the waveform and the spectrum of the recording. The onset of a word-initial voiced stop (if applicable) was defined as the onset of its voicing. The onset of vowel voicing was identified at the start of the periodicity in the waveform, after the release of the stop. First (F1) and second (F2) formants of vowels were identified visually from the spectrum of the recording.

Consonant system

Similar to American English, all consonants in Bengali are pulmonic. The consonant inventory of Bengali and English are of similar sizes, although both languages have unique sounds that are not found in the other language. In **Table 1,** consonants in Bengali are sorted according to their places and manners of articulation. Examples are given in **Table 2.** The major differences between Bengali and American English are highlighted in the following texts.

Table 1. Consonant phonemes of Bengali.

Consonant phonemes are arranged by the manner of articulation (rows) and place of articulation (columns). In the same field, the bottom consonants are aspirated; the right consonants are voiced. Table adapted with minor modifications from Khan (2010).

	Bilabial	Dental	Alveolar	Post-alveolar	Velar	Glottal
Nasal	m	n			ŋ	
Plosive	p b (pʰ) bʱ	t̪ d̪ t̪ʰ d̪ʱ	t d tʰ dʱ		k g kʰ gʱ	
Affricate				tɕ dʑ tɕʰ dʑʱ		
Fricative	(ɸ)		s	ʃ		h
Rhotic		r				
Lateral approximant		l				

Table 2. Examples of consonant phonemes.

Adapted from Khan (2010).

Consonant	Example	Consonant	Example
m	/amar/ ‘my’	k	/kaʃ/ ‘wild sugarcane’
n	/anar/ ‘for bringing’	k^h	/k^haʃ/ ‘you eat’
ŋ	/daŋar/ ‘of dry land’	g	/gaʃ/ ‘you sing’
p	/pati/ ‘grass mat’	$g^ɦ$	/$g^ɦ$aʃ/ ‘grass’
b	/bati/ ‘bowl’	tɕ	/tɕal/ ‘rice’
$b^ɦ$	/$b^ɦ$ati/ ‘kiln’	$tɕ^h$	/$tɕ^h$al/ ‘peel’
ʈ	/ʈak/ ‘shelf’	dʑ	/dʑal/ ‘net’
$ʈ^h$	/$ʈ^h$ak/ ‘let it be’	$dʑ^ɦ$	/$dʑ^ɦ$al/ ‘spicy’
ɖ	/ɖag/ ‘stain’	ɸ	/ɸaʃ/ ‘noose’
$ɖ^ɦ$	/$ɖ^ɦ$ak:a/ ‘shove’	s	/saɸ/ ‘clean’
t	/tak/ ‘bald spot’	ʃ	/ʃaʃ/ ‘breath’
t^h	/t^hakur/ ‘lord’	h	/haʃ/ ‘duck’
d	/dak/ ‘call!’	r	/raʃ/ ‘reduction’
$d^ɦ$	/$d^ɦ$ak/ ‘cover!’	l	/laʃ/ ‘corpse’

Plosives (stops) and affricates

Phonemic contrasts based on both voicing and aspiration are found in Bengali plosives and affricates. Therefore, there are four distinct phonemes for each set of plosives or affricates (e.g. post-alveolar plosives /t, d, t^h, $d^ɦ$/), except for bilabial plosives (explained below). In comparison, English plosives and affricates only contrast in voicing phonemically: voiced plosives and affricates are often devoiced in speech and therefore their distinctions lie in the degree of aspiration (e.g. / t^h, d/ are often realized as [t^h, t]). **Table 3** and **Figure 1** (right panel) show the voice onset time (VOT) of Bengali word-initial stops.

Table 3. Voice onset time (VOT) of Bengali word-initial stops.

The voice onset time (VOT) was measured from a native speaker of a West Bengali dialect.

Stop	VOT (ms)
/p/	11.219
/b/	-124.792
/bʱ/	-141.807
/ʈ/	10.699
/ʈʰ/	95.02
/ɖ/	-177.994
/ɖʱ/	-11.308
/t/	9.283
/tʰ/	106.536
/d/	-178.094
/dʱ/	-6.198
/k/	17.578
/kʰ/	138.547
/g/	-176.679
/gʱ/	-148.59

Unlike in English, unaspirated voiced stops /b, ɖ, d, g/ in Bengali are fully voiced as indicated by negative VOT. **Figure 1** (left panel) shows the presence of prevoicing in fully voiced stops /g, gʱ/. Unaspirated voiceless stops /p, ʈ, t, k/ have short VOT around 10-20 ms. Aspirated voiced stops / ɖʱ, dʱ, gʱ/ (except /bʱ/) have less negative VOT compared to unaspirated voiced stops. In contrast, aspirated voiceless stops / ʈʰ, tʰ, kʰ/ have much longer VOT compared to unaspirated voiceless stops.

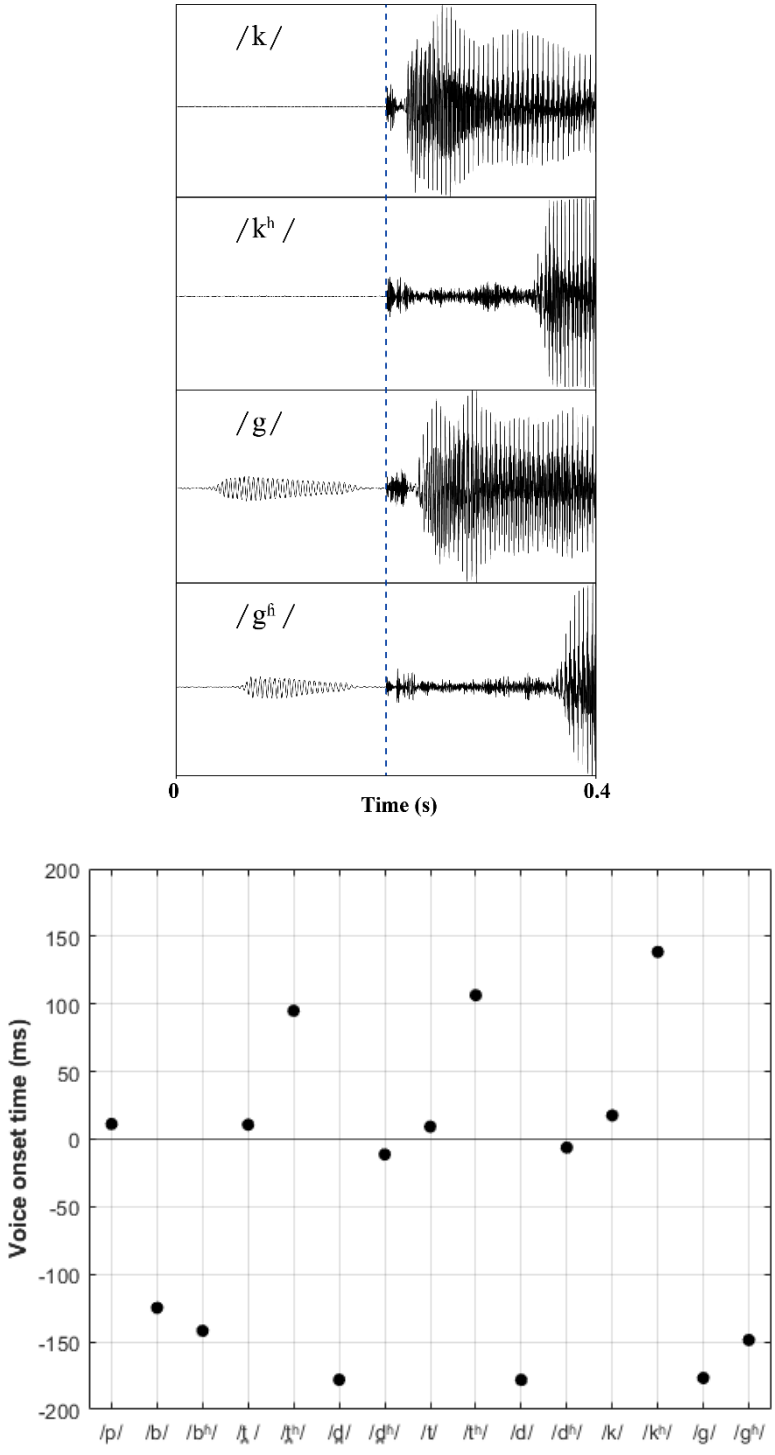

Figure 1. Voice onset time (VOT) of Bengali word-initial stops. Top: Sections of sample waves (400 ms long) containing /k, k^h, g, $g^ɦ$/ were extracted from the audio recording. The dotted line indicates the release of stops, where waves are aligned. Bottom: Data in Table 1 is plotted here for easier comparison.

Nasals

Both Bengali and American English have the nasal consonants /m, n, ŋ/. Note that nasalized vowels do not exist in the Standard Bangladeshi form of Bengali and therefore are not included in this paper.

Fricatives

Aspirated bilabial plosive /p^h/ can be realized as [p^h] or [ɸ~f] (Dasgupta, 2014; Khan, 2010). The closest English consonant to this sound is the voiceless labiodental fricative /f/. Unlike English, Standard Bengali does not have phonemic contrast based on voicing for fricatives (e.g. English /f/ vs. /v/ and /s/ vs. /z/).

Rhotic

In Bengali, /r/ can be realized as [r, ɾ or ɹ], depending on the speakers and the position of the sound (e.g. [ɾ] occurs word-medially and -finally) (Dasgupta, 2014; Khan, 2010). Both [ɾ] and [ɹ] are commonly found in American English, though [ɾ] is a realization of /t/ when it occurs word-medially (e.g. *better* ['bɛɾə˞]).

Approximants

Unlike English, Bengali has only one approximant phoneme /l/. However, /w/ and /j/ may occur as allophones of vowel phonemes /u, o/ and /i/ correspondingly (e.g. /de(w)ul/ ‘temple’ and /di(j)o/ ‘give’) (Dasgupta, 2014).

Vowel system

Monophthongs

Bengali has a relatively small vowel inventory compared to American English: There are seven monophthongs in Bengali, while American English has about twelve (the exact number depends on the dialect and definition of monophthongs). These seven vowels are presented in a vowel diagram in **Figure 2** (Khan, 2010) and **Figure 3** (measured from a native speaker of a West Bengali dialect)**.** A minimal set for the vowels is provided in **Table 4**.

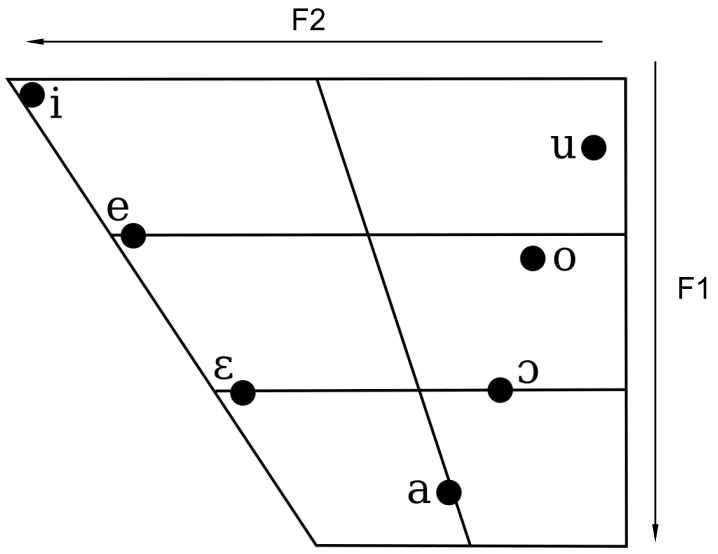

Figure 2. Monophthongs in Bengali. Vowels were recorded from six speakers and are arranged based on their F1 and F2 frequencies (increasing F1 frequency from top to bottom; increasing F2 frequency from right to left). Vowel diagram from Khan (2010).

Conventionally, the IPA symbols for the closest cardinal vowels are assigned to the corresponding vowels of a language. Therefore, some of the IPA symbols used in **Figure 2** might differ from other studies. To the best of my knowledge, most studies choose the IPA symbols [i, e, ɔ, o, u] for the phonemes /i, e, ɔ, o, u/. However, both [ɛ] and [æ], as well as both [a] and [ɑ] (and

sometimes [ɐ]) can be found in the literature to describe [ɛ ~ æ] and [a ~ ɐ ~ ɑ] respectively, since these vowels may be realized as different allophones by the speakers involved in different studies (Khan, 2010; Barman, 2011; Mostafa, 2013; Dasgupta, 2014).

Table 4. Minimal set for Bengali vowels.

Examples from Ferguson, C. A., & Chowdhury, M. (1960).

Vowel	Example
i	/k^{h}il/ 'waste, bolt'
e	/k^{h}el/ 'game'
ɛ	/k^{h}ɛl/ 'play!' (inf.)
a	/k^{h}al/ 'canal'
ɔ	/k^{h}ɔl/ 'bond'
o	/k^{h}ol/ 'shell'
u	/k^{h}ul/ 'open' (root)

All Bengali vowel phonemes can be found in American English. However, although Bengali and American English share some of the IPA vowel symbols, the actual position of each of the vowels in Bengali differs from that of its corresponding vowel (i.e. vowel denoted by the same IPA symbol) in American English (see **Figure 2** or **Figure 3**). For example, /i/ in Bengali has a higher F2 frequency (i.e. more advanced) than in English, and /a/ has a higher F1 frequency (i.e. more open). Nevertheless, such differences may not be noticeable to the untrained ear.

Some vowel phonemes in American English are absent in Bengali. These phonemes are the lax vowels /ɪ, ʊ/ and the central vowels /ə, ɚ, ʌ/. Bengali does not have phonemic contrasts based

on vowel tensing, although some of the vowels may have a lax or a tense allophone (e.g. /u/ may be realized as [u] or [ʊ] by different speakers). Bengali has only one central vowel /a/ (realized as [a ~ ɐ ~ ɑ]), although its allophones could be advanced or retracted as mentioned above.

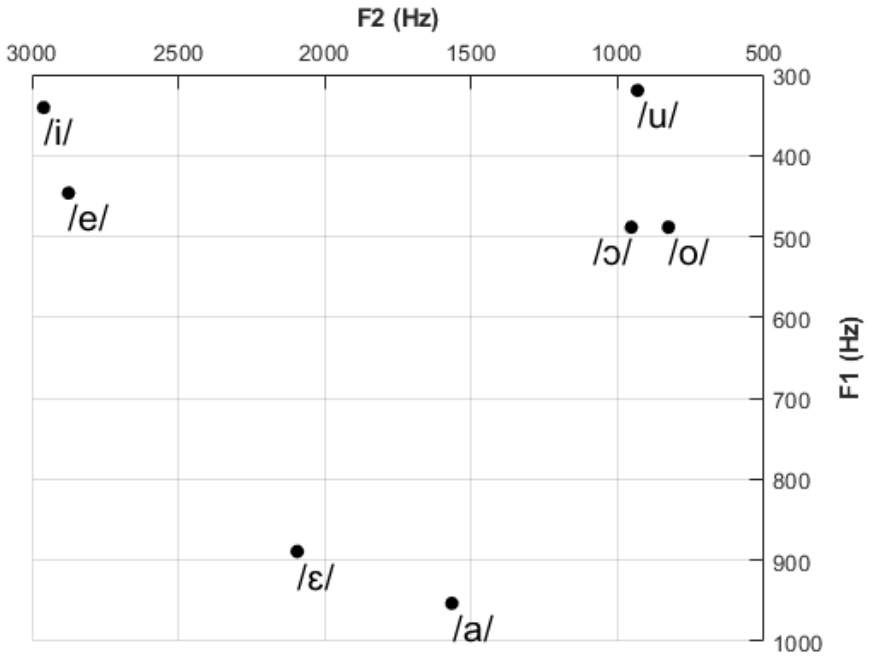

Figure 3. F1 and F2 frequency of Bengali vowels. Data shown here was measured from a native speaker of a West Bengali dialect. Note that /ɔ/ and /o/ are very close to each other. This could be due to variations between dialects or accents.

Diphthongs

While there are fewer vowels in Bengali than in English, the diphthong inventory of Bengali is remarkably larger than that of American English. The five diphthongs in American English are /aɪ, ɔɪ, aʊ, eɪ, oʊ/, though the last two diphthongs may be treated as monophthongs /e/ and /o/. In Bengali, a consensus on the exact number of diphthongs has not been established. The number can range from 17 to 31 in different studies (Barman, 2011). **Table 5** shows a list of 16 diphthongs that are commonly found in many studies.

Table 5. Bengali diphthongs and examples.

Table adapted from Khan (2010). *Note: /oo/ is pronounced [oo̯], where the second vowel is a semivowel. The diphthongs listed here may not be accepted by all linguists. For example, Chatterji (1986) did not include /oo/ as a diphthong in their paper.

Diphthong	Example
/iu/	[piu] (name)
/ei/	[nej] 'I take'
/eu/	[dew] 'wave'
/ɛe/	[nɛe̯] 'justice'
/ɛo/	[nɛo̯] 'you take'
/ai/	[paj] 'I find'
/ae/	[pae̯] 'finds'
/au/	[paw] 'sliced bread'
/ao/	[pao̯] 'you find'
/ɔe/	[ʃɔe̯] 'tolerates'
/ɔo/	[bɔo̯] 'you flow'
/oi/	[ʃoj] 'signature'
/oe/	[ʃoe̯] 'lies down'
/ou/	[bow] 'bride'
/oo/*	[ʃoo̯] 'you lie down'
/ui/	[ʃuj] 'needle'

Syllable structure

General structure

The most common syllable pattern in Bengali is CV (Kar, 2009). Similar to English, syllable coda is allowed in Bengali. Therefore, the syllable pattern CVC can occur in Bengali (e.g. /ʈ̪ak/ 'shelf') (Khan, 2010). Other possible syllable patterns are V, VC, VV, CVV, CCV, CCVC, CVVC, CCVV, CCVVC, CVCC, CCCV, CCCVC, VVC and CCCVV (Kar, 2009). Please note that not all patterns with consonant clusters occur natively in Bengali (see below).

Onset and final consonant clusters

Consonant clusters do not appear word-initially in native Bengali words. However, they are commonly found in loanwords such as those borrowed from Sanskrit or English (e.g. CCVC, such as /gram/ 'village'). Final consonant clusters occur very rarely in Bengali, and most of the words containing final consonant clusters are borrowed from English (Masica, 1991).

Medial consonant clusters

Medial consonant clusters appear frequently in Bengali (Mostafa 2013). In addition, gemination (doubling of consonants) can occur in almost every consonant (Khan 2010), such as /b^{ɦ}ago/ '(you) go away' and /b^{ɦ}aggo/ 'luck' (Mostafa 2013). Such a phenomenon also occurs in American English phonetically, although consonant length is not contrastive in English.

Prosody

Stress and trochee

Bengali is often described as "musical" partly due to its distinct stress and intonation pattern (Chatterji, 1921). One of the most notable features in Bengali is probably that many words are trochaic: Unlike in English where word-level stress can be found in all positions, the stress in Bengali often occurs in the first syllable (Chatterji, 1921). In contrast, stress in English is not fully predictable, but rather phonemic, and changing the position of the stress can change the meaning of a word. For example, *decrease* is a verb when pronounced /dɪˈkriːs/ but a noun when pronounced /ˈdiːkriːs/ (shifting stress from the second to the first syllable). Stress is mostly non-phonemic in Bengali and its patterns can be predicted by phonological rules related to stressing (Chatterji, 1921).

Intonation

In Bengali, most words in a simple declarative sentence are produced with a rising tone except for the final word of the sentence (**Figure 4**). Therefore, a typical sentence has words alternating between low and high tones and a falling tone at the end of the sentence (Hayes and Lahiri, 1991). Such a pattern does not exist in English, where rising tones occur typically at the end of yes-no questions.

As shown in **Figure 4**, the pitch of the sentence varies at the word level, and it ends with a falling tone (e.g. the last word in the example, *maryadā*). Almost all of the multisyllabic words here feature a rising tone towards the end of the word (e.g. *sādhāraṇata, bānlādēśī, bhāratīẏadēra, kāchē, tāṁra, raẏēchē, biśēsa*). This intonation pattern may contribute to the perception of Bengali as a musical-sounding language mentioned above.

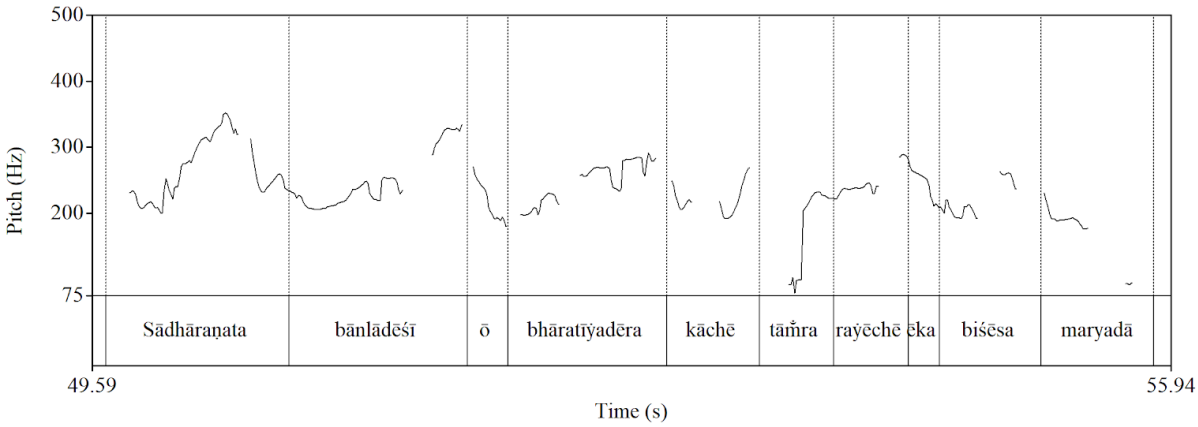

Figure 4. Pitch contour of a declarative sentence in Bengali. The participant was asked to read a short text in Bengali. The sentence with medium length and in the middle of the text is used as an example here. In this figure, the Bengali sentence is written in romanized script. See methodology for Bengali script and English translation.

Sample text for pitch analysis (Retrieved from Banglatip. 2021, October 26)

Bengali script and romanization
সব ক্লাসের বা শ্রেণির উপর ভিত্তি করে এখানে একটি সংক্ষিপ্তভাবে রচনা দেওয়া হলো, যেটি কবি রবীন্দ্রনাথ ঠাকুরকে কেন্দ্র করেই লিখা। Saba klāsēra bā śrēṇira upara bhitti karē ēkhānē ēkaṭi saṅkṣiptabhābē racanā dē'ōẏā halō, yēṭi kabi rabīndranātha ṭhākurakē kēndra karē'i likhā.
চলুন তাহলে সে সম্পর্কে জানা যাক। Caluna tāhalē sē samparkē jānā yāka.
ভূমিকা : কবি হিসেবে, খ্যাতি হিসেবে রবীন্দ্রনাথ ঠাকুরের রয়েছে অমূল্য কৃতি ও অর্জন। Bhūmikā: Kabi hisēbē, khyāti hisēbē rabīndranātha ṭhākurēra raẏēchē amūlya kr̥ti ō arjana.
এশীয় মহাদেশে কবি রবীন্দ্রনাথ ঠাকুরের রয়েছে এক বিশেষ মর্যাদা। Ēśīẏa mahādēśē kabi rabīndranātha ṭhākurēra raẏēchē ēka biśēṣa maryādā.
যদিও তিনি পৃথিবীব্যাপী বিশ্বকবি হিসেবে খেতাব পান,কিন্তু তাঁর লেখায় গল্প-কবিতায়, তিনি চিরকাল অমর হয়ে থাকবেন বাঙ্গালীদের মনে প্রাণে। Yadi'ō tini pr̥thibībyāpī biśbakabi hisēbē khētāba pāna,kintu tāṁra lēkhāẏa galpa-kabitāẏa, tini cirakāla amara haẏē thākabēna bāṅgālīdēra manē prāṇē.
যা কোনো দিন ক্ষয় বা ভুলে যাওয়ার নয়। Yā kōnō dina kṣaẏa bā bhulē yā'ōẏāra naẏa.
বাংলাদেশী ও ভারতীয়দের কাছে তাঁর রয়েছে এক বিশেস মর্যদা। Sādhāraṇata bānlādēśī ō bhāratīẏadēra kāchē tāṁra raẏēchē ēka biśēsa maryadā.
বাহ্যিকভাবে যদিও তিনি মাত্র ৮০ বছর বেঁচে ছিলেন, কিন্তু বাঙ্গালীদের মনে প্রাণে তিনি এখন অবধি বেঁচে রয়েছেন এব ভবিষ্যৎও তিনি তাঁর লেখার মাধ্যমে চির অম্লান হয়ে থাকবে। Bāhyikabhābē yadi'ō tini mātra 80 bachara bēṁcē chilēna, kintu bāṅgālīdēra manē prāṇē tini ēkhana abadhi bēṁcē raẏēchēna ēba bhabiṣyaṯō tini tāṁra lēkhāra mādhyamē cira amlāna haẏē thākabē.
তার লিখে যাওয়া লেখাগুলো এখনো বাংলা সাহিত্যের বিভিন্ন শাখাগুলোকে উজ্জীবিত রেখেছে। Tāra likhē yā'ōẏā lēkhāgulō ēkhanō bānlā sāhityēra bibhinna śākhāgulōkē ujjbībita rēkhēchē.
বাংল সাহিত্য, ভাষা ও সংস্কৃতিতে সমাসীন রয়েছেন তাঁর লিখে যাওয়া লেখাগুলো। Bānla sāhitya, bhāṣā ō sanskr̥titē samāsīna raẏēchēna tāṁra likhē yā'ōẏā lēkhāgulō.

English translation
Here is a brief essay based on all the classes, which is centered on the poet Rabindranath Tagore. Let's find out now. Introduction: As a poet, Rabindranath Tagore has invaluable achievements. Poet Rabindranath Tagore has a special status in the Asian continent. Although he is regarded as a world poet, in his stories and poems he will remain immortal in the minds of Bengalis forever. Which is not to be eroded or forgotten any day. He has a special status among Bangladeshis and Indians in general. Outwardly, although he lived only 70 years, in the minds of Bengalis he is still alive and the future will be immortalized through his writings. His writings still enliven the various branches of Bengali literature. His writings are contemporary in Bengali literature, language and culture.

References

Banglatip. (2021, October 26). Essay on Rabindranath Tagore in Bengali - রবীন্দ্রনাথ ঠাকুর রচনা. Retrieved December 12, 2021, from https://banglatip.com/essay-on-rabindranath-tagore-in-bengali

Barman, B. (2011). A contrastive analysis of English and Bangla phonemics. Dhaka University Journal of Linguistics. 2. 10.3329/dujl.v2i4.6898.

Behrman E., Santra A., Sarkar S., Roy P., Yadav R., Dutta S., Ghosal A. (2021). Dialect Identification of the Bengali Language. In Tyagi A. K. (Ed.). *Data Science and Data Analytics,* 1st ed., pp. 142-170. Chapman and Hall/CRC.

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. Glot International 5:9/10, 341-345.

Chatterji, S. (1921). Bengali Phonetics. Bulletin of the School of Oriental and African Studies, 2(1), 1-25. doi:10.1017/S0041977X0010179X

Chatterji, S. K. (1986). A Bengali phonetic reader. Calcutta: Rupa.

Dasgupta, P. (2014). Bangla. In Cardona, G. and Jain, D. (Eds.). *The Indo-Aryan Languages,* pp. 386-428. Routledge.

Ferguson, C. A., and Chowdhury, M. (1960). The Phonemes of Bengali. Language, 36(1), 22–59. doi:10.2307/410622
Hayes, B. and Lahiri, A. (1991). Bengali intonational phonology. Nat Lang Linguist Theory 9, 47–96. doi:10.1007/BF00133326

Kar, Somdev (2009). The syllable structure of Bangla in Optimality Theory and its application to the analysis of verbal inflectional paradigms in Distributed Morphology (Doctoral dissertation or Master's thesis, Universität Tübingen, Tübingen, Germany). Retrieved from https://d-nb.info/994257015

Khan, S. (2010). Bengali (Bangladeshi Standard). Journal of the International Phonetic Association, 40, pp. 221-225. doi:10.1017/S0025100310000071

Masica, C. P. (1991). The Indo-Aryan Languages, pp. 126. Cambridge University Press.

Mostafa, T. (2013). A Contrastive Analysis between Bangla and English Phonology: Some Pedagogical Recommendations. European Conference on Language Learning, Brighton, UK.

United Nations. (2021). International mother language day. United Nations. Retrieved November 17, 2021, from https://www.un.org/en/observances/mother-language-day