# VOICE ONSET TIME OF NATIVE- AND NONNATIVE-ENGLISH SPEAKERS: A CASE STUDY OF BENGALI SPEAKERS

Kaoru Tomita

*Yamagata University, Faculty of Literature and Social Sciences, Kojirakawa-machi, Yamagata, JAPAN*
*email: kaoru@human.kj.yamagata-u.ac.jp*

Discrimination of voiced or unvoiced sounds is made at phonological level and at the same time, it works as phonetic levels. An acoustic feature that works to discriminate voiced and unvoiced sounds is proposed by Abramson, et al (1970, 569), and it is called as voice onset time (VOT). That is the time between points of burst and beginning of vocal cord vibration. Voiced sounds hold short VOTs or minus ones. That is, vocal cord vibration has begun before burst starts for some voiced sounds. Unvoiced sounds hold long VOT.

VOTs are values that can be used to know which sounds are voiced or unvoiced sound. They could be factors to decide voiced or unvoiced sounds but they are not the only factors. Formant 0 values are to be used to construct voiced or unvoiced sounds. The formers, however, are used for measurement of voiced or unvoiced sound for this study, as they are major factors to decide voiced or unvoiced sounds.

As phonological level, Bengali has many numbers of discrimination for voiced or unvoiced sounds than English. It holds /$p^h$/-/p/-/$b^h$/-/b/ for plosive bilabial sounds. On the bases of this special phonological feature, it is predicted that VOT for unvoiced sounds of Bengali without aspiration, for example /p/, would shorter than English as Bengali speakers need to spare longer VOT for aspirated unvoiced sounds, such as /$p^h$/ or /$t^h$/. Besides that, it is predicted that VOT for voiced sounds without aspiration would shorter than that of English as Bengali speakers need to spare longer VOT for aspirated voiced sounds, such as /$b^h$/ or /$d^h$/.

## 1. Introduction

Phonological features play an important role for acquiring languages. It could be seen that this is the case especially when people are learning foreign languages. Among several language features, such as phonological ones, syntactic ones, semantic ones and pragmatic ones, the first ones is claimed to affect processes of language learning for long periods. Foreign accented pronunciation is found in speech produced by even advanced learners.

Some features that are proposed as universal ones are acquired or learned by every speaker of world languages. In learning languages as native ones, some of these features are acquired at earlier stages of learning and other features are acquired at later stages. On the other hand, this order for acquisition does not reflect on learning foreign languages exactly the same ways. In learning languages as foreign ones, some of the features that look simple and are supposed to be acquired at earlier stages, are not learned very well even at later stages. They may remain not to be acquainted very well and learners usually make mistakes on these features repeatedly for the rest of their lives.

Phonological features are divided into two main groups. One is a prosodic feature, such as stress, tone, intonation and duration. The other is a segmental feature, such as consonants and vowels. Prosodic features are used to classify world languages. Some languages discriminate words with different tone levels and others do not. Some languages discriminate words with different durations and others do not. Segmental features are also referred to for classification of world languages. For example, some languages have sound systems of small number of vowels and others have large number of ones. Much the same is true on consonant systems. All over the world, we can find around two hundred types of vowels and six hundred types of consonants. Their quality differences and the number of vowels and consonants each language hold make up its original vowel systems and consonant systems.

Consonant systems are mainly divided into a manner of articulation and a place of articulation. Besides these two factors, the other one, voiced or unvoiced is supposed to be a main phonological feature. Of course there are languages that do not hold symmetry for voiced and unvoiced sounds, such as Arabic and Chinese. Arabic has voiced plosive bilabial sound, /b/, and it does not have unvoiced plosive bilabial one, /p/. It has unvoiced plosive velar sound, /k/, and it does not have voiced plosive velar one, /g/. Chinese has two unvoiced plosive bilabial sounds, /p/ and /pʰ/, and it does not have voiced plosive bilabial sound /b/.

Different languages have different consonant systems for voiced and unvoiced discrimination. English discriminate /p/-/b/, /t/-/d/, /k/-/g/, /tʃ/-/dʒ/, /f/-/v/. /θ/-/ð/, /s/-/z/ and /ʃ/-/ʒ/. Bengali discriminate /pʰ/-/p/-/bʰ/-/b/, /tʰ/-/t/-/dʰ/-/d/, /ʈʰ/-/ʈ/-/ɖʰ/-/ ɖ/, /kʰ/-/k/-/gʰ/-/g/,  /tʃʰ/-/tʃ/-/dʒʰ/-/dʒ/, and /s/-/z/.

Discrimination of voice or unvoiced sounds is made at phonological level and at the same time, it works as phonetic levels. An acoustic feature that works to discriminates voiced and unvoiced sounds is proposed by Abramson, et al (1970, 569), and it is called as voice onset time (VOT). This is the time between points of burst to beginning of vocal cord vibration. Voiced sounds hold short VOT or minus one. That is, vocal cord vibration has begun before burst starts for some voiced sounds. Unvoiced sounds hold long VOT.

VOTs are values that are to be used to know which sounds are voiced or unvoiced sound. They could be factors to decide voiced or unvoiced sounds but they are not the only factors. Formant 0 values are to be used to construct voiced or unvoiced sounds. The formers, however, are used for measurement of voiced or unvoiced sound for this study, as they are major factors to decide voiced and unvoiced sounds.

As a phonological level, Bengali has many numbers of discrimination for voiced and unvoiced sounds than English does. It holds /pʰ/-/p/-/bʰ/-/b/ for plosive bilabial sounds. On the bases of this special phonological feature, it is predicted that VOT for unvoiced sounds of Bengali without aspiration, for example /p/, would shorter than that of English as Bengali speakers need to spare longer VOT for aspirated unvoiced sounds, such as /pʰ/ or /tʰ/. Besides that, it is predicted that VOT for voiced sounds without aspiration would shorter than that of English as Bengali speakers need to spare longer VOT for aspirated voiced sounds, such as /bʰ/ or /dʰ/.


## 2.   Methods

VOT values for /p/, /t/ and /k/ produced by native Bengali speakers are predicted to be shorter than the ones produced by native English speakers. VOT values for /b/, /d/ and /g/ produced by native Bengali speakers are predicted to be longer than the ones produced native English speakers.

### 2.1   Proceedings

English words that start with /p/, /t/, /k/, /b/, /d/ or /g/ are read by native English speakers and Bengali learners of English.  Bengali words that start with /p/, /t/, /k/, /b/, /d/ or /g/ are read by native Bengali speakers.

### 2.1.1  A case study 1

A word sets used for case study 1 is listed in Table 1

Table 1 A word set used for a case study 1

|  | unvoiced | voiced |
|---|---|---|
| bilabial | pat | bat |
| alveolar | tab | dab |
| velar | cat | gat |

Each stimulus item is printed in a carrier phrase, I said "      ". or Did you say "      "?. It is composed into a dialog presented below:

Table 2  A dialog used for a case study 1

| A: What did you say? |
|---|
| B: I said "pat". |
| A: Did you say"bat"? |
| B: I said "pat". |

### 2.1.2  A case study 2

English word sets used for a case study 1 and new Bengali word sets are prepared for an experiment of a case study 2. Bengali words are presented in Table 3.

Table 3 A word set used for a case study 2

|  | Voiceless | voiced |
|---|---|---|
| bilabial | পন<br>/pola/<br>*read* | বন<br>/bola/<br>*forest* |
| alveolar | টাকা<br>/taka/<br>*money* | ঢাকা<br>/daka/<br>*call* |
| velar | কলা<br>/kola/<br>*banana* | গলা<br>/gola/<br>*neck* |

Each stimulus item is printed in a Bengali carrier phrase, that has the same meaning of English phrase, I said "    ". or Did you say "    "?. It is composed into a dialog presented below:

Table 4 A dialog used for a case study 2

| | | | |
|---|---|---|---|
| A: | তুমি | কি | বলেছ ? |
| | *you* | *what* | *said* |
| B: | আমি | বলেছি | " কলা ", |
| | *I* | *said* | *banana* |
| A: | তুমি | কি | বলেছ "গলা" ? |
| | *you* | *what* | *said* *neck* |
| B: | আমি | বলেছি | "কলা", |
| | *I* | *said* | *banana* |

## 2.2 Subjects

### 2.2.1 A case study 1

Three male and three female speakers of American English (EM or EF) who majored in Japanese cultural studies participated in the experiment. They came from Oklahoma, U.S.A. as exchange students with one year term. They were from 22 to 24 years old.

### 2.2.2 A case study 2

One male and one female speakers of Bengali (BM or BF) who majored in technology participated in the experiment. They came from Dhaka, Bangladesh as exchange researchers with two years term. They were 31 and 32.

## 2.3 Recording procedure

Recording was conducted in a dyad. Each speaker was presented with test sheets and asked to clearly produce test items ten times.

## 2.4 Measurements

Time between points of burst to beginning of vocal cord vibration was measured with Praat.

## 3. Results

## 3.1 Statistics for a case study 1

Results of VOTs for English /p/ and /b/ by native English speakers are presented in Table 5.

Table 5 Consonant discrimination for English /t/ and /d/ by native English speakers [msec.]

|  | tab (S.D.) | dab (S.D.) | N | Mean | Z | P | Comparison |
|---|---|---|---|---|---|---|---|
| EM1 VOT | 112.93(17.63) | 12.70(5.58) | 30 | 62.81 | -4.78 | <.000 | /d/</t/ |
| EM2 VOT | 97.03(29.41) | 12.97(2.89) | 30 | 55.00 | -4.83 | <.000 | /d/</t/ |
| EM3 VOT | 46.00(15.04) | 11.77(2.71) | 30 | 28.88 | -4.78 | <.000 | /d/</t/ |
| EF1 VOT | 101.67(15.85) | 11.80(2.28) | 30 | 56.73 | -4.78 | <.000 | /b/</p/ |
| EF2 VOT | 102.17(30.37) | 11.53(2.25) | 30 | 56.85 | -4.78 | <.000 | /b/</p/ |

## 3.2 Statistics for a case study 2

Results of VOTs for English /p/ and /b/ by native Bengali speakers are presented in Table 6.

Table 6 Consonant discrimination for English /t/ and /d/ by non-native speakers [msec.]

|  | /t/ (S.D.) | /d/ (S.D.) | N | Mean | Z | p | Comparison |
|---|---|---|---|---|---|---|---|
| BM1 VOT | 20.57(6.86) | 20.33(8.36) | 30 | 20.45 | -.98 | NS | /d/, /t/ |
| BF1 VOT | 33.33(7.80) | 27.17(8.30) | 30 | 30.25 | -2.96 | <.003 | /d/</t/ |

By BM, /t/ and /d/ are not discriminated very well.

## 3.3 Statistics for distance between unvoiced and voiced sounds

Discrimination of unvoiced and voiced sounds is measured with distance between these two sounds. Table 7 presents mean distance between unvoiced and voiced sounds of English words.

Table 7 Mean distance between unvoiced and voiced sounds [msec.]

|  | EM1 | EM2 | EM3 | EF1 | EF2 | BM | BF |
|---|---|---|---|---|---|---|---|
| p-b | 99.50(18.10) | 65.63(19.02) | 36.00(10.61) | 69.55(14.85) | 96.83(25.13) | 7.60(5.48) | 7.73(5.99) |
| t-d | 84.41(29.84) | 34.23(15.49) | 100.23(19.09) | 89.87(16.75) | 90.63(29.72) | .23(4.25) | 6.17(11.62) |
| k-g | 71.69(27.60) | 32.07(16.42) | 104.17(28.25) | 48.50(11.57) | 67.60(32.11) | 9.13(8.83) | 30.00(16.61) |

table continued

| Mean | F-value | p-value | Comparison |
|---|---|---|---|
| 54.69 | 180.72 | <.000 | BM, BF <EM3<EM2, EF1<EF2, EM1 |
| 57.96 | 139.30 | <.000 | BM, BF<EM2<EM1<EF1, EF2, EM3 |
| 51.86 | 63.61 | <.000 | BM<BF<EM2, EF1<EF2, EM1<EM3 |

Results show that native English speakers discriminate unvoiced and voiced sounds clearly but Bengali learners of English do not discriminate them very well.

## 3.4 Statistics for sounds produced by English and Bengali speakers

English unvoiced sounds produced by native English speakers and native Bengali speakers and Bengali unvoiced sounds produced by native Bengali speakers are listed in Table 8.

Table 8 Mean VOTs for unvoiced sounds [msec.]

|  | EM1 | EM2 | EM3 | EF1 | EF2 | BM | BF |
|---|---|---|---|---|---|---|---|
| p | 77.14(16.54) | 45.07(10.24) | 103.80(18.35) | 77.43(14.16) | 103.90(25.03) | 17.73(4.69) | 19.80(4.17) |
| t | 97.38(29.87) | 46.00(15.04) | 112.93(17.63) | 101.67(15.85) | 102.17(30.37) | 20.57(6.86) | 33.33(7.80) |
| k | 93.38(27.04) | 55.20(17.03) | 127.93(25.29) | 71.93(10.63) | 89.83(32.55) | 33.97(5.72) | 53.03(15.04) |

table continued

| BMB | BFB | Mean | F-value | p-value | Comparison |
|---|---|---|---|---|---|
| 18.80(3.89) | 19.97(7.32) | 53.73 | 225.27 | <.000 | BM, BMB, BF, BFB<EM2<EM1, EF1<EM3, EF2 |
| 16.00(2.70) | 14.40(3.84) | 60.49 | 175.92 | <.000 | BFB, BMB, BM<BF, EM2<EM1<EF1, EF2, EM3 |
| 46.30(6.34) | 41.27(13.79) | 68.09 | 75.79 | <.000 | BM<BFB, BMB, BF, EM2<EF1<EF2, EM1<EM3 |

English voiced sounds produced by native English speakers and native Bengali speakers and Bengali unvoiced sounds produced by native Bengali speakers are listed in Table 9.

Table 9 Mean VOTs for voiced sounds [msec.]

|  | EM1 | EM2 | EM3 | EF1 | EF2 | BM | BF |
|---|---|---|---|---|---|---|---|
| b | 10.93(4.32) | 9.07(1.68) | 4.30(2.35) | 7.50(2.22) | 7.07(1.59) | 10.13(4.45) | 12.07(4.17) |
| d | 12.97(2.94) | 11.77(2.71) | 12.70(5.58) | 11.80(2.28) | 11.53(2.25) | 20.33(8.36) | 27.17(8.30) |
| g | 21.69(4.47) | 23.13(5.33) | 23.77(9.22) | 23.43(2.95) | 22.23(5.85) | 24.83(6.20) | 23.03(8.66) |

table continued

| BMB | BFB | Mean | F-value | p-value | Comparison |
|---|---|---|---|---|---|
| 12.57(3.93) | 7.80(3.33) | 9.04 | 19.42 | <.000 | EM3<EF2, EF1, BFB, EM2<BM, EM1, BF, BMB |
| 8.97(1.62) | 9.07(2.79) | 14.03 | 45.86 | <.000 | BMB<BFB, EF2, EM2, EF1, EM3, EM1<BM<BF |
| 33.40(5.89) | 25.63(6.61) | 24.56 | 9.05 | <.000 | EM1, EF2, BF,EM2, EF1, EM3, BM, BFB<BMB |

Results show that for both unvoiced and voiced sounds, Bengali learners of English produce English consonants that sound like their Bengali consonants.

## 4.  Discussion and conclusion

A predictions that VOTs for unvoiced sounds of Bengali without aspiration would shorter than that of English is proved with showing data obtained from the language experiment conducted for Bengali learners of English and native English speakers. Prediction that VOTs for voiced sounds without aspiration would shorter than that of English is also proved. The subjects of the experiment for this study present their tendency for producing English unvoiced or voiced wounds with their own Bengali unvoiced or voiced sounds.

## REFERENCES

1  Abramson, A.S., and Lisker, L. Discriminability along the voicing continuum: Cross-language tests. In *Proceedings of the Sixth International Congress of Phonetic Sciences*. Prague: Academia, 569-573. (1970).

2  Reets, H. andJongman, A. *Phonetics: Transcription, Production, Acoustics, and Perception*. Wiley-Blackwell. (2009).

3  Ioup, G. Exploring the role of age in the acquisition of a second language phonology. J.G.H. Edwards and M.L.Zampini (Eds.) *Phonology and Second Language Acquisition*. Philadelphia: John Benjamins Publishing Company. (2008).