

Role of markedness in perception of Bengali stops

Sreeparna Sarkar*

Abstract. Markedness is a theory that was developed based on segmental patterns observed in speech output and has primarily been addressed in regard to speech production in previous studies. According to the Markedness Theory, marked segments are more difficult to produce due to an additional property or “mark” which require more articulatory effort. However, its effects on speech perception is not discussed in the previous literature. This study examines the role of markedness in perception with Bengali stops. Bengali stops involve two types of markedness or additional properties, voicing and aspiration. Voiced stops (represented as D) are marked with respect to voiceless stops (represented as T) and aspirated stops (TH) are marked with respect to unaspirated stops (T). Voiced aspirated stops bear both the additional properties (represented as DH). While the absence of a marked property may make segments easier to produce than those with the property, the question addressed here is whether the same holds true for perception. This study investigates the possibility that the opposite is what is observed. That is the presence of additional properties may make segments more audible and identifiable. Additionally, this study also investigates whether the combination of multiple marks, in the Bengali DH stops, lead to a cumulative effect on perception with the best perceptual results for the DH stops. The results from this study show that this is in fact the case.

Keywords. speech perception; Markedness Theory; Bengali stops; voicing; aspiration.

1. Introduction. In the classical view of markedness, marked segments bear a “mark” or an extra property with respect to the corresponding unmarked segments (Trubetzkoy 1939; Jakobson 1941). For instance, some examples of unmarked vs marked pairs are oral vs nasal vowels and voiceless vs voiced obstruent. According to the markedness theory, these pairs are not dual oppositions, but one member of the pair bears an extra property compared to the other. An implication of this might be that the extra property may make the marked segments somewhat harder to produce requiring more articulatory effort than the unmarked ones. The unmarked segments are also said to have a broader distribution within and across languages due to more complexity in their production and are also acquired later in language acquisition (Jakobson 1941; Carlisle 1988). What is less clear, however, is whether the mark affects speech perception in the same way as it does speech production. Since the markedness theory has been mostly studied from the perspective of speech production in the past, its effect on perception is less understood. The reason this is an interesting question is because the additional properties associated with the marked segments may provide additional cues to the listeners for correct identification of those sounds, even if they are harder to produce. Therefore, there are two possible expected patterns for the effect of markedness on perception. One possibility is that markedness may make a sound more difficult to perceive following the same pattern of

* I would like to thank Dr. Irene Vogel for her comments and feedback and the members of the Phonology Lab at the University of Delaware. Author: Sreeparna Sarkar, University of Delaware (sree@udel.edu).

production. The other possibility is that markedness makes sounds more perceptually salient due to the extra additional properties which provide additional acoustic information to the listeners and make them more audible and easily identifiable with higher perceptual accuracy.

The current study tests the second possibility, that is, the presence of additional properties makes segments more audible and easily identifiable with higher perceptual accuracy compared to unmarked sounds. The language of investigation in this study is Bengali. Bengali provides a situation where the stops can be studied in relation to the markedness theory predictions. Therefore, the current study experimentally tests the effect of markedness on perception with the Bengali four-way stops.

2. Bengali stops as sets of marked and unmarked pairs. Bengali is an Indo-Aryan language spoken in India and Bangladesh. Like many other Indo-Aryan languages that have descended from Sanskrit, Bengali exhibits the rich four-way stop contrast involving the combination of two kinds of extra properties, voicing and aspiration. The voiced stops are marked with respect to voiceless stops and aspirated stops are marked with respect to unaspirated stops. The resulting four categories of stops as shown in Table 1. The voiced unaspirated stops (henceforth represented as D) and voiceless aspirated stops (henceforth represented as TH) have one additional mark compared to the voiceless unaspirated stops (henceforth represented as T). The voiced aspirated stop (henceforth represented as DH) category has two marks compared to the voiceless unaspirated stops.

	Voiceless	Voiced
Unaspirated	Voiceless Unaspirated	Voiced Unaspirated
Aspirated	Voiceless Aspirated	Voiced Aspirated

Table 1. Table showing four categories of Bengali stops

The marks increase incrementally in two steps from the unmarked voiceless unaspirated stops to the voiced aspirated stops. The increment in markedness is shown in Table 2.

	Voicing	Aspiration	Resulting Stop Type
Voiceless Unaspirated (T)	-	-	T
Voiced Unaspirated (D)	+ voicing		D
Voiceless Aspirated (TH)		+aspiration	TH
Voiced Aspirated (DH)	+voicing	+ aspiration	DH

Table 2. Table showing Bengali stops with increments in markedness properties

The resulting sets of marked and unmarked pairs are:
Voiceless (T, TH) vs Voiced (D, DH)
Aspirated (TH, DH) vs Unaspirated (T, D)

Additionally, the distribution of the stops in Bengali also correlates with the predicted pattern of markedness theory, that is, the DH category with the most marks have the most limited distribution in the language. They do not appear in every syllable position. For instance, they are available in the initial syllable of some words, but their availability decreases in the second syllable and are almost non-existent in third syllables. This is followed by the TH category with slightly more availability in the second syllable but almost no availability in third syllables. then D and finally the T category which has a broad distribution in the language (Fergusson and Choudhury 1960).

The stops are available in various places of articulation, bilabial, dental, post-alveolar, velar. All the available places of articulation are included in this study. The alveolar place of articulation in capital letters shown in Table 1 and 2 are for representation purposes and does not mean that the stops included in this study are limited to only one place of articulation.

3. Research questions and hypotheses. The Bengali stop perception experiment that was designed to investigate the effect of markedness on the perception is based on the four-types of Bengali stops that provide a situation to compare marked and unmarked pairs. The experiment addressed three main research questions. First, does voicing as a markedness property increase the perceptual accuracy of the stops? Second, does aspiration as a markedness property increase the perceptual accuracy of the stops? And third, does the combination of voicing and aspiration show a cumulative effect with the voiced aspirated category showing the best perceptibility? The study tested the following three hypotheses in order to address these questions. The first and the second hypotheses address the question whether stress and focus individually increase perceptual accuracy or not. The third hypothesis addresses the question whether stress and focus have a cumulative effect on the perception of the stops resulting in even higher perceptual accuracy than what stress and focus have individually.

Hypothesis 1: Enhanced articulation in syllables with lexical stress leads to better perceptual results in that position compared to non-stressed syllables.

Hypothesis 2: Enhanced articulation in focus leads to better perception results for syllables produced in a focus position.

Hypothesis 3: There is higher success in perception of initial syllable of focused words in Bengali due to a cumulative effect of stress and focus on the syllable.

3.1. PARTICIPANTS. That participants of this experiment were 50 native speakers of Bengali, 16 male and 34 female. The participants were aged between 18-32, born and raised in West Bengal with Bengali as their first language. Given that most people in India are also able to speak Hindi and English, all the participants in this study were also fluent in Hindi and English. Since the experiment required the participants to be able to identify the CV syllable written in Bengali script participants were required to be able to read the Bengali script fluently. Everyone who participated in this study were expert readers of the Bengali script. Additionally, all the participants were university educated working professionals. All the participants were residing in India at the time of their participation.

3.2. STIMULI DESIGN AND SYLLABLE EXTRACTION PROCEDURE. The stimuli for the perception experiment were created by extracting CV syllables from previously recorded three syllable real Bengali words CVCVCV produced by 9 native speakers of Bengali. The onsets of the CV syllables were the four kinds of Bengali stops i.e., voiceless unaspirated (T), voiceless aspirated (TH), voiced unaspirated (D) and voiced aspirated (DH) stops from varying places of articulation. The vowel in all cases were the vowel /a/. Example (1) below shows a three-syllable real Bengali word.

- (1) /paṭata/
'The leaf'

The first and second syllables from these words were extracted for this study. For instance, in (1), the syllables that were extracted were /pa/ and /ṭa/. The third syllable was not included in this study. One of the reasons why the third syllable was not included in the study is that the DH stops are not available as onsets in the third syllable. Interestingly, this follows the predicted pattern of distribution of marked segments in the markedness theory. The most readily available stop type as the onset in the third syllable is the T stop type.

The CV syllables were extracted using PRAAT (Boersma and Weenink 2018). For the extraction procedure, both the wave form and the spectrogram were examined in each case. The target syllables were identified as the string from the start of the consonant to the end of the vowel. The start of the C was considered to be the start of the stop closure, and the end of the CV syllable was the end of the /a/ vowel, i.e., where the periodic structure of the vowel /a/ ended. In the case of the stop onset in the initial syllable, in some cases there were long pauses after the end of the previous word, making it impossible to detect the beginning of the stop closure, especially for the voiceless stops. In such cases, the stop closure was identified as 80ms to the left of the point of release into the vowel. This duration, while only an estimate, was used as it was a typical duration for closures in cases without long initial pauses. Since the durations of the segments themselves were not of concern here, such an estimation was deemed adequate to provide whatever information was needed for the perception of the syllables. An example of the extraction process for the word '**paṭata**' is shown in Figure 1. The first target syllable, /pa/ in this case, was extracted from the start of the stop closure until the end of the periodic cycles of the vowel /a/. This syllable is an example of a CV syllable that a participant heard during the experiment. The second target syllable in this word is /ṭa/. The same extraction procedure was followed for the second syllable, except there were no chances of any pauses in the second syllable, therefore, the start of the stop closure was where the previous vowel's periodic cycles ended, and a visible stop closure appeared. The dental place of articulation in the onset of the second syllable of figure 1 is shown with an asterisk for typographical purposes on PRAAT. The asterisk is used for typographical simplicity to represent the dental place of articulation instead of the IPA dental diacritic symbol for the rest of this study.

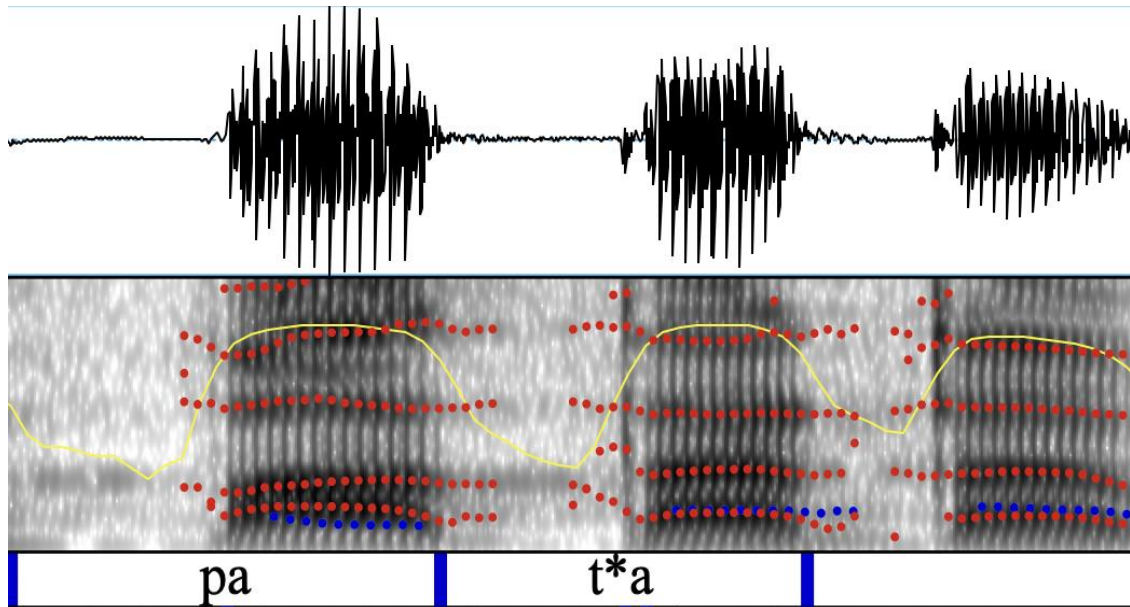


Figure 1. An example of the stimuli extraction procedure for the word 'paṭata on PRAAT

3.3. EXPERIMENT DESIGN AND PROCEDURE. A total of sixteen voiceless unaspirated, sixteen voiceless aspirated, sixteen voiced unaspirated and sixteen voiced aspirated stops were extracted from first and second syllables. That is, eight instances of CV syllables with each stop type as the C were extracted from syllable one. The same procedure was followed for syllable 2, yielding 64 target syllables in total, shown in Table 3. There were eighteen filler items which had the same CV structure as the target syllables. The C of the fillers were other kinds of consonants available in the language and included nasals, j, r and l. The vowels were the same /a/.

An example of table formatting is illustrated by Table 1. It uses the same formatting as Figure 1 above, but with Table # as the caption label. Use only simple horizontal lines as separators, if desired.

	T	TH	D	DH	Total
Syll 1	8	8	8	8	32
Syll 2	8	8	8	8	32
					64

Table 3. Table showing total number of syllables included in the study

Even though the stimuli were extracted from recordings of 9 speakers, for the perception experiment, each participant was presented with randomly selected stimuli from 3 speakers. This was done to ensure that the experiment is not too long for each participant. Each participant of the perception study responded to $64 \times 3 = 192$ sound trials in the experiment.

3.3. DATA COLLECTION PROCEDURE. Data was collected through a Multiple Forced Choice Task Experiment. The experiment was built using the PRAAT MFC Perception Experiment Program. In this task the participants heard the sound tokens and selected the sound they heard out of 6 possible choices. Each participant responded to 192 target trials.

An innovation in the data collection procedure was implemented in order to collect the perception data from participants. This was necessary because the participants were residing in India at the time of their participation in the experiment. While in-person data collection would be ideal, however, due to the ongoing Covid-19 pandemic and travel restrictions, travel to India was not possible, hence, certain modifications had to be made to collect data online. In order to conduct the experiment online, the Zoom platform was used. For smooth running of the experiment, participants were instructed to ensure access to stable internet connection throughout experiment session, find a noise-free room, and use earphones/headphones connected to the computer in order to hear the sound clearly.

When the participants logged on to the Zoom platform, the experimenter shared the PRAAT experiment screen with the participant. The participants heard the target sounds twice for each trial and simultaneously saw the 6 choices on the experiment screen and had to tell the experimenter, which syllable they heard. A practice session was conducted before the actual experiment session, to ensure the audio and video systems were working smoothly and also to provide the participants with a chance to understand the procedure thoroughly. An example of what they saw on the screen while simultaneously hearing the sound trial is shown in Table 4. The IPA symbols were not shown to them, they only saw the Bengali characters.

পা /pa/	মা /ma/	বা /ba/	লা /la/	ফা /p ^h a/	ভা /b ^h a/
------------	------------	------------	------------	--------------------------	--------------------------

Table 4. Selection choices for the participants of the perception experiment

One of the main reasons for conducting the data collection procedure through Zoom was because it mimicked a lab situation where the experimenter could monitor the whole experiment session and track any technical issues or troubleshoot any problems during the experiment. Additionally, not all of the participants were equipped to run the PRAAT experiment on their own, hence sending the program to them was not an option. Moreover, using this procedure, it was easier to ensure that all the participants heard and responded to the stimuli in a consistent manner.

4. Results. Data was analyzed by calculating the sum of correct responses in each category and then calculating the percentage of perceptual accuracy. Results from 50 native speakers of Bengali show a higher perceptual accuracy for syllables containing the voiced stops compared to the voiceless stops. The perceptual accuracy of the voiced stops was 80% and the perceptual accuracy of the voiceless stops was 76% (shown in figure 2). This 4% difference in perceptual accuracy was statistically significant as shown by a linear mixed effects model in R Studio, showing a t value= 4.022 and p value <0.05. This result confirms the first hypothesis of this study.

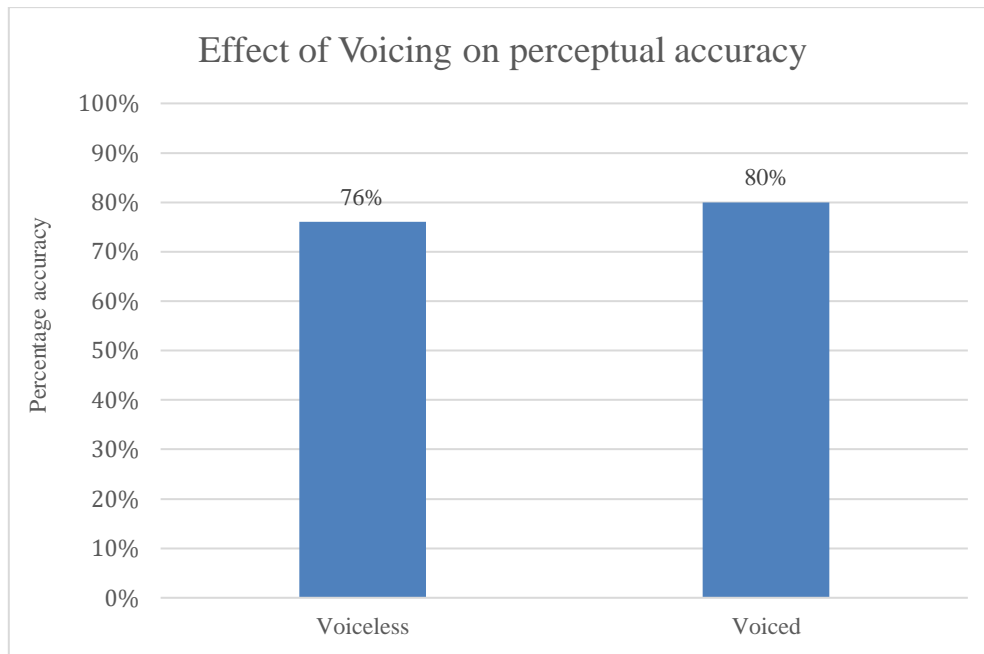


Figure 2. Higher perceptual accuracy of voiced stops compared to the voiceless stops

Next, the effect of aspiration was tested. Results from fifty native speakers of Bengali show a higher perceptual accuracy for syllables containing the aspirated stops compared to the unaspirated stops. The perceptual accuracy of the syllables containing aspirated stops was 82% and the perceptual accuracy of the syllables containing unaspirated stops was 74.27% (shown in figure 3). This 7.73% difference in perceptual accuracy was statistically significant as shown by a linear mixed effects model test in R Studio, showing a t value= 7.625, and p value <0.05. This result confirms the second hypothesis of this study.

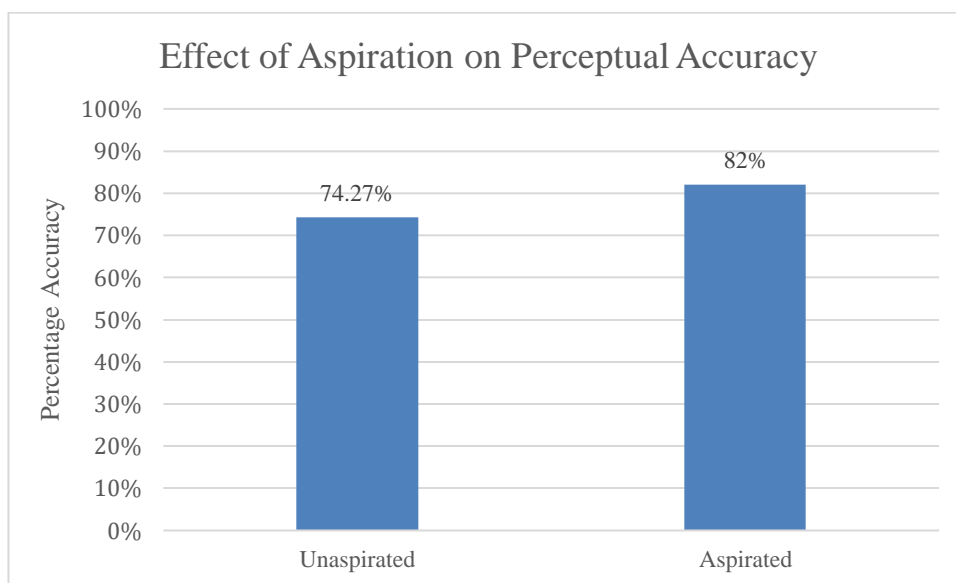


Figure 3. Higher perceptual accuracy of aspirated stops compared to the unaspirated stops

Finally, for the cumulative effect of two marks, the results showed that the DH category had the highest perceptual accuracy percentage compared to all the stop categories, shown in figure 4. As predicted, the T stop had the least perceptual accuracy percentage compared to all the stop categories and the difference between the T category and the DH category was the highest with 11.9% difference. A binomial logistic regression analysis shows that the 11.9% difference between T and DH is statistically a highly significant difference with a t value of 7.433 and p value < 0.001. This result confirms the third and final hypothesis of this study.

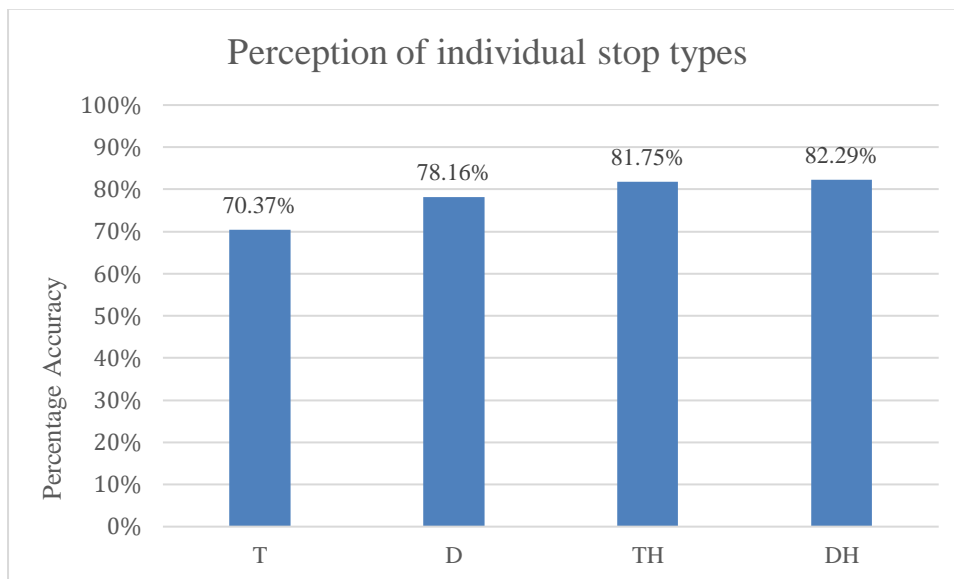


Figure 4. Perception of individual stop categories. Perceptual accuracy of each individual stop type increases with the addition of each additional mark; with the highest perceptual accuracy seen in the DH stop category with two marks, voicing and aspiration.

5. Discussion. First both the independent properties, voicing and aspiration, were tested which showed a significant effect of both of the marked properties on perceptual accuracy. The cumulative effect of both the marks were also tested and results showed a significant effect of voicing and aspiration combined with the highest perceptual accuracy of the DH stops and the lowest perceptual accuracy of the T stops. This confirmed the third hypothesis. However, in order to check the actual content of the patterns, it was informative to check the actual accuracy data. Hence, the percentage accuracies were individually checked which showed a higher improvement in perception due to aspiration compared to the improvement due to voicing. This is also evident from Figure 4, even though there is an improvement in perception due to voicing, the effect due to aspiration is substantially more, indicating that the effect of two different kinds of markedness properties on perception, is not the same. The improved effect may be due to two reasons. First, it could be because of the longer duration associated with aspiration which help the listeners identify the sounds with more ease. The second reason may be due to the noise associated with aspirated stops. Acoustically aspiration involved noise which may also help the listeners identify them more easily. It may also be due to a combination of both the factors. These possibilities need to be systematically tested with experimental data in future research. Additionally, Kingston (1993) suggested a hierarchical relationship between markedness properties in large inventory languages. The suggestion states that the difference between languages with larger phonemic inventories and smaller phonemic inventories lies in the fact that

larger inventory languages have more marked sounds in addition to the unmarked sounds, which are also present in the smaller inventory languages (Lindblom and Maddieson 1988). The languages with larger phonemic inventories, add more properties on the basic contrasts. Kingston (1993) suggests that marked and unmarked sounds in language inventories with larger sizes would exhibit the effects of markedness incrementally from the basic unmarked structure. In this view, Kingston (1993) also suggests that in languages with even larger inventories, another step of incremental effect of the yet another “mark” should appear. The sounds that appear in these languages, have even more marked sounds than the intermediate inventory languages. Therefore, if English is a language with voiced and voiceless contrast compared to a language that has only voiceless, then languages such as Burmese are even larger with three-way contrast in their stop system which gives them voiceless, voiced and the aspirated category stops. Bengali is an even larger inventory language with the voiced aspirated stops. Therefore, in this case, the aspiration property is more marked than the voicing property and voiced aspirated is even more marked than the aspirated property. This proposal by Kingston (1993) was not evaluated previously with experimental data, from either production or perception perspective. The current study shows that the proposal holds true, at least for the perception of the Bengali stops as evidenced from the results.

6. Conclusions. The novel finding in this study is that the systematic study of Bengali stops with marked and unmarked properties show that the marked stops are perceived better than the unmarked stops. An additional finding is that, while both the properties improve the perceptual accuracy of the stops significantly, improvement is more in the case of the aspiration property compared to the voicing property. This indicates that different marks have different effects on perception of segments, the stops in this case and there is a hierarchical relationship between the different types of marks. While this was proposed in previous literature, it was not studied with experimental data. This study clearly shows the differences in the effects of two different types of markedness properties. Aspiration, which is less common compared to voicing, has a stronger effect on the perception of the aspirated stops compared to the voicing property. Overall, it was revealed that corresponding to the distributional patterns predicted by the markedness theory that is supported by previous production studies, markedness has an effect on speech perception. The effect is that of making segments perceptually clearer. This is also true for the cumulative effect of multiple markedness properties, as evidenced by the Bengali voiced aspirated stops.

References

- Carlisle, Robert S. 1988. The effect of markedness on epenthesis in Spanish/ English interlanguage phonology. *Issues and Development in English and Applied Linguistics (IDEAL)* 3. 15–23.
- Ferguson, Charles A. & Munier Chowdhury. 1960. The phonemes of Bengali. *Language* 36(1). 22–59. <https://doi.org/10.2307/410622>.
- Hayes, Bruce & Aditi Lahiri. 1991. Bengali intonational phonology. *Nat Lang Linguist Theory* 9. 47–96. <https://doi.org/10.1007/BF00133326>.
- Jakobson, Roman. 1941. *Child language, aphasia and phonological universals*. The Hague: Mouton.
- Khan, Sameer U. 2010. Illustrations of the IPA Bengali (Standard Bangladeshi). *Journal of the International Phonetic Association* 40(2). 221–225.
- Kingston, John. 1993. Size, structure and markedness in phonological inventories. *University of Massachusetts Occasional papers in Linguistics* 16. 4.
- Trubetzkoy, Nikolai. 1939. *Principles of phonology*. Paris: Klickensieck.