

Perception of Lexical Tones in Sylheti

Amalesh Gope¹, Shakuntala Mahanta¹

¹Phonetics and Phonology lab, Department of HSS, IITG, India

g.amalesh@iitg.ernet.in, smahanta@iitg.ernet.in

Abstract

Tone as a phonological category is capable of distinguishing the meanings of two lexical words by varying the pitch height and/or contour differences. In a tone language, it is indispensable for a hearer to perceive the f_0 fluctuations (and the phonetic and phonological properties) associated with distinct speech signals that distinguish the lexical meanings (Yip 2004, Francis et. al 2003). To qualify as distinct tones, speech signals of (contrasting) words (with distinct tonal property, viz., high and low) must contain large enough f_0 fluctuations between them so that the hearer can use the pitch difference as a cue to distinguish the lexical meanings (viz., high and low). This paper investigates the process of tone perception in Sylheti. A series of historical developments such as the loss of breathy voice contrasts, spirantization and deaffrication predominantly led to a high tone in Sylheti ([dàx] ‘drum’ [d^h>d], [dàx] ‘roar/call’) (Gope & Mahanta, 2015, 2014). The results indicate that the Sylheti speakers by and large rely on f_0 fluctuation in identifying the contrastive tones. Further, the trend of perception of contrastive tones also suggest that tones in Sylheti is perceived (mostly) in categorical manner.

Index Terms: tone, perception, cues, Sylheti, pitch

1. Introduction

Tone as a phonological category is capable of distinguishing the meanings of two lexical words (or utterances) by varying the pitch height and/or contour differences. It is argued that the nature of tonal system of a given language must ‘converge with fundamental processes associated with the production and perception of tones’ (Gandour 1979). In a tone language, therefore, it is equally important for a hearer to perceive the f_0 fluctuations (and the phonetic and phonological properties) associated with distinct speech signals that distinguish the lexical meanings (Yip 2004, Francis et. al 2003). Keeping that in mind, this paper investigates the way tone is perceived in Sylheti by the native speaker.

A series of historical developments such as the loss of breathy voice contrasts, spirantization and deaffrication predominantly led to a high tone in Sylheti ([dàx] ‘drum’ [d^h>d], [dàx] ‘roar/call’, [zá] ‘hot’ [d₃^h>z], [zà] ‘net’ [d₃>z]) (Gope & Mahanta, 2015, 2014).

1.1. Perception of tone

Generally it is assumed that listeners employ and exploit phonetic and/or phonological features associated with a speech signal (at least in part) in order to perceive and distinguish that signal from another one. According to Studdert-Kennedy (1975), “perception entails the analysis of the acoustic syllable, by means of its acoustic features, into the abstract perceptual structure of features and phonemes that characterize the morpheme”.

Researchers however, often do vary in their interpretation of the way tone is perceived by the native speakers of a particular language. While it is widely accepted that the consonants (especially the stop consonants) are mostly perceived in categorical manner (Liberman et. al 1961, Liberman et. al. 1957), the perception of vowel contrasts does not seem to follow the same trend (Abramson 1976, Stevens et. al 1969). The concept of ‘categorical perception’ was developed by Liberman et.al (1957). The categorical perception of speech supports the idea that the perception of two tokens with comparable acoustic variations depends on whether these tokens are heard as the members of same or different categories. Thus, categorical perception is dependent on listeners’ capability of identifying or discriminating a particular token belonging to a specific category of speech sounds. Generally it is assumed that listeners (with the experience of a given language) are more proficient of locating specific categorical boundaries in order to discriminate a particular token than identifying the similarities between two speech sounds with corresponding acoustic differences (Reetz and Jongman 2009). Wang (1976), while studying the perception of Mandarin Chinese observed that native speakers differentiate high rising tone (Tone 2) from a high level tone (Tone 1) in a categorical manner. Abramson (1979), on the other hand, reported that Thai level tones are not perceived categorically. Interestingly, Francis et. al (2003) observed that contour tones in Cantonese are perceived in categorical manner, however, the level tones in Cantonese do not follow the same pattern. This paper examines the properties of perceptual cues of Sylheti tones and investigates whether tones in Sylheti are perceived in categorical manner or not. In the following section we will discuss the details of the experimental design, the methods adopted for preparing the synthesized stimuli, and the details of procedure employed for the current experiment.

2. Experimental design

2.1. Stimulus

To prepare the synthesized stimuli, we first recorded 8 monosyllabic pairs (chosen from the dataset used in Gope & Mahanta 2014) from two native speakers (1 male and 1 female) of Sylheti (Table 1).

Table 1. The dataset displaying the list of monosyllabic words considered for the perceptual experiment.

Sylheti words	Gloss	Sylheti words with a history of underlying aspiration	Gloss
[ga]	‘body’	[ga]	‘wound’
[ban]	‘tie’	[ban]	‘pretend’
[ʃan]	‘donate’	[ʃan]	‘paddy’
[zal]	‘net’	[zal]	‘chilly hot’

All the target words were embedded in a fixed sentence frame of “I am saying X” [ami X xɔiar], X being the target word. The duration, intensity and pitch values of the selected words were manually crosschecked in Praat. We observed distinct f_0 pattern of the contrastive pairs which confirms the presence of lexical tone in Sylheti (Figure 1). The duration and intensity values were found to be identical for the contrastive pairs. Presence of similar phenomena was also reported in Gope & Mahanta (2014).

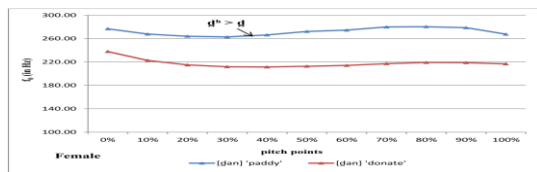


Figure 1: Non-normalized averaged pitch track for [dan] ($n=3$) is drawn using the percentage wise pitch values produced by the female speaker

To generate artificial stimuli, we first marked the pitch onset and pitch offset of the target words and measured the duration values of the contrastive pairs. The pitch contour of each target word was then divided into 5 parts (at every 20% of the TBU/rhyme) by using the following formula:

[Total duration (of the TBU) * X/ 100], where X is the point of calculation (such as 20%, 40%, ..). This was done manually and each durational point was marked separately. To prepare the synthetic stimulus, pitch contour of each TBU was shifted up or down to match a particular point (say the first 20% or the first 40% and so on) of the pitch contour of their contrastive counterparts. So for example, stimulus 1 was created by manipulating and exchanging the first 20% of a high tone contour with the first 20% (starting from the onset till the first 20% of the rhyme) of its low tone counterparts; the remaining portion of the target contour (from 21% till the offset [100%] of the low tone contour) remained unchanged. Thus, 5 stimulus (stimulus 1- starting from the onset till first 20% of the low tone contour, stimulus 2- starting from the onset till 40% of the low tone contour, stimulus 3- starting from the onset till 60% of the low tone contour, stimulus 4- starting from the onset till 80% of the of the low tone contour, and stimulus 5- the entire low tone contour (starting from the onset till the offset), of each target words were resynthesized by exchanging the exact points (for example, 20%, 40%, 60%, 80% and entire high tone contour- 100%) of the contrastive high tone contour (Figure 2).

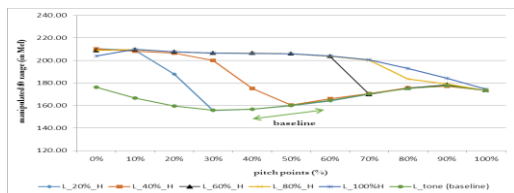


Figure 2: Synthesized contours of first five stimulus set. The original baseline or low tone contour is also shown. The contours are averaged across all the tokens (3 tokens for each word) of all the low tone words (4 words) considered in this experiment

The mean f_0 of the original baseline/low tone was found to be 195.56 Hz (169.72 Mel). The artificial stimulus was thus gradually increased (at every 20%) to match the high tone counterparts. For example, mean f_0 of stimulus 1 was 12.70 Hz

(8.88 Mel) higher than the original low tone contour. The details of f_0 differences between the synthesized stimuli and the original baseline (contours of the original low tone or high tone) are discussed in section 3 (Results and Discussion).

Following the same procedure 5 other stimuli were also created by manipulating the contours of high tone words into a systematic low tone TBU (Figure 3). All the manipulations were done manually by using the Praat pitch manipulation function and the manipulated (target) stimuli were resynthesized to create the sound files used in this experiment.

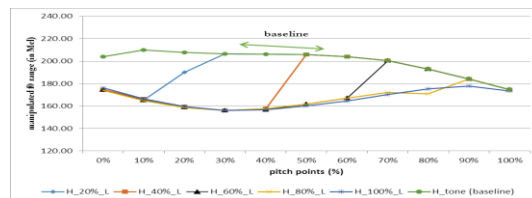


Figure 3: Synthesized contours of final five stimulus set. The original baseline or high tone contour is also shown. The contours are averaged across all the tokens (3 tokens for each word) of all the high tone words (4 words) considered in this experiment

Note that only the contours of the target words (Table 5-1) were manipulated and resynthesized, the remaining portion of the entire (embedded) sentence [ami X xɔiar], including the portion of the onset consonant (if any) of the target words were kept unchanged.

2.2. Subjects

Twelve native Sylheti speakers, (6 male and 6 female) participated in this perceptual experiment. In this experiment we deliberately wanted to involve both the younger generation speakers as well as the older generation speakers to have a better understanding about the status of lexical tone in Sylheti. The first six subjects, aged between 40 and 63 years (3 male and 3 female), represent the middle aged to old generation, while the remaining 6 subjects aged between 21 and 29 years (3 male and 3 female) represent the younger generation.

2.3. Experimental procedure

The target stimulus embedded in the natural sentence carrier frame [ami X xɔiar], ‘I am saying X’ (X being the target word), were played on a notebook connected with a headphone. Each stimulus was embedded with three options - the real meaning, the contrastive meaning, and the third option being “NOT SURE”. Each participant heard 3 repetitions of each stimulus (which were randomized) for a total of 120 tokens (8 words * 3 repetitions * 5 manipulation [each rhyme of those 8 words was manipulated at every 20% of the total duration]). Altogether 1440 responses from 10 stimuli were examined.

All the subjects were instructed to listen and respond to each item and choose one of the three options displayed on the computer screen that best represents the meaning of the target word/stimulus that they have heard. Subjects were allowed to listen to each stimulus up to three times (if required).

In our experiment, instead of following the model of categorical perception directly we have modified our experiment and left the option with the native speakers (listeners) whether to identify or discriminate a particular token belonging to a specific tone category. We argue that, in a given context (in our case it is the exploitation of the f_0 of the rhyme) if a native speaker identifies a particular word ‘X’ as ‘A’ (say for example a word with a higher f_0) and not as ‘B’

or ‘C’, then, ‘A’ is identified as belonging to one particular category (here as high tone). On the other hand, when that particular context of the same word ‘X’ is changed or altered (lowering the f_0 of the rhyme to a certain extent) and the resultant word is perceived as ‘B’ over ‘A’ or ‘C’, then option ‘B’ is being discriminated as belonging to the other category (‘A’- high tone). This process also provides the opportunity to examine whether these stimuli are perceived in a categorically or not. We hypothesized that stimulus with less discrimination might be perceived as the token belonging to the same category, i.e., a token (say a low tone contour) with a manipulation of only first 20% of its contour (into first 20% of contrastive high tone) might not be perceived as a different word (since the manipulated contour will have a very similar shape of that of a low tone), similarly, a token with large discrimination (for example a manipulation of 80% of its total contour) would be perceived as a different word. Thus, the more the fluctuation is, the more it is likely to be perceived as contrastive counterpart.

3. Results and discussion

3.1. Manipulation of low tone contour into high tone contour (at every 20% of the contour)

Stimulus 1: To create stimulus 1, the first 20% of a low tone contour is resynthesized and substituted with the first 20% of contrastive high tone counterpart, the remaining portion of the low tone contour (from the 21% till the offset of the contour) remained unchanged. On average the mean f_0 value of the first 20% (manipulated portion) of the synthesized stimulus (245.66 Hz or 202.39 Mel, first 20%), is 49.3 Hz or 35.02 Mel higher than the first 20% of the original baseline or low tone contour. Responses to stimulus 1 gathered from all the subjects (across all age groups) show that the manipulation of first 20% of the total contour of a low tone was still perceived as a low tone word. On average, 88.89% of the total responses identified stimulus 1 as a low tone word, whereas only 9.72% of the total responses from all the speakers identified it as a high tone word. 1.39% of the total response was marked as ‘Not Sure’.

Stimulus 2: Stimulus 2 represents the second manipulated f_0 contour- the mean f_0 value of the first 40% (manipulated portion) of the synthesized stimulus (243.03 Hz or 200.01 Mel) is 51.80 Hz or 37.07 Mel higher than the first 40% of the original baseline or low tone contour (190.23 Hz or 162.94 Mel). On the other hand, when the entire contour of the synthesized stimulus and the original baseline or low tone contour are compared, it is observed that on average the synthesized stimulus (stimulus 2) is only 23.34 Hz or 16.99 Mel (mean f_0) higher than the original baseline or the low tone TBU.

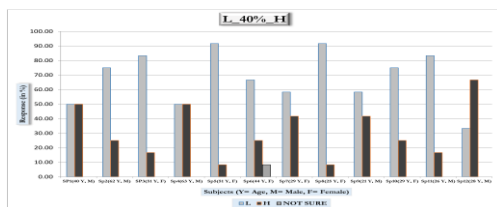


Figure 4: Responses (in % shown in Y axis) of the speakers (shown in X axis) to Stimulus 1 (SP= Speaker/Subject, Y= Age, M=Male, F=Female). Total Responses considered = 144 (4 words * 3 repetitions * 12 subjects, L= (identified as) low tone word, H= (identified as) high tone words

Responses to stimulus 2 (as shown in Figure 4) gathered from all the subjects (across all age groups) show that the manipulation of first 40% of the total contour of a low tone into the same range (first 40%) of their respective high tone counterparts resulted in mixed responses. While a few subjects (subject no 5, 8, 3, and 11) perceived these stimulus mostly as low tone words, a few others (subject no 1, 4, 7, and 9) gave mixed responses. Subject 12 surprisingly identified these tokens mostly as high tone words (nearly 70% of his total response identified stimulus 2 as high tone word). On average, 68.06% of the total responses identified Stimulus 2 as low tone word, whereas 31.25% of the total responses from all the speakers identified it as high tone counterpart. 0.69% of the total response was marked as ‘Not Sure’.

Stimulus 3: Stimulus 3 is the third manipulated contour where the first 60% of a low tone contour is substituted with the first 60% of contrastive high tone contour. Responses to stimulus 3 gathered from all the subjects (across all age groups) indicate that manipulation of first 60% of a low tone into the same range (first 60%) of their respective high tone counterparts created large enough f_0 fluctuations for the native speakers to perceive this stimulus as high tone. While most of the subjects (subject no 1, 4, 5, 7, 8, 10, and 11) have identified these tokens mostly as high tone words, a few subjects perceived these tokens mostly as low tone word (speaker no 2 and 3) [more than 50% and 70% respectively]. Subject 6, 9 and 12 on the other hand gave mixed responses.

Stimulus 4: The first 80% of stimulus 4 bears the f_0 contour of high tone contour while the final 20% is that of the (original) low tone counterpart. The synthesized stimulus is 46.23 Hz (33.2 Mel) [mean f_0 of the entire contour] higher than the original low tone baseline/contour. We hypothesized that since the contours of these low tone tokens are transformed into (almost) identical high tone counterparts, these tokens would be perceived mostly as the high tone tokens.

As expected, the manipulation of first 80% of a low tone contour into the same range (first 80%) of their respective high tone counterparts have been perceived mostly as high tone tokens. On average, only 13.19% of the total responses perceived stimulus 7 as low tone word, whereas as much as 86.11% of the total responses from all the speakers identified it as high tone token. 1.39% of the total responses were marked as ‘Not Sure’ (Figure 5).

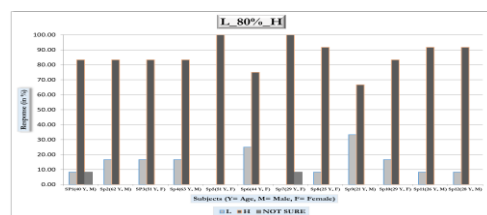


Figure 5: Responses (in % shown in Y axis) of the speakers (shown in X axis) to Stimulus 4 (SP= Speaker/Subject, Y= Age, M=Male, F=Female). Total Responses considered = 144 (4 words * 3 repetitions * 12 subjects, L= (identified as) low tone word, H= (identified as) high tone words

Stimulus 5: The entire contour (100%) of a low tone is exchanged and resynthesized with the contour of the contrastive high tone. As per our hypothesis, most of these tokens should be perceived as high tone words. On average, the synthesized stimulus is 47.86 Hz (34.42 Mel) [mean f_0] higher than the original baseline or low tone contour. The overall responses indicate that the manipulation of the entire

low tone contour into high tone counterpart (the entire contour, from the onset [0%] till the offset [100%] of the TBU) has been perceived mostly as high tone words. On average, only 10.42% of the total responses identified stimulus 5 as low tone word, whereas 89.58% of the total responses were in favor of high tone word.

3.2. Manipulation of high tone contour into low tone contour (at every 20% of the contour)

Stimulus 6: The first 20% of a high tone contour is resynthesized by lowering and exchanging that portion with the first 20% of a contrastive low tone contour. The synthesized portion (first 20%) of stimulus 6 is 37.40 Hz or 30.64 Mel (mean f_0) lower than the original baseline or high tone contour. However, the overall synthesized contour (starting from the onset [0%] till the offset [100%] of the TBU) is 11.17 Hz (7.84 Mel) [mean f_0] lower than original baseline or high tone contour. We assumed that the f_0 fluctuation between the synthesized stimulus (stimulus 6) and original baseline of high tone contour might not be large enough to be perceived as different words. On average, 90.97% of the total responses perceived stimulus 6 as a high tone word, whereas only 8.33% of the total responses from all the subjects identified it as a low tone word. A total of 0.69% of the responses was marked as ‘Not Sure’.

Stimulus 7: The first 40% of a high tone contour is exchanged with the first 40% of contrastive low tone contours. The remaining portion (from 41% till the offset (100%) of the TBU) of the high tone contour was unchanged. On average, the synthesized portion (first 40%) of stimulus 7 is 52.11 Hz (43.98 Mel) [mean f_0] lower than the original baseline or the high tone contour; however, the entire contour (from 0% till 100% of the TBU) is only 25.54 Hz (18.04 Mel) [mean f_0] lower than the original high tone contour. Responses to stimulus 7 (Figure 6) gathered from all the subjects (across all age groups) indicate that most of the listeners have heard stimulus 7 mostly as high tone word. Subject 9 appears to be a bit confused and gave mixed responses (half of his responses were marked as low tone word). Overall 78.47% of the total responses perceived stimulus 7 as a high tone word, whereas 20.83% of the total responses from all the speakers perceived it as a low tone word. 0.69% of the total response was marked as ‘Not Sure’.

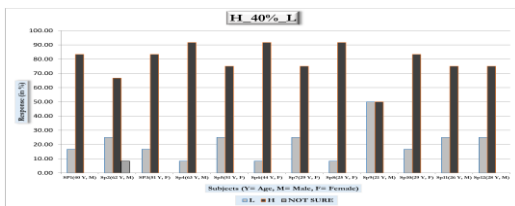


Figure 6: Responses (in % shown in Y axis) of the speakers (shown in X axis) to Stimulus 4 (SP= Speaker/Subject, Y= Age, M=Male, F=Female). Total Responses considered = 144 (4 words * 3 repetitions * 12 subjects, L= (identified as) low tone word, H= (identified as) high tone words)

Stimulus 8: The first 60% of a high tone contour is lowered and resynthesized to match the first 60% of a low tone counterpart; the remaining portion (from the 61% till 100%) of the high tone contour remained unchanged. The synthesized portion (first 60%) of stimulus 8 appears to be 54.36 Hz or 43.46 Mel [mean f_0] lower than the original baseline or high tone contour whereas the entire synthesized contour (from 0% till 100%) is 37.32 Hz (26.54 Mel) [mean f_0] lower than the

original baseline. The overall results indicate that the f_0 fluctuation generated in the synthesized contour of stimulus 8 was large enough for most of the listeners (subject no 5, 7, 8, 9, and 11) to hear these tokens mostly as low tone words. On the other hand, a few listeners heard these tokens mostly as high tone word (subject 2, 6 and 10) [more than 50%]. Subject 1, 4, and 12 gave mixed responses (50% each). On average, 58.33% of the total responses heard stimulus 8 as a low tone word, whereas 41.67% of the total responses from all the subjects heard stimulus 8 as a high tone word.

Stimulus 9: The synthesized portion (first 80%) of stimulus 9 is 52.43 Hz (39.36 Mel) [mean f_0] lower than the original (high tone) baseline. However, the entire contour (from 0% till 100% of the TBU) of stimulus 9 is 46.11 Hz (33.02 Mel) [mean f_0 of the entire contour] lower than the original high tone baseline, thus making the synthesized contour a low tone contour. Hence, we can assume that most of these tokens would be heard as representing the low tone word.

As expected the overall results indicate that all the subjects heard stimulus 9 mostly as low tone words (Figure 7). On average, 92.36% of the total responses perceived stimulus 9 as low tone words, whereas only 6.94% of the total responses from all the speakers perceived it as high tone words. 0.69% of the total responses were marked as ‘Not Sure’.

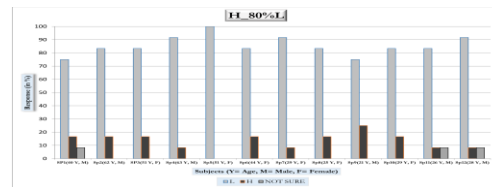


Figure 7: Responses (in % shown in Y axis) of the speakers (shown in X axis) to Stimulus 4 (SP= Speaker/Subject, Y= Age, M=Male, F=Female). Total Responses considered = 144 (4 words * 3 repetitions * 12 subjects, L= (identified as) low tone word, H= (identified as) high tone words)

Stimulus 10: The synthesized contour of stimulus 10 is 47.86 Hz (34.42 Mel) [mean f_0 of the entire contour] lower than the contrastive high tone counterpart. We assume that the tokens of stimulus 10 would be heard mostly as low tone words. As expected, all of the responses to stimulus 10 gathered from all the subjects (across all age groups) suggest that these tokens were heard (mostly) as low tone words. On average, 92.36% of the total responses perceived stimulus 10 as a low tone word, whereas only 6.94% of the total responses of all the subjects perceived it as a high tone word. 0.69% of the total response was marked as ‘Not Sure’.

4. Conclusion

From the results discussed above it is reasonably clear that the f_0 fluctuation of the contrastive words bears the most prominent perceptual cue in identifying contrastive tones in Sylheti. Further, the trend of perception of contrastive tones also suggest that tones in Sylheti is perceived (mostly) in categorical manner. The results also suggest that in addition to f_0 fluctuations, listeners of Sylheti also rely on f_0 fluctuation made till a minimum portion of the total TBU- for example, manipulations made till 60% (and above) of the total contour are likely to be perceived as contrastive tone; thus also confirming the importance of the length of the TBU to a certain extent. Further, the contrastive contours (the entire TBU: starting from the onset) must contain a minimum f_0 fluctuation of 37.32 Hz (26.94 Mel) [mean f_0] and the range of

f_0 fluctuation must be continued (at least) till 60% of the total TBU to be able to be perceived as different words.

5. References

- [1] Yip, M., *Tone*. Cambridge: Cambridge University Press, 2004.
- [2] Francis, Alexander L., Valter Ciocca, and Brenda Kei Chit Ng. "On the (non) categorical perception of lexical tones." *Perception & Psychophysics* 65.7: 1029-1044, 2003.
- [3] Gope, Amalesh, and Shakuntala Mahanta. "An Acoustic Analysis of Sylheti Phonemes" in the XVIII *ICPhS*. (pp 10-14). 2015.
- [4] Gope, Amalesh, and Shakuntala Mahanta. "Lexical Tones in Sylheti" in *Fourth International Symposium on Tonal Aspects of Languages*. 2014. (pp 10-14). 2014.
- [5] Gandour, Jack. "Perceptual Dimensions of Tone: Thai." *South-east Asian Linguistic Studies* Vol. 3. Ed. Nguyen Dang Liem. 3.: 277-300, 1979.
- [6] Gandour, Jack. "Perceptual Dimensions of Cantonese Tones: a Multidimensional Scaling Reanalysis of Fok's Tone Confusion Data." *South-east Asian Linguistic Studies* Vol. 4. Ed. Nguyen Dang Liem. 4.: 415-429, 1979.
- [7] Gandour, Jack. "Perceptual Dimensions of Cantonese Tones: a Multidimensional Scaling Reanalysis of Fok's Tone Confusion Data." *South-east Asian Linguistic Studies* Vol. 4. Ed. Nguyen Dang Liem. 4.: 415-429, 1979.
- [8] Liberman, A. M., Harris, K. S., Eimas, P. D., Lisker, L., & Bastian, J. An effect of learning on speech perception: The discrimination of durations of silence with and without phonemic significance. *Language & Speech*, 4, 175-195. 1961.
- [9] Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358-368. 1957.
- [10] Abramson, A. S. Thai tones as a reference system. In W. Gething, G. Harris, & P. Kullavanijaya (Eds.), *Tai linguistics in honor of Fang-Kuei Li* (pp. 1-12). Bangkok: Chulalongkorn University Press. 1976.
- [11] Stevens, K. N., Liberman, A. M., Studdert-Kennedy, M., & Öhman, S. E. G. Crosslanguage study of vowel perception. *Language & Speech*, 12, 1-23, 1969.
- [12] Reetz, Henning, and Allard Jongman. *Phonetics: Transcription, production, acoustics, and perception*. Vol. 34. John Wiley & Sons, 2011.
- [13] Wang, W. S.-Y. The phonological features of tone. *International Journal of American Linguistics*, 33, 93-105, 1967.
- [14] Wang, W. S.-Y. (1976). Language change. In S. R. Harnad, H. D. Steklis, & J. Lancaster (Eds.), *Origins and evolution of language and speech* (Annals of the New York Academy of Sciences, Vol. 280, pp. 61-72). New York: New York Academy of Sciences. 1967.
- [15] Boersma, Paul & Weenink, David., Praat: doing phonetics by computer (Version 5.3.04_win32) [Computer program], 2008., Online: [http://www.praat.org/Bradley 1982.html](http://www.praat.org/Bradley%201982.html), Retrieved on December 22, 2011.
- [16] Esposito, C.M "The Effects of Linguistic Experience on the Perception of Phonation" in *Journal of Phonetics*. Vol. 38(2), pp. 303–316, 2010a.
- [17] Esposito, C.M. "The Perception of Pathologically-disordered Phonation by Gujarati, English and Spanish Listeners" in. *Language and Speech*. Vol. 54(3), pp. 1–16, 2011.
- [18] Bickley, C. Acoustic analysis and perception of breathy vowels. In *Speech communication group working papers*, pp. 71-82. Cambridge: Massachusetts Institute of Technology. 1982.