Research Article

A study on the role of temporal integration of VOT, vowel length and noise in the perception of bilabial stops

JAIISH(2014)
Vol 33, pp. 51-58

Saransh Jain[a], N. P. Nataraja[a] and Sudha P. Nair[a]

Affiliations

[a]JSS Institute of Speech and Hearing

Corresponding Author

Saransh Jain
Department Of Audiology
JSS Institute of Speech and Hearing
saranshavi@gmail.com

Key Words

Temporal Integration
Voicing
Categorical Perception
Vowel Length
VOT

Abstract

The present study aimed at investigating the role of temporal integration of voice onset time (VOT), vowel length and noise in the perception of bilabial stop consonants. For this purpose, a 9 point continuum of voiced to voiceless sound was constructed in a synthetic speech token and each point along the continuum was associated with long and short vowel separately. The stimuli were presented to 30 normal hearing individuals in quiet and noisy situation. Categorical perception was assessed for the identification of perceptual boundary and the reaction time task. The results revealed that vowel length had significant effect on perception of bilabial stop voicing only in quiet but not in noisy situations. Thus, it was concluded that temporal integration of VOT and vowel length are important cues in quiet situation, but in the presence of noise, individuals appear to use some other cues to perceive voicing of bilabial stop consonants.

Introduction

Speech perception is the process of decoding the acoustic dimensions of speech signal, which is thought to depend on the use of language processes (Escudero, 2009). Research on speech perception seeks to understand the way listeners recognize speech sounds and use this information to understand spoken language (Boothroyd, 1998). The process of speech perception relies on various acoustic cues which are either perceived simultaneously (in trading relation; Repp, 1983), or are integrated with respect to time (as in temporal integration).

In studies that have been carried out to assess the process of speech perception, temporal integration refers to how chunks of information arriving at the ears at different times are linked together by the listener in mapping speech sounds into meaningful words (Nguyen & Hawkins, 2003). The process of temporal integration requires online amalgamation of various cues with respect to time. This phenomenon is vital and fundamental to almost all the models of auditory function (Viemeister & Wakefield, 1991). According to the models of auditory perception, the acoustic cues in consonants result in rapid spectral and temporal changes and the processing of these changes depend upon the input within the short temporal window (Poeppel, 2003). This short time window duration is barely of a few millisecond for the processing of the consonantal information (Plomp, 1964; Penner & Cudahy, 1973; Viemeister, 1979). In such a short duration, the acoustic cues are integrated to extract overall meaning. This observable fact is an important part of consonant perception. Thus, it is evident from the review that the temporal nature of the acoustic cues plays important role in consonant perception. Some of the important dynamic cues in consonant perception are preceding and/or following vowel duration, voice onset time, formant transition duration, burst duration and closure duration.

The vowel duration has been found to be an essential cue in the perception of different categories of speech sounds. Research exploring the role of vowel length has revealed that the duration of the vowel following the stop consonant is vital for the discrimination of voiced and voiceless sound (Raphael, 1972). It has been suggested that vowels were of shorter duration when preceded by voiceless consonants and are of longer duration following a voiced consonant (Peterson & Lehiste, 1960). Miller and Liberman (1979) also demonstrated subject's sensitivity to this relationship by assessing the identification of two |b|-|w| manner continua with different vowel lengths and found that category boundary shifted depending on the rate of speech.

Another important cue that is utilized to distinguish between voiced and voiceless stop sound is VOT (Lisker & Abramson, 1964). While voiced

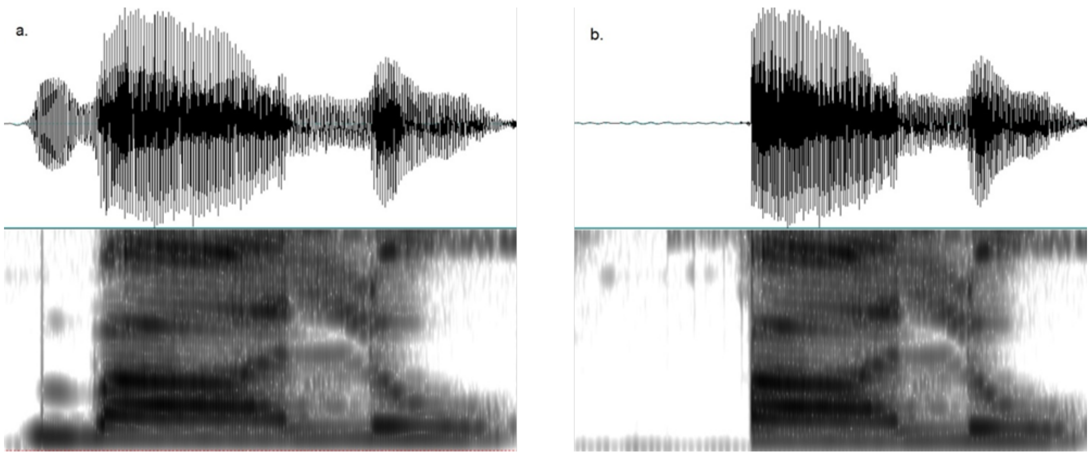

Figure 1: The waveform and the sound spectrogram of the word (a) |balu| and (b) |palu|.

consonants have been found to have leading VOT, voiceless consonants have lagging VOT. McMurray in 2004 studied the temporal integration of VOT with the vowel length in the perception of voicing. The results of the experiment revealed that there was a significant effect of vowel length in the categorical perception of voicing. In the voicing continua, as the VOT was decreased, the perception shifted more towards the voiceless sound, and so with the short vowel length. The results were considered sufficient to justify the role of temporal integration of earlier cues like VOT with the later cues of vowel length. Many other investigators have found results in agreement with the above mentioned studies in various languages like English (Delattre, 1962); German (Fourakis & Iverson, 1984), Japanese (Homma, 1981), Tamil (Balasubramanian, 1981), Malayalam (Velayudhan, 1971), and Kannada (Savithri, 1986). As evident from the literature cited above, it may be inferred that temporal integration of VOT and vowel length plays an important role in the perception of stop consonants.

In McMurray's study, where the author strongly demonstrated that role of temporal integration in the perception of voicing in quiet situation, no such study has been reported in literature to demonstrate the effect of temporal integration in the perception of voicing in the presence of noise (McMurray & Aslin, 2004). Perception of speech has also been found to be influenced by the presence of noise. Alwan (1992) reported that in the presence of noise the perception of stop consonant was affected. Researchers have also reported that the categorical perception of stop consonant was affected in the presence of noise. Lisker and Abramson (1964) reported that the perception of voicing feature of the syllable initial plosive is affected in the presence of noise. Others reported that the alveolar-velar categorical boundary shifted to a higher frequency location in the presence of noise (Syrdal-Lasky, 1978). Under such conditions temporal integration may facilitate correct perception. On the other hand, Ali, Spiegel and Mueller (2002) found that voicing detection for the stop consonants was hardly affected in the presence of Gaussian white noise. Similar results were also obtained by Miller and Nicely (1955) and Wang and Bilger (1973). In an extensive study by Parikh and Loizou (2005), comparison of the perception of |pa|-|ta|-|ka| and |ba|-|da|-|ga| in the presence of speech shaped noise and multi talker babble at various SNR levels of -5, 0, 5 and 10 dB was carried out, and it was found that the stop consonant identification remained high even at -5dB SNR. Thus, there is a wide discrepancy in literature with reference to the perception of voicing of stop consonants in the presence of noise.

In addition, in most of the studies where researchers reported that noise has minimal effect on the perception of voicing feature of stop consonants, the integration of various temporal cues were not considered. Thus, these results can't be generalized to temporal integration experiments. Consequently, taking into account the lack of sufficient research literature examining the effect of the noise on the integration of various temporal cues, the present study was formulated to investigate the role of temporal integration in the perception of voicing in syllable initial bilabial plosive in the presence of noise.

Method

Participants: A total of 30 individuals (15 males and 15 females) in the age range of 18-25 years participated in the present study. All the subjects were native Kannada speakers. The hearing sensitivity was assessed using pure tone audiometry and speech audiometry and only those individuals with normal hearing sensitivity (PTA $\leq$ 25 dB HL; SIS $\geq$ 90%, re: ANSI, 1996) were selected for the present study. The selected participants were also assessed to have no associated speech, language, neurological, or psychological problems.

Stimulus: A voicing continuum was generated by considering a synthesized speech token, with voiced plosive |ba| at the initial position of the word |balu| using the Praat software (version 5.1.41). This speech token was considered as the first point of the voicing continuum. The remaining points along the continuum were constructed by truncating the voice onset time (VOT) systematically from the voiced sound in 5 ms steps till the onset of the stop burst and replacing the truncated VOT segment at each point with aspiration. This yielded a 9 point voicing continuum with voiced plosive |balu| at one endpoint and voiceless plosive |palu| at the other endpoint (Figure 1). The length of the vowel following stop consonant was kept constant as 250 ms (long vowel) in one set of the continuum. In another set, the vowel length was reduced to 100 ms (short vowel) using Praat software. These specific values of the vowel duration were taken up for this study based on the findings of Nataraja (2000) and Raphel (1975). Nataraja (2000) found that the vowel duration following voiced consonant was approximately 220-260 ms in Kannada language. Similarly, Raphel (1975) reported that the difference in the vowel duration between voiced and voiceless consonant to be approximately 150-200 ms for bilabial stops.

After generating both the continuums, each stimulus in the continuum was normalized to 70 dB SPL (average most comfortable level). In order to see the effect of noise, 5 Kannada speaker multi talker speech babble was recorded and presented with the continuum at signal to noise ratios (SNR) of +2, 0 and -2 dB using the Colea toolbox (A Matlab based software tool for speech analysis; Loizou, 2009). The reference for these SNR values was obtained from the study by Alwan (1992).

Procedure: The entire experiment was conducted in a standardized sound treated dual room setup with minimal reverberation as well as minimal visual distraction. Informed written consent was obtained from each participant prior to the commencement of the study and in brief, the purpose of the study was explained to them in their native language. The stimulus was presented using DMDX reaction time software installed in a laptop and delivered via headphones. The output of the headphones was calibrated using B&K 2280 (me-

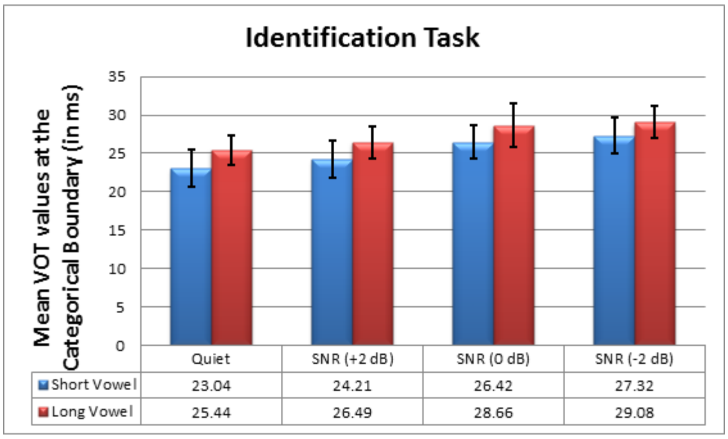

Figure 2: Mean (SD) identification scores of VOT at the categorical boundary as a function of various SNR levels and vowel length.

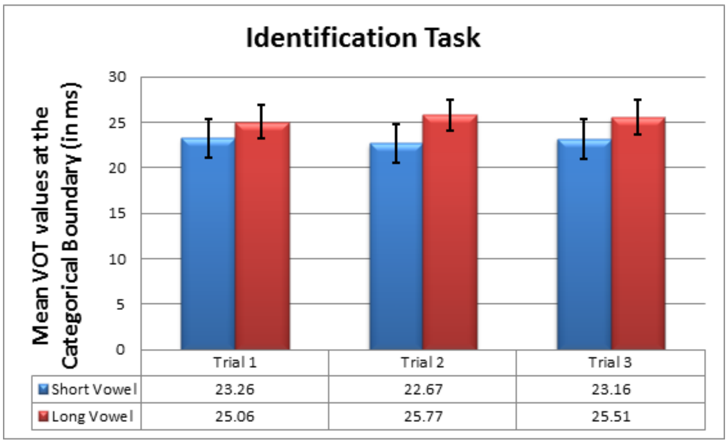

Figure 3: Mean (SD) identification scores of VOT at the categorical boundary as a function of repetition with respect to each vowel length.

diator) sound level meter and the volume of the laptop was set to produce an output of 70 dB SPL on SLM for a 1KHz pure tone routed through a 2cc coupler. Initially, 5-10 stimuli were presented as practice trials in order to familiarize the participants to the stimuli as well as test procedure. To ensure for the reliability, each stimulus was presented thrice. The stimuli were presented in random order within each of the continuum. An inter stimulus interval of 3000 ms (3 seconds) was kept. Thus, a total of 9 stimuli in each continuum were presented in random order thrice wherein one set contained longer vowel duration and the other set a short vowel duration. Further, this complete set of stimuli was presented in quiet as well as at 3 SNR levels. Hence, the stimulus arrangement was as follows: [9 point VOT continuum*2 Vowel Length*4 Levels (3 noise and 1 quiet)] 3 Trials. In totality, 216 stimuli were presented to each participant. The participants were instructed to listen to the stimulus carefully and repeat it as fast as possible. The responses were recorded using a microphone attached to the laptop for the identification and the reaction time task. The complete testing was carried out in two sessions of approximately 45-50 minutes each.

In the identification task, the percentage correct response of the voiced/voiceless bilabial plosive across the continuum was estimated. This was plotted as an averaged response at each VOT step and the intersection of the line graph drawn for |balu| and |palu| was used to derive the categorical boundary. The categorical boundary was measured using logistic regression model and linear or non linear interpolation function. Individual participant's responses for all the trials were averaged across each step along the continuum and the response line was fit with either linear or non linear regression depending on the distribution of the responses using Prism software (version 5.03; GraphPad Software Inc.). Once the line was fit to the curve, 50% probability response was interpolated. The boundary was defined as the point on the graph, where 50% correct responses for either |palu| or |balu| was obtained. In the reaction time task, the minimum time required for responding to each stimulus was noted. The mean reaction time at the categorical boundary was calculated as the average reaction time for the region of ambiguity (Pisoni & Tash,

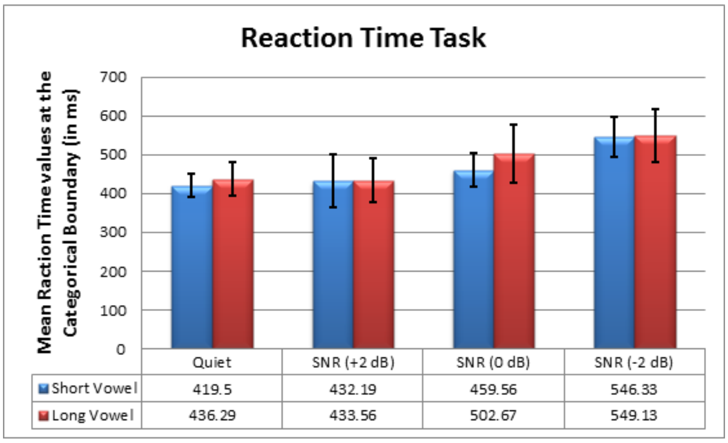

Figure 4: Mean (SD) reaction time scores at the categorical boundary as a function of various SNR levels and vowel length

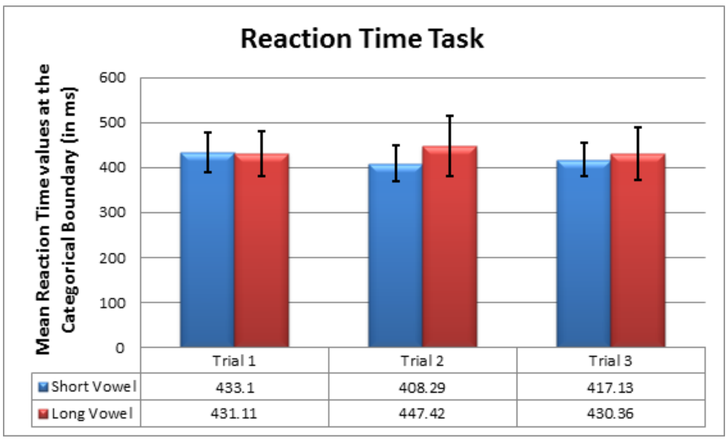

Figure 5: Mean (SD) identification scores of VOT at the categorical boundary as a function of repetition with respect to each vowel length.

1974). Ambiguous region is defined as the points along the continuum where the listener was unsure about the exact categorization of the speech sound. This is the region around the categorical boundary (McMurray & Spivey, 1999).

Results

In summarizing the procedure and analysis of the obtained data, 9 steps voicing continuum of the initial bilabial plosive was presented to each participant. The continuum was presented in quiet as well as at three different SNR levels, viz. +2dB, 0dB and -2dB SNR levels, each time associated with short and long vowels separately, and each stimulus was presented thrice. The results were noted separately for identification task and reaction time task. The identification task was utilized to obtain the categorical boundary. It was determined as the point along the continuum where the perception was changed from either voiced to voiceless or from voiceless to voiced sound. Using the logistic regression model and linear or non linear interpolation of the interaction value, the VOT value at the categorical boundary was obtained.

The mean and the standard deviation of the VOT values at the categorical boundary as the function of SNR and vowel length are illustrated in Figure 2 and as a function of repetitions are represented in Figure 3. The main effect of SNR, vowel length and repetition as well as the interaction of these independent variables on the identification task was measured using 3 factor mixed ANOVA. Table 1 indicate the results of three factor mixed ANOVA for the identification of the perceptual boundary.

As is evident from the Figure 2, the mean identification scores at the categorical boundary were poorer in the presence of noise in comparison to the quiet situation when the stimulus was associated with both long as well as short vowel duration. Similarly, the mean identification scores at the categorical boundary were different for the stimulus association with long vowel in comparison to that with short vowel length. The categorical boundary shifted towards the perception of voiceless sound when the stimuli were associated with short vowel. Additionally, no marked differences in the mean scores were observed as a function of repetition of the stimuli, indicating consistency in the responses across trials (Figure 3).

The results tabulated in Table 1 may be explained in terms of SNR and vowel length individually having a significant effect on the identification of voicing parameter of bilabial plosive at the categorical boundary, however, no such effect was observed as a function of repetition of the stimuli. However, the interaction between SNR and vowel length as well as vowel length and repetition was not statistically significant. The overall interaction of all the three independent variables on the identification task showed significant effect.

The results of the reaction time task was also expressed in terms of the main effect of each independent variable on the reaction time, as well as the interaction effect of these variables on the reaction time at the categorical boundary. Table 2 indicate the mean reaction time for the perception and categorization of bilabial stop at each point along the continuum, tabulated separately for long as well as short vowel duration. The mean reaction time at the categorical boundary as a function of SNR and vowel length is shown in figure 4, and as a function of repetition along the vowel length is illustrated in Figure 5.

From Figure 4, it is possible to make of that the mean reaction time at the categorical boundary was affected as a function of SNR. The reaction time was found to be minimum for quiet situation, and maximum for -2dB SNR level. On the other hand no marked differences in the mean values were observed as a function of vowel length. Similarly, the mean values were also comparable across trials for both long as well as short vowel length as evident in Figure 5.

The 3 factor multiple ANOVA results are tabulated in Table 1. It may be worthy to note than similar to the identification scores, there was a significant main effect on the reaction time as a function of SNR, and no statistically significant effect of repetition on the reaction time was observed. However, contrary to the identification task results, no significant effect of vowel length on reaction time was observed. The interaction between SNR and vowel length revealed significant effect, whereas no such significant effect was observed as a function of interaction between vowel length and repetition, SNR and repetition, as well as combined interaction of SNR, vowel length and repetition for the reaction time task.

Discussion

The VOT for initial bilabial plosives and vowel length were found to be related. This relationship has been reported to be present across different languages. The results of the present study are in agreement with the findings of McMurray and Aslin (2004) who found that with short vowel length, when the VOT was increased, the perceptual boundary shifted towards the perception of the voiceless sound. Similar findings have also been reported by several other investigators in different languages. Delattre, (1962); House, (1961); Umeda, (1975); Whitehead and Jones, (1976); Fox and Terbeek, (1977); Crystal and House, (1988);

Davis and Summers, (1989); and Jong, (1991) found the shift in the categorical boundary with varying vowel length in English. Fourakis and Iverson, (1984) found the similar effect in German, Homma, (1981) in Japanese; O'Shaughnessy (1981) in French; Lyberg (1981) in Swedish; Balasubramanian (1981) in Tamil; Velayudhan (1971) in Malayalam; Shalev, Ladefoged and Bhaskara Rao (1994) in Toda; and Savithri (1986) in Kannada. Contrary to these findings, Mitleb (1984) did not find significant differences in perception of voicing with varying vowel length. The author attributed their findings to the fact that Arabic language does not have a clear voiced-voiceless distinction of plosives. The findings of the present study highlighted the role of temporal integration in voiced-voiceless distinction in quiet situations.

The results of the study also indicated that noise plays an important role in the perception of initial bilabial plosives. Presence of noise at various SNR levels was found to significantly affect voicing perception. Further, the effect was more evident at less redundant SNR levels. Formby, Childers and Lalwani (1996) also reported that mean identification scores for voice-voiceless distinction was poor for the stimuli presented at 0 dB SNR in comparison to +12 dB SNR. In another such study, Fledman, Griffith and Morgan (2009) found that there was a gradual change in the perception of voicing in the presence of noise compared with no noise condition. Syrdal-Lasky (1978) compared the identification of |p|-|t|-|k| and |b|-|d|-|g| at three intensity levels i.e. 75 dB, 92 dB and 92 dB+60 dB of noise and found poor voiced-voiceless discrimination at the 92 dB+60 dB noise condition.

Contrary to the present findings, Ali, Spiegel and Mueller (2002) found that voicing detection was hardly affected by the presence of Gaussian white noise. Similar results were obtained by Miller and Nicely (1955) and Wang and Bilger (1973) also. In an extensive study by Parikh and Loizou (2005), they compared the perception of |pa|-|ta|-|ka| and |ba|-|da|-|ga| in the presence of speech shaped noise and multi talker babble at various SNR levels, of -5, 0, 5 and 10 dB, and they found that stop consonant identification remained high even at -5dB SNR. These findings are difficult to explain and may be attributed to stimuli and the methodological variation amongst the studies. It may also be noteworthy that in none of these findings, vowel length was considered and hence, they were not true measurements of temporal integration, but only of the perception of voicing with variable VOT. So, the findings of these studies can't be generalized to the experiments assessing the role of temporal integration in the perception of voicing.

However, contrary to the individual effect of SNR and vowel length especially on the perception of voicing, the interaction of SNR with vowel length showed no significant change. This indicates that in the presence of noise, vowel length does not play a major role in the perception of voicing parameter of initial bilabial plosive. Under such condition, i.e., in the presence of background noise, there must be some other acoustic cue for the perception of voicing. According to Jiang, Chen and Alwan (2006), at reduced SNRs, listeners may use F1 onset transition as a cue for stop voicing.

Another finding obtained from the present study was increased reaction time at the categorical boundary in comparison to the extremes of the continuum. The obtained results were in agreement with the findings of Pisoni and Tash (1974), where they found increased mean reaction time as the stimuli approaches to the ambiguous region i.e. the region near the category boundary. Similar results were obtained by other researchers (Whalen, 1991; Whalen, Abramson, Lisker & Mody, 1993; Blumstein, Myers & Rissman, 2005).

An additional observation in this experiment was that there was no significant increase in the mean reaction time for the responses with variable

Table 1: Results of 3 factor mixed ANOVA showing difference in the identification of perceptual boundary and the reaction time in the voicing continuum for initial bilabial plosive consonants as a function of SNR, vowel length and repetition of the stimuli

Source	Degree of Freedom	Identification Task		Reaction Time Task	
		F-value	p-value	F-value	p-value
SNR	2	4.171	0.042ˆ	2.973	0.031ˆ
Vowel Length	1	6.350	0.017ˆ	1.035	0.079
Repetition	2	0.212	0.809	0.279	0.147
SNR*Vowel Length	2	1.117	0.069	4.255	0.001ˆ
SNR*Repetition	4	1.135	0.058	1.479	0.059
Vowel Length*Repetition	2	1.983	0.048ˆ	0.685	0.219
SNR*Vowel Length*Repetition	4	3.679	0.032ˆ	1.751	0.057

ˆSignificant at 0.05 level.

Table 2: The reaction time values for each point along the continuum with long and short vowel in quiet and noisy situation

Voicing Continuum	Quiet		SNR = +2		SNR = 0		SNR = -2	
	Long Vowel	Short Vowel	Long Vowel	Short Vowel	Long Vowel	Short Vowel	Long Vowel	Short Vowel
Point 1	232.37	227.30	339.31	336.08	307.67	353.33	418.87	479.07
Point 2	292.43	271.62	292.45	291.69	364.67	336.33	438.93	445.08
Point 3	352.41	350.69	355.03	353.33	361.33	364.33	453.53	441.80
Point 4	262.32	260.53	380.31	373.33	456.67	417.33	516.80	577.60
Point 5	443.72	419.82	444.21	443.07	548.67	501.33	581.47	515.07
Point 6	428.87	419.19	422.91	421.31	347.80	354.53	432.13	425.60
Point 7	219.41	207.81	219.01	218.98	451.67	395.33	435.07	487.07
Point 8	217.48	202.21	223.27	222.08	406.06	399.67	413.07	483.01
Point 9	241.49	224.37	230.61	225.10	333.33	387.27	425.60	487.67
Mean	298.94	287.06	323.01	320.55	397.54	389.94	457.27	482.44

vowel duration in the perception of voiceless sound. Unlike the effect of vowel duration on reaction time, introduction of noise had significant effect when the overall mean reaction time and it was found to be higher than that for the quiet conditions. Further, in reduced SNR level conditions i.e. to 0 and -2 dB, a negligible shift in the perceptual boundary was observed. This further strengthens the previous findings and inference from the identification task, that in the presence of noise, vowel length may not seem as a cue for the perception of voicing, and that individuals might rely on some other cues for detecting voice-voiceless distinction.

In summary, the present research finding in coordination with the available review of literature point towards the possible existence of some other cue in the perception of voicing feature of initial bilabial plosive in the presence of noise and that temporal integration of VOT and vowel length may be sufficient for the perception in quiet situation but not in the presence of noise.

References

Ali, A. M. A., Spiegel, J. V. & Mueller, P. (2002). Automatic detection and classification of stop consonants using an acoustic phonetic feature based system. *Proceedings of XIV International Congress on Phonetic Sciences*, (pp. 1709-1712), Berkeley.

Alwan, A. A. H. (1992). Modeling speech perception in noise. The stop consonants as a case study. *RLE Technical Report No. 569.* Unpublished master's dissertation submitted to Massachusetts Institute of Technology, Cambridge, Massachusetts.

Balasubramanian, T. (1981). Duration of vowels in Tamil. *Journal of Phonetics, 9*, 151-161.

Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience, 17*(9), 1353-1366.

Boothroyd, A. (1998). The perception of speech by children with hearing loss. In A. Weisel (Ed.), *Issues unresolved. New perspectives on language and deaf education* (pp. 103-116). Washington, D.C.: Gallaudet University Press.

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America, 83*(4), 1553-1573.

Davis, S., & Summers, W. V. (1989). Vowel length and closure duration in word-medial VC sequences, *Journal of Phonetics, 17*, 339-353.

Delattre, P. (1962). Some factors of vowel duration and their cross-linguistic validity. *Journal of the Acoustical Society of America, 34,* 1141-1142.

Escudero, P. (2009). The linguistic perception of similar L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 151-190). Berlin, Germany: Walter deGruyter GmbH & Co.

Feldman, N. H.,Â Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review, 116*(4), 752-782.

Formby, C., Childers, D. G., & Lalwani, A. L. (1996). Labeling and discrimination of a synthetic fricative continuum in noise: A study of absolute duration and relative onset time cues. *Journal of Speech & Hearing Research, 39,* 4-18.

Fourakis, M., & Iverson, G. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica, 41*, 140-149.

Fox, R. A., & Terbeek, D. (1977). Dental flaps, vowel duration and rule ordering in American English, *Journal of Phonetics, 5*, 27-34.

Homma, Y. (1981). Durational relationship between Japanese stops and vowels. *Journal of Phonetics 9,* 273-281.

House, A. S. (1961). On vowel duration in English. *Journal of the Acoustical Society of America, 33*, 1174-1178.

Jiang, J., Chen, M., & Alwan, A. (2006). On the perception of voicing in syllable-initial plosives in noise. *Journal of the Acoustical Society of America, 119,* 1092-1105.

Jong, P. F. (1991). *The measurement of attention: The construction of attention tests for childre.* Unpublished Master's Dissertation submitted to Vrije University, Amsterdam

Lisker, L., & Abramson, A. S. (1964). A cross language study of voicing in initial stops: acoustic measurements. *Word, 20,* 384-422.

Loizou, P. (2009). COLEA: a matlab software tool for speech analysis. Retrieved February 18, 2014 from *http://www.utdallas.edu/ loizou/speech/colea.htm.*

Lyberg, B. (1981). *Temporal properties of spoken Swedish.* Monograph from the Institute of Linguistics, vol. 6, University of Stockholm.

McMurray, B. & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy, 6*(2), 203-229

McMurray, B., & Spivay, M. J. (1999). The categorical perception of consonants: the interaction of learning and processing. *Proceedings of the Chicago Linguistics Society, 35,* 205-219.

Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusion among some English consonants. *Journal of the Acoustical Society of America, 27,* 338-352.

Miller, J. L., & Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semi-vowel. *Perception & Psychophysics, 25*, 457-465.

Mitleb, F. (1984). Voicing effect on vowel duration is not an absolute universal. *Journal of Phonetics, 12,* 23-27.

Nataraja, N. P. (2000). Transformation of speech of hearing impaired. Departmental project report submitted to All India Institute of Speech and Hearing.

Nguyen, N., & Hawkins, S., (2003). Temporal integration in the perception of speech. *Journal of Phonetics, 31(3/4)*, 279-287.

O'Shaugnessy, D. (1981). A study of French vowel and consonant duration? *Journal of Phonetics, 9,* 385-406.

Penner, M. J., & Cudahy, E. (1973). Critical masking interval: a temporal analog of the critical band. *Journal of the Acoustical Society of America, 54,* 1530-1534.

Peterson, G.E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of Acoustical Society of America, 32*, 693-703.

Pisoni, D., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics, 15* , 289-290.

Plomp, R. (1964). The rate of decay of auditory sensation. *Journal of the Acoustical Society of America, 36*(2), 277-282.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication, 41,* 245-255.

Raphael, L. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*(3), 1296-1303.

Repp, B. (1983). ?Bidirectional contrast effect in the perception of VC-CV sequences, *Perception and Psychophysics, 33 (2)*, 147-155.

Savithri, S. R. (1986). Duration analysis of Kannada vowels. *Journal of the Acoustical Society of India, 14*(2), 34-40.

Shalev, M., Ladefoged, P., & Bhaskararao, P. (1994). Phonetics of Toda. *PILC Journal of Dravidic Studies, 4,* 19-56.

Syrdal-Lasky, A. (1978). Effects of intensity on the categorical perception of stop consonants and isolated second formant transitions. *Perception & Psychophysics, 23*, 420-432.

Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America 58*(2), 434-445.

Velayudhan, S. (1971). *Vowel duration in Malayalam: An acoustic phonetic study.* Trivandrum: Dravidian Linguistic Association of India.

Viemeister, N. F. (1979). Temporal modulation transfer functions based on modulation thresholds. *Journal of the Acoustical Society of America 66 (5),* 1364-1380.

Viemeister, N. F., & Wakefield, G. H. (1991). Temporal integration and multiple looks. *Journal of the Acoustical Society of America 90 (2),* 858-865.

Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. *Journal of the Acoustical Society of America, 54*(5), 1248-1266.

Parikh, G., & Loizou, P. (2005). The influence of noise on vowel and consonant cues. *Journal of the Acoustical Society of America, 118,* 3874-3888.

Whalen, D. (1991). Perception of the English /s/-/S/ distinction relies on fricative noises and transitions, not on brief spectral slices. *Journal of the Acoustical Society of America, 90,* 1776-1785.

Whalen, D., Abramson, A., Lisker, L., & Mody, M. (1993). F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America, 4*(1), 2152-2159.

Whitehead, R. L., & Jones, K. O. (1976). Influence of consonant environment on duration of vowels produced by normal hearing, hearing impaired and deaf adult speakers. *Journal of Acoustical Society of America, 60*, 513-515.