Effectiveness of Time- Expansion of Consonants for the Benefit of Hearing Impaired Listeners

Dr N H SHOBHA

Department of Electronics & Communication Engineering, Methodist college of Engineering & Technology
Hyderabad, India

ABSTRACT

The current paper investigated the efficacy of consonant duration modification or consonant time- expansion on speech intelligibility for the impaired listeners. This work is an extension of author's previous papers and projects a comprehensive summary of the same along with the correlation of results with the current work. This paper highlights the importance of the lengthening an important acoustic attribute of clear speech - 'consonant-Duration'. A case for synthetic clear speech in the context of hearing impairment was argued for speech intelligibility. Consonant recognition in noise free and noisy situations using the non-sense syllable test (NST) was investigated upon (i) Hearing Impaired subjects, and (ii) Normal hearing subjects under simulated hearing impairment. Stops and Fricative consonants of English language with cardinal vowels were processed for time expansion. Non-uniform algorithm for waveform expansion involving selection and expansion of acoustic features was adopted. The Burst Duration, Voice Onset Time, Formant Transition Duration for Stops; and Fricative Noise Duration and Formant Transition Duration for Fricatives, were time-expanded by 50 - 100% of original duration. The speech perception in noise tests were quantified in terms of relative information transmission measures. The results reported that among the various segments selected for modifications, only Burst Duration of Stop-vowels has positive benefit on speech intelligibility improvement for the hearing impaired subject while the rest of them lead to perceptual errors.

Keywords- Burst Duration, Hearing impaired, Clear speech, Consonant Duration, Speech Intelligibility

1. INTRODUCTION

Good hearing is essential not only for communication and business, but also for many of the subtle joys that make life worth living. Persons with hearing impairment are less fortunate, as they often they have the greatest difficulty understanding speech in noisy environments. Sensorineural hearing losses (SNHL) are due to reduced sensitivity of the neural receptor that distorts the perception of sounds. The listeners with SNHL may suffer increased susceptibility to forward masking, making it more likely the vowels will mask energy in weaker adjacent consonants. SNHL are not amenable to medical intervention and patients need to use the hearing-aids for speech perception.

When confronted with difficult environments or when speaking to hearing-impaired person, humans instinctively change the way they speak and adopt a speaking style called clear speech. The acoustic analysis show that naturally produced clear speech typically involves a wide range of acoustic & articulatory adjustments referred as special attributes [1-3] such as decrease in speaking rate (longer segments), wider dynamic range, greater sound-pressure levels, more salient stop releases, greater rms intensity of non-silent portions (release burst, frication, and/or aspiration) of obstruent consonants.

Studies involving the difference between 'clear' and 'conversational' speech suggest that it may prove beneficial to attend to the temporal characteristics of speech for the benefit of SNHL hearing-impaired listeners. The clear speech modifications are aimed at providing the listener with more salient acoustic cues in the speech signal that may enhance their ability to access and comprehend the message. In this direction, the present paper focused on speech intelligibility enhancement based on Consonant Duration modifications (CDM).

2. EXPERIMENTATION

2.1 Speech Material

People suffering from hearing loss are often said to have greatest difficulty in identifying short speech sounds such as Stops and Fricative consonants. Hence nonsense syllables involving Stop consonants - /p, t, k, b, d, g/, and the Fricative consonants - /f, θ, s, v, ð, z/, with the accompanying cardinal vowels /a, ε, o/ and /a, i, u/ respectively were used as target stimuli.

The baseline stimuli were categorized as Stop-Vowel or /SV/ set and Fricative-Vowel or /FV/ set, with 18 syllables each grouped into two sub-sets, Voiceless sub-set and Voiced subset. The stimuli sets in /SV/ context were (i) Voiceless sub-set: /pa, p ε , po, ta, t ε , to, ka, k ε , ko/ / (ii) Voiced sub-set: /ba, b ε , bo, da, d ε, do, ga, g ε , go/ . Further in /FV/ context were , (i) Voiceless sub-set: / fa, fi, fu, θa, θi, θu, sa, si, su/and (ii) Voiced sub-set: /va, vi, vu, ða, ði, ðu, za, zi, zu/.

Stops are produced by first forming a complete closure in the vocal tract via a constriction at the place of constriction, during which there is a either silence or a low-frequency hum called ‘voice bar’. The vocal tract is then opened suddenly releasing the pressure built up behind the constriction; this is characterized acoustically by a transient and /or a short duration noise Burst [4]. The period between the release of the stop and the beginning of voicing in the vowel is called the Voice Onset Time - VOT. During this period there is a silence and/or aspiration noise. The time interval between the onset of the following vowel and the instance when a formant frequency reaches its steady-state value is called the Formant Transition Duration- FTD. These temporal segments are visualized in the spectrograms shown in Figure 1.

Fricatives are produced when the turbulent air-flow occurs at a point of constriction in the vocal tract. Fricative consonants are characterized by a turbulent Noise, and may consist of the noise alone or may consist of the noise together with vocal cord vibration. The main acoustic cue that has been reported to affect perception of fricatives include-Noise Duration-ND, as well as adjacent Formant Transition Duration-FTD. The ND is said to extend from fricative onset time to offset time. The time interval between the onset of the following vowel and the instance when a formant frequency reaches its steady state value is called Formant transition.

Fricatives and plosives bursts are both characterized by high frequency random noise, which occurs on the opening of oral cavity. Plosives are characterized by highly transient cues, release burst very brief; whilst the noise spectrum of a fricative is quiet a great deal longer and rises to its target amplitude more gradually than a plosive does.

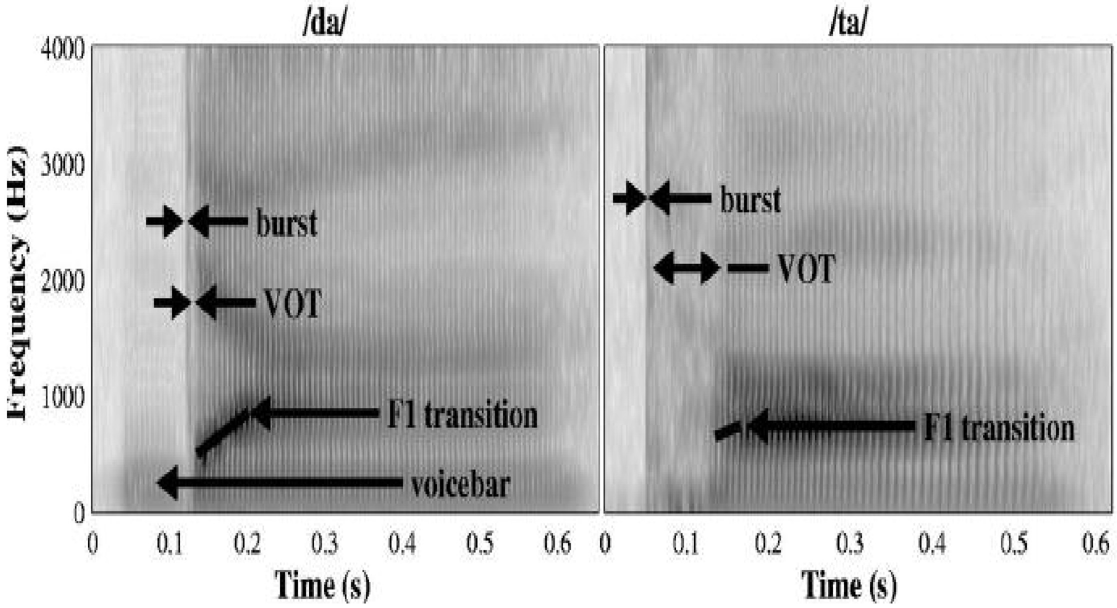

Figure 1. Spectrogram indicating acoustic segments of Voiced and Unvoiced Stops [5]

2.2 Listeners

The listeners were distinguished as Type 1 and Type 2 based on their hearing thresholds. Type 1 listeners were Hearing Impaired (HI) listeners whereas Type 2 were Normal Hearing (NH) listeners in simulated hearing impairment. Type 1 Listeners: One female and three male senior citizens in the age group of 58 to 62 yrs with hearing impairment ranging from moderate to severe loss (45 dB to 85 dB) in both ears, participated in the listening tests (as per PTA tests). Figure 2 displays the hearing thresholds for different stages of hearing loss. Type 2 Listeners: Two female and two male subjects with normal hearing with Pure Tone Audiometer (PTA) thresholds within 25dB of the normal hearing standards participated in the listening tests.

2.3 Speech Signal Processing

The experimentation spans into two phases, in the first phase the speech stimuli were processed for time-expansion using non-uniform algorithm for wave form expansion involving selection and modification of acoustic features. In the second phase the developed database was subjected to perception tests leading to the evaluation of speech intelligibility measures. A detailed explanation is as follows.

In the first phase, the recorded speech syllables were subjected to resynthesis using the procedure of LPC (linear prediction) analysis-synthesis as provided in PRAAT [7]. PRAAT (also the Dutch word for ‘talk”) is a free, multiplatform, scientific software program for the analysis of speech in phonetics. Resynthesis leads to synthetic copy rendering efficient and independent manipulation of the spectral, temporal and intensity characteristics. The LPC method is based upon the source-filter model; LPC analysis does it by estimating the vocal tract resonances from a signal’s waveform, removing their effects from the speech signal (inverse filtering) in order to get the source signal (or residue). After the process of resynthesis, the synthesized tokens (baseline syllables) were normalized to 70 dB IL to avoid the signal clipping in subsequent processing stages.

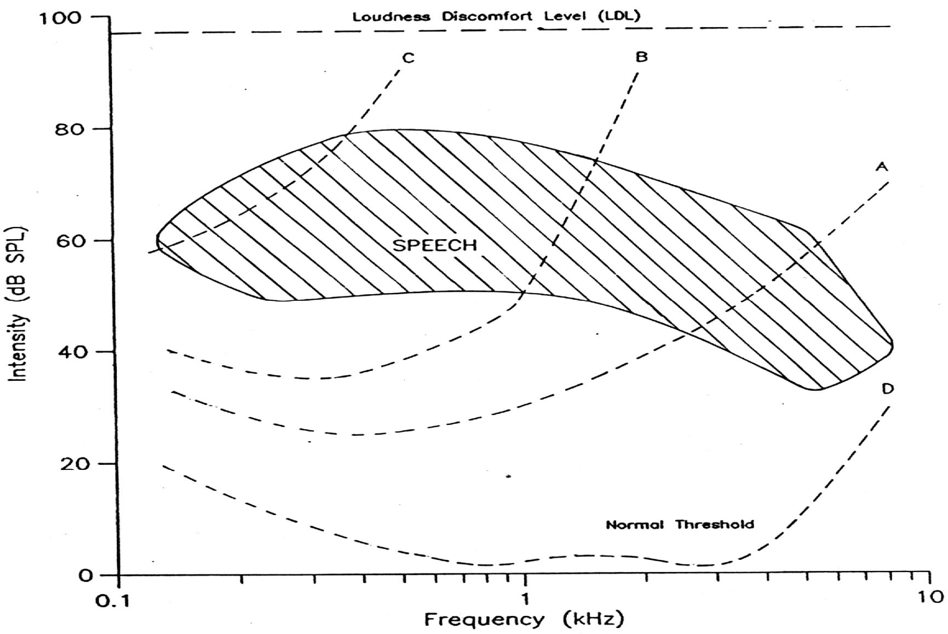

Figure 2. Sensorineural Hearing Thresholds: Hatched portion- Normal speech spectrum , Curve A-Mild to Moderate HL;B-Severe HL, C-Profound HL;D-Normal Hearing [6]

In the first stage of signal processing, baseline stimuli were temporally processed to generate stimuli sets at different levels of CD Modifications. The CD manipulations were based on PSOLA (Pitch-Synchronous Overlap and Add) algorithm using PRAAT. The basic PSOLA algorithm consists of three steps: analysis-modification-synthesis. The PSOLA analysis-modification-synthesis method belongs to the general class of STFT (Short-Time Fourier Transform) analysis-synthesis method, where the original pitch is being preserved during the processing [8]. PRAAT scripts were run for accomplishing all the above processing steps.

For /SV/ set, the acoustic segments for modification were: Burst Duration (BD), Voice Onset Time (VOT), and Formant Transition Duration (FTD). The release burst was identified as the short segment characterized as a 'spike' in the time domain and a sudden, sharp vertical line in the spectrogram [9]. The segmentation of a burst was performed visually by examining both the waveform and the spectrogram. VOT was identified as duration from the end of burst to the beginning of the vowel (the beginning of first waveform period) [10]. It is to be noted that the silence or closure interval of plosives cannot be defined for isolated CV syllables. The current stage of investigation reported that, the release burst was longer for voiceless than voiced plosives; VOT durations were longer for velars than alveolar, which in turn were longer than for labials; formant transitions were longer for voiced than voiceless plosives.

For /FV/ set, the main acoustic cues-Noise Duration (ND) and Formant Transition Duration (FTD) were reported to affect perception of fricatives [11]. ND is the high frequency noise measured as the difference between the fricative onset time and fricative offset time. The onset time is the point at which the high frequency energy appeared on the spectrogram and/or the point at which the number of zero crossings rapidly increased, while the offset time is the intensity minimum immediately preceding the onset of vowel periodicity, for voiceless fricative the earliest pitch period exhibiting a change in waveform from that seen throughout the initial frication, zero crossing of the preceding pitch was designated as fricative offset [12].

The Formant Transition Durations (FTD) were measured by simultaneous consultation of time domain waveform, spectrogram, linear-predictive coding (LPC) spectra, and Short-Time Fast-Fourier transform (ST-FFT) spectra [8] The LPC spectrum was constituted for a prediction order of 10 (at least twice as the number of spectral peaks that we want to detect), analysis window of 12.5 ms and 5 ms step, +6dB/octave filtering above 50 Hz. The three formants were originally located by examining the LPC spectra, FFT spectra, and spectrogram. The steady-state point of the vowel was centered at 100 ms after the onset. Formant analysis was performed for the detection of formant transition duration. After proper settings, formant contour was extracted and the formant values were written to a text file. Utilizing this data, the duration of the transitions and their onset and offset points were determined, and we then applied a time warp to all formants over the determined duration of the transition. The acoustic segmentations and measurements were done using PRAAT software. The temporal waveforms and spectrograms are presented in Figures 3- 5 for /SV/, Figures 6, 7 for /FV/ syllables below. The top panel displays the baseline (original) syllable while the middle panel 50% time expanded and the bottom panel 100% time expanded waveforms. The length of acoustic segments for /pa/ syllable was determined experimentally as Burst Duration: 4.5ms, VOT: 9ms, FTD: 30ms, hence time-expansions with 50%/ 100% (compared to original duration) are evident from the figures.

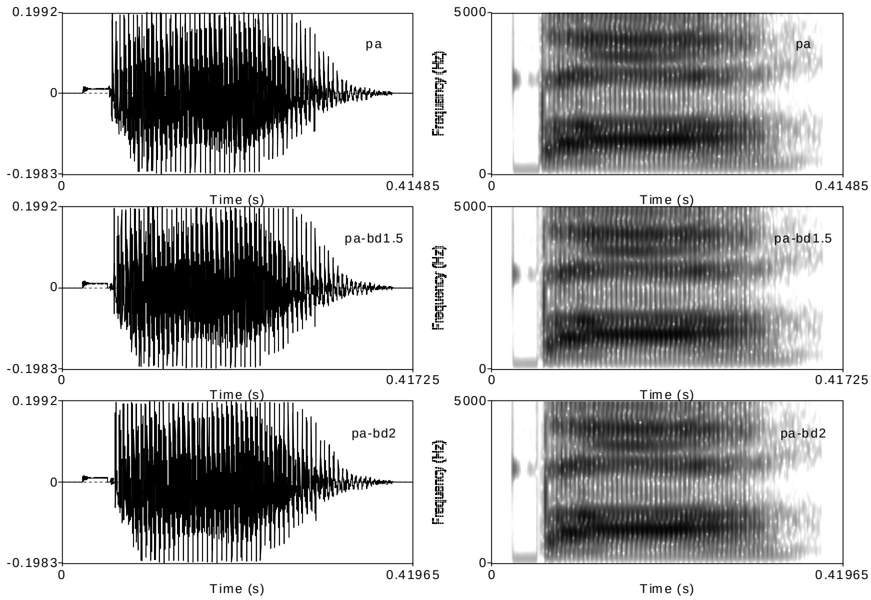

Figure 3. BDM Paradigm: Temporal waveforms and spectrograms for /pa/ syllable

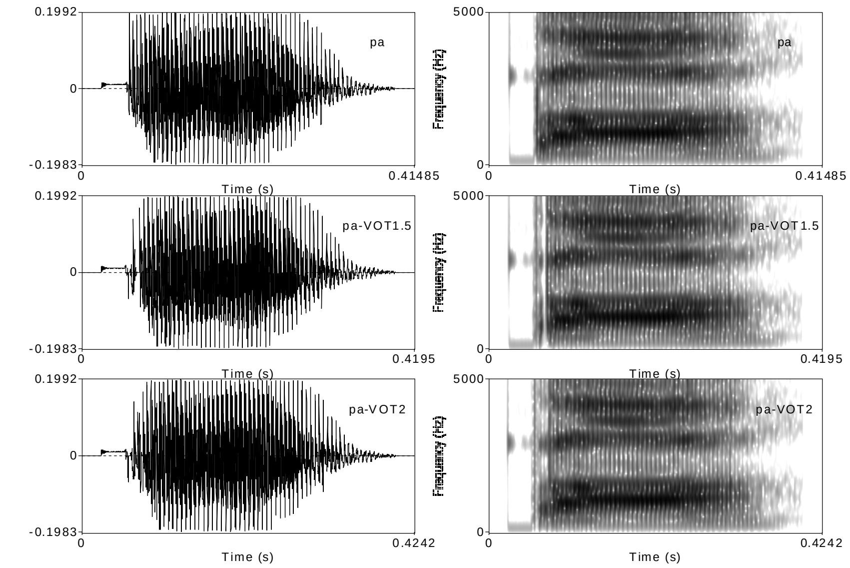

Figure 4. VOTM Paradigm: Temporal waveforms and spectrograms for /pa/ syllable

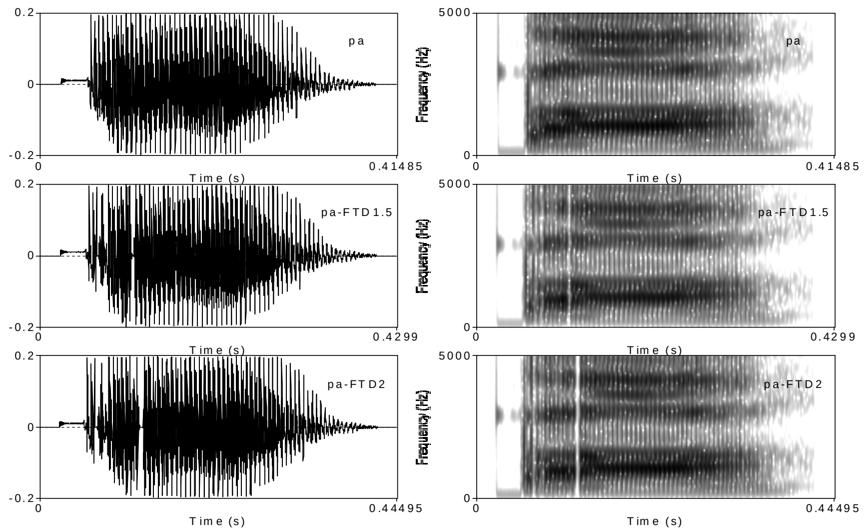

Figure 5. FTDM Paradigm: Temporal waveforms and spectrograms for /pa/ syllable

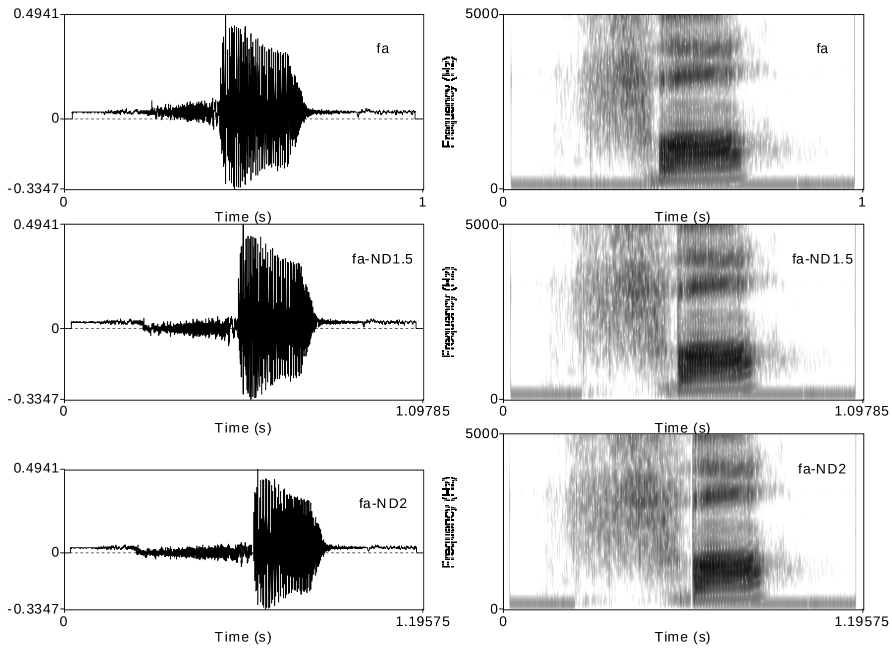

Figure 6. NDM Paradigm: Temporal waveforms and spectrograms for /fa/ syllable

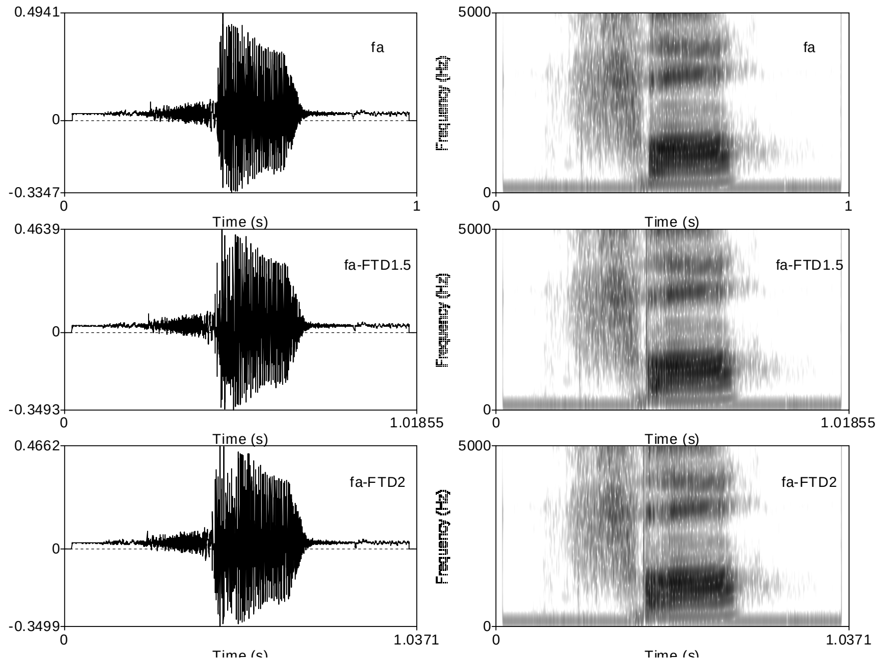

Fig.7 FTDM Paradigm: Temporal waveforms and spectrograms for /fa/ syllable

2.4 Speech in Noise Task

For Type 1 context, time-expanded stimuli from previous stage in the presence of no-noise served as test stimuli. The stimuli corpus for /SV/ context held 162 tokens (18*3*3) categorized under three schemes (BDM/ VOTM/ FTDM) and single listening no-noise condition ; also /FV/- CD modification held 108 tokens (18*3*2), categorized under two schemes (NDM/ FTDM) and single listening no-noise condition.

For Type 2 context, Hearing-Impairment in normal-hearing listeners was simulated. To provide a simple model for some of the perceptual effects of hearing impairment, gaussian white-noise masker was mixed with the stimuli. In other words, in order to simulate some effects of hearing impairment, three noise levels, no-masking noise, +12 dB and +6dB SNR were adopted. The PRAAT scripts were run for the generation of noise and for simulation process of mixing noise and stimuli. The stimuli corpus for /SV/ context held, 486 tokens (18*3*3*3) categorized under Three schemes (BDM/ VOTM/ FTDM) and 9(3*3) Listening conditions ; also /FV/ context held, 324 tokens (18*3*3*2), categorized under Two schemes (NDM/ FTDM) and 9(3*3) listening conditions.

3. EXPERIMENTAL SESSIONS

The second phase of the work utilized the processed speech data with the developed facility to evaluate the effects of time-expansion on speech perception in noise (SPIN). The SPIN experimentation was conducted using computerized testing procedure developed using MATLAB. Stimuli were presented using the computerized testing procedure at the most comfortable listening level of 75 dB to 85 dB SPL for the listeners. The test procedure used a similar protocol for all the experiments. Each experiment referred to single paradigm for eg BDM1/ BDM1.5/ BDM2 in no-noise/ +12dB/,+6dB masking level. Subjects were played token with ten randomized replications, they were prompted to choose from the set of choices displayed on the computer screen. Each run lasted for 20 min-25 min, spanning a period of nearly 4hrs for the entire experiment for one listener. Results were cast into three groups of six by six confusion matrices (CM) per run; sub-matrices (3*3) can be derived for analyzing the effect on the production-based categories.

3.1 Speech Intelligibility Measures

Speech discrimination test results were summarized as the percentage of correct responses for many experimental runs. The diagonal cell entries in the stimulus-response confusion matrix correspond to the correct responses and the off-diagonal entries correspond to the confusion errors. The results of each run were cast into stimulus-response confusion matrix, which were evaluated for percent-correct recognition data based on two intelligibility measures: Recognition scores and Relative information transmitted. While Recognition Scores are easiest to calculate and interpret, Relative Information Transmitted measures the covariance between the stimuli and responses and hence takes into account the relatedness of the two. The Relative information transmitted of the input stimulus X and output response Y are defined in terms of the mean logarithmic probability MLP given by,

$$I(X;Y) = -\sum_i \sum_j p(x_i, y_j)\ \log_2\left(\frac{p(x_i)\,p(y_j)}{p(x_i, y_j)}\right) bits \quad ..\ (1)$$

The Relative Information Transmission (RIT) from X to Y is given by ,

$$I_{tr}(X;Y) = \frac{I\ (X;Y)}{I_s(X)} \quad ..(2)$$

Where, $Is(x)$ is the information measure of the input-stimulus in terms of MLP.

3.2. Results of CD Modifications on Stop-Vowel Syllables

3.2.1. Speech Perception in Type- 1 Listeners

The consonant recognition scores averaged across Type 1 listeners quantized in terms of Relative Information Transmitted scores are presented in Table 1 below. The scoring pattern reported that BDM paradigm has significant improvement in 50% and 100% lengthening (w r t original length) for voiced and voiceless stops in vowel contexts - /a/, /e/ with maximum intelligibility benefit equal to +26 percent points . The VOTM paradigm reported significant improvement in 50% lengthening in vowel contexts - /a/ , /e/,/o, with maximum intelligibility benefit of +33 percent points. The FTDM paradigm reported low improvement under 50% for voiceless stops in /a/ context only, with intelligibility benefit equal to +16 percent points.

Table 1 CD Modifications **-** HI Subjects**-** / SV/ Syllables/**-** RIT scores pattern with BDM/VOTM/FTDM Modification schemes for voiceless/voiced syllables

Information transmission analysis:/SV/:Test CV9										
		Relative information transmitted (%)								
	Vowel Context	No-masking noise			No-masking noise			No-masking noise		
		BDM (%)			VOTM (%)			FTDM (%)		
Test Stimuli		0	50	100	0	50	100	0	50	100
Unvoiced stop-vowels	/a/	14	20	40	14	33	13	14	30	17
	/e/	44	53	43	44	47	50	44	40	33
	/o/	53	50	57	53	83	60	53	53	27
Voiced stop-vowels	/a/	59	60	80	59	60	33	59	33	43
	/e/	27	33	50	27	60	0	27	0	17
	/o/	100	20	23	100	20	0	100	10	3

3.2.2. Speech Perception in Type- 2 Listeners

The Consonant Duration Modifications on /SV/ set have established mixed results**.** The conclusion summary of paper [13] based on the perceptual and statistical results suggested that BDM by 50% (compared to original duration) reported significant intelligibility benefit, while VOTM by 50% and 100% could not report significant benefit in the presence of noise, hence may be detrimental. The third scheme-FTD lengthening by 100% reported significant benefit in the presence of no-noise, and +12dB noise**.** The above results suggested that of the three acoustic segments considered here, lengthening BD and FTD yielded positive results. At lower SNRs, Burst Duration and Formant Transition Duration are found to be dominant cues for lengthening consonant duration. A detailed analysis is presented in previous papers [13].

3.3 RESULTS OF CD MODIFICATIONS ON FRICATIVE-VOWEL SYLLABLES

3.3.1 Speech Perception in Type-1 Listeners

The consonant recognition scores averaged across listeners and in terms of Relative Information transmitted scores are presented in Table 2. The scoring pattern reported no significant benefit with FTDM paradigm while NDM paradigm reported low benefit in 50% & 100 % with vowel context/i/ only.

3.3.2 Speech Perception in Type- 2 Listeners

The conclusion summary of paper [14] based on the perceptual and statistical results suggested that Consonant Duration Modifications on /FV/ set have reported that fricative NDM reported no significant benefit in the absence or in the presence of noise; hence such alterations may be detrimental. On the other hand, FTDM by 50% (compared to original duration) reported consistent significant benefit in the presence of +6dB noise. The above results suggested that of the two acoustic segments considered here, Based on the consistency of scores, 50% FTD can be treated as a candidate for consonant duration modification in lower SNR (higher level of SNHL), but not lengthening ND. A detailed analysis is presented in the previous paper **[**14].

Table 2 CD Modifications- HI_Subjects- / FV/ Syllables: RIT scores pattern with NDM/FTDM modification schemes for voiceless/voiced syllables

Information transmission analysis :/FV/:CV9							
	Vowel Context	Relative information transmitted (%)					
		No-masking noise			No-masking noise		
		NDM (%)			FTDM (%)		
Test Stimuli		0	50	100	0	50	100
Unvoiced Fricative-vowels	/a/	93	77	57	93	87	100
	/i/	50	60	70	50	23	23
	/u/	73	77	67	73	60	73
Voiced Fricative-vowels	/a/	100	83	93	100	77	83
	/i/	57	50	40	57	3	13
	/u/	100	97	90	100	87	77

4. DISCUSSION

In spite of not a very encouraging background, an attempt to perform duration modifications was being investigated in current paper. As it was felt that it would be more effective to apply temporal modifications to enhance speech intelligibility only to consonant distinctions that are cued by duration differences. Thus, it is worthwhile to explore a few of recent studies in correlation with the present investigational results. However, it is crucial to understand that the variability in benefits across different studies depends on speaker- listener effects, signal-dependant effects, implementation of clear speech strategies etc.

Krause and Braida [15] assessed the role of speaking rate on intelligibility by eliciting clear speech at normal (conversational) speaking rates naturally, thereby avoiding some of the pitfalls of signal-processing techniques. Clear speech at normal/conversational speaking rates increased intelligibility for normal hearing listeners with simulated hearing losses(in noise), albeit to a slightly smaller degree compared to clear speech at slow rates. The intelligibility tests indicated 18-percentage-points advantage for clear/slow relative to conv/normal speech; while a 14-point advantage for clear/normal relative to conv/normal speech, and 12-point advantage for clear/slow relative to conv/slow speech. The study concluded that at normal rates, none of the alternative speaking modes tested provided as large or as consistent of an intelligibility advantage over conversational speech as clear speech.

Liu and Zeng [16] employed time-scaling algorithms to digitally compress (clear sentences) and stretch (conversational sentences) the experimental stimuli by inserting silent intervals into the conversational sentences to cause the overall duration to be equal to that of the clear speech. They concluded that, time compression was found to be more detrimental than time stretching in terms of processing reversibility and degree of performance degradation. Consistent with previous findings, they found that clear speech was more advantageous over conversational speech for normal-hearing listeners. In contrast to previous studies, they found that an increase in the number of pauses has a beneficial effect on intelligibility.

Smiljanic and Bradlow [17], [18] conducted an in-depth comparison of clear speech production and perception in English and Croatian. Sentences-in-noise test results showed that spontaneously produced clear speech enhanced intelligibility equally for both English and Croatian listeners in their respective native languages. Acoustic analyses revealed that talkers of both languages enhanced the overall salience of speech signal through a decrease in speaking rate, expansion of pitch range, expanded F1 x F2 vowel space.

Maniwa and Jongman [19] conducted experiments to characterize the adaptations in the clear speech production of American English fricatives in a carefully controlled range of communication situations. The study demonstrated that there were systematic acoustic-phonetic modifications in the production of clear fricatives. There were consistent overall style effects, several of which (consonant duration, spectral peak frequency, and spectral moments) were consistent with previous findings and a few (notably consonant-vowel intensity ratio) of which were not.

In the present investigation, time expansion on consonants demonstrated that, for the hearing-impaired listeners longer stops (with Burst Duration alteration) has been beneficial for intelligibility improvement, with a benefit of +26 percent points(for stop consonants), while the rest of the modifications such as VOT alteration (for stop consonants) and Formant transition alteration(for stops and fricatives), Noise Duration alteration (for fricative consonants) degraded consonant intelligibility or lead to perceptual errors. On the other hand, for normal listeners in simulated impairment ,Burst Duration by 50% (for stop consonants), FTD alteration by 100% (for stop consonants) , FTD by 50% (for fricative consonants) reported significant benefits. Reliable intelligibility advantages were not reported for VOT alteration (for stop consonants) and Noise Duration alteration (for fricative consonants). The tests reported maximum intelligibility benefit in relative information transmitted for overall consonant recognition as equal to +13 percent points (for stop consonants with BD lengthening), +8 percent points (for stop consonants with FTD lengthening), and +25 percent points (for fricative consonants with FTD lengthening).

5 CONCLUSIONS

Finally, Based on the correlation of experiment results for HI and NH listeners, it was asserted that only longer stops (with Burst Duration alteration) as beneficial cue for intelligibility. However, VOT alteration (for stop consonants) and Noise Duration alteration (for fricative consonants), Formant Transition Duration (for stops and fricatives) degraded consonant intelligibility. Hence the findings suggested that efforts to emphasize potentially weak consonant duration through ‘Burst duration’ should be beneficial in surmounting some of the speech recognition difficulties of hearing impaired listeners with sensorineural hearing loss. The finding also suggested the fact that hearing-impaired listeners may have difficulty integrating amplitude and spectral cues, and may generally place less weight on formant transitions than normal hearing listeners. In agreement with the previous work[20}, it can be concluded that the listeners with sensorineural hearing loss have elevated thresholds, and reduced dynamic range in regions relevant to fricative perception.

REFERENCES

[1] Liu, S., and Zeng, F., “Temporal properties in clear speech perception”, Journal of Acoustic Society of America, 120(1), 424-432, 2006
[2] J C Krause & L D Braida, “ Investigating alternative forms of clear speech:the effects of speaking rate and speaking mode on intelligibility”, Journal of Acoustic Society of America., vol 112, 2165-2172, 2002.
[3] J C Krause, L D Braida., “Accoustic properties of naturally produced clear speech at normal speaking rates”, Journal of Acoustic Society of America., vol 115 (1), 362-378, 2004.
[4] Stevens, K.N.: “Acoustic phonetics”, MIT Press, Cambridge, MA, 1998.
[5] Jintao Jiang, Chen, M., and Alwan, A., “On the perception of voicing in syllable-initial plosives in noise”, Journal of Acoustic Society of America , 119(2),1092-1105, 2006.

[6] Chaba, "Speech perception aids for the hearing-impaired people: current status & needed research", Journal of Acoustic Society of America , 90, 637-685, 1991.
[7] P Boersma, & D Weenik, PRAAT: Doing Phonetics by computer, http://www.praat.org
[8] Liu, S. and Fan-Gang Zeng, "Temporal properties in clear speech", Journal of Acoustic Society of America , 120(1), pp 424-432, 2006.
[9] Jiang, J. Chen, M. and Alwan, A, "On the perception of voicing in syllable-initial plosives in noise", Journal of Acoustic Society of America , 119(2), pp 1092-1105, 2005.
[10] Thomas, T.G., " Experimental Evaluation of improvement in speech perception with consonantal intensity and Duration modification", Ph.D. Thesis. Department of Electrical Engineering. Indian Institute of Technology, Bombay, 1996.
[11] Van Heuven, V.J.. "Reversal of the rise-time cue in the affricative-fricative contrast: An experiment on the silence of sound", The Psychophysics of speech perception, Martinus Nijhoff, Dordrecht, pp 181-187. 1987.
[12] Jongman,A.,Spence,M.,Wang,Y.,Kim,B., and Schenck,D., "Perceptual properties of English fricatives', Journal of Acoustic Society of America.,105, 1401(A),1999.
[13] Shobha, N.H., Thomas, T.G., and Subbarao, K., "Effect of Consonant Duration Modifications on Speech Perception in Noise-I", International Journal of. Electronics Engineering (ISSN NO. 0973-7383), Vol-I, Issue-II, pp 179-184, 2009.
[14] Shobha, N.H., Thomas, T.G., and Subbarao, K. "Effect of Consonant Duration Modifications on Speech Perception in Noise- II", International Journal of. Electronics Engineering (ISSN NO. 0973-7383). 2010.
[15] Krause, J.C., & Braida, L.D. (2002). "Investigating alternative forms of clear speech: the effects of speaking rate & speaking mode on intelligibility", Journal of Acoustic Society of America, 112, 2165-2172.
[16] Liu, S., & Zeng, F.(2006), "Temporal properties in clear speech perception", Journal of Acoustic Society of America 120(1), 424-432.
[17] Smiljanic, R., & Bradlow, A.R. (2005), "Production & perception of clear speech in Croatian & English", Journal of Acoustic Society of America., 118(3), 1677-1688.
[18] Smiljanic, R., & Bradlow, A.R.(2008), "Stability of temporal contrasts across speaking styles in Croatian & English", Journal of Phonetics., 36(1), 91-113.
[19] Maniwa, K., & Jongman, A. (2009). "Acoustic characteristics of clearly spoken English fricatives", Journal of Acoustic Society of America, 125(6), 3962-3973.
[20] Dubno, Dirks, & Langhoffer, "Evaluation of hearing-impaired listeners using nonsense syllable test II. Syllable recognition & consonant confusion patterns", Journal of Speech & Hearing Research, 25, 141-148,1982.

BIOGRAPHY OF AUTHOR

Dr N H SHOBHA obtained her Bachelor's in Electronics and Communication Engineering from Mysore University, Master's in Power Electronics from Bangalore University. She obtained her PhD from Osmania University, Hyderabad in Speech Perception/ Psychoacoustics' domain. Her thesis titled 'Effects of Consonant-Vowel Intensity Ratio and Consonant-Duration Modifications on Speech Intelligibility for the Hearing-Impaired' addressed the social problem of hearing-impairment and focused on developing appropriate speech signal processing mechanisms to improve speech perception for the advantage of those individuals. She has rich experience in teaching, research and administration in Karnataka and Hyderabad. Presently, working as Professor and Head in ECE dept. at Methodist College of Engineering, Hyderabad. Her expertise includes Speech signal processing, Digital signal processing, Signals and systems, Microprocessors, Automatic control systems. She has published 13 research papers indexed in IEEE,IEI, AES . She is the life member of ISTE, JASA, JASI