EXPERIMENTAL EVALUATION OF IMPROVEMENT IN CONSONANT RECOGNITION FOR THE HEARING-IMPAIRED LISTENERS: ROLE OF CONSONANT-VOWEL INTENSITY RATIO

[1] NH SHOBHA, [2] TG THOMAS, [3] K SUBBARAO

[1]Research Scholar, Department of ECE, Osmania University, Hyderabad, India-500007

[2] Prof., Department of ECE, Osmania University, Hyderabad, India-500007

[3] Prof., Department of ECE, Osmania University, Hyderabad, India-500007

Shobhanh7@gmail.com ,tgth@hotmail.com, kakarlasubbarao@yahoo.com

ABSTRACT

One of the most common complaints expressed by individuals with hearing loss is the difficulty in understanding speech when listening in background noise. This paper highlights the importance of enhancement of an important acoustic attribute of clear speech -'Consonant-Vowel intensity ratio'. A case for synthetic clear speech in the context of hearing impairment was developed. Consonant recognition in noise free and noisy situations using the Nonsense syllable Test (NST) was investigated in 5 normal hearing subjects under simulated hearing-impairment. The fricative consonants of English language in CV context were being processed for consonant-vowel intensity ratio modifications from 0 to +12 dB at +3 dB step. The Speech Perception In Noise (SPIN) tests were quantified in terms of information transmission analysis measures, in the presence of white noise-masker at three noise levels,0 dB, +12 dB, and +6 dB. C/V intensity ratio modifications of +6, to +12 dBs were found to improve speech intelligibility in simulated low level sensorineural loss. A maximum intelligibility benefit of 21% points was reported for +12 dB C/V modification, also the perception of consonant place of articulation reported better improvement of intelligibility compared to consonant voicing.

Keywords: Hearing loss, clear speech, speech perception in noise, consonant-vowel intensity ratio, information transmission analysis

INTRODUCTION

The speech that people hear is often degraded by the addition of competing speech and non-speech signals. People suffering from hearing loss often have the greatest difficulty understanding speech in noisy environments. When speech communication becomes difficult, a talker may adopt a different style of speech, 'clear speech': Such speech which is deliberately made clear is more intelligible than every day's speech style called conversational speech. Speakers naturally revise their speech when talking to impaired listeners or in adverse environments.

Clear speech and its effects on intelligibility improvements have been studied extensively for more than two decades. Those studies have demonstrated a significant intelligibility advantage for clear speech over conversational speech in both normal-hearing and hearing-impaired listeners across a wide range of listening conditions including quiet, noisy, and reverberant backgrounds [1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11]. The clear speech is said to include some of the special attributes like longer and more frequent inter-word pauses, more salient consonant contrasts (enhanced consonant-vowel intensity ratio, CVR), slower speaking rates, longer formant transitions, less vowel reduction [1, 12, 13, and 14].Preprocessing speech with those acoustic modifications is expected to improve speech intelligibility for impaired listeners [1, 12, 13, 14] and speech development in HI children [15]. Two important temporal attributes of clear speech that are found to increase at phoneme level are the consonant-vowel Intensity ratio (CVR) and the consonant duration (CD). The process of strengthening CVR and CD increases the salience of the consonant cues to weaken the masking effect or in other words results in reducing the vowel emphasis. Several works have reported that adjustment of the C-V intensity ratio (CVR) can yield significant improvement in consonant recognition for hearing-impaired (HI) listeners in quiet, and for normal hearing and HI listeners in noise. [16, 17, 18, 19, 20]. However, the exact acoustic cues those are responsible for the clear speech advantage remain largely elusive. The present study focuses on the role of consonant-vowel intensity (CVR) modification using synthesized speech syllables, on normal-hearing listeners under the simulated hearing impaired environment

In the past literature, even though clear speech attributes and their acoustic analysis as factors of intelligibility improvements were widely studied, they are quantified poorly. Each of those works offered a different explanation, insight, and a different set of limitations. A majority of them have employed natural/recorded utterances [16, 17, 18, and 21] as stimuli, though some speech technology approaches have been followed as well. Much research has predominantly focused on the perception of stop consonants [22, 23]. In contrast, fricatives have been studied in much less detail. Moreover, it is uncertain whether the classification metrics proposed for stop consonants can be successfully applied to fricatives. Our previous work has focused on the perception of stop consonants [22] and the current work extends the research on to fricatives by providing a detailed perceptual analysis and their statistical implications. The work quantified the consonant recognition by establishing, (i) perceptual comparison between voiceless and voiced consonants, (ii) Effect of vowel contexts on consonant recognition, (iii) perception of consonant-voicing (iv) perception of consonant place of articulation, in noise free and noisy environments.

People suffering from hearing loss are often said to have greatest difficulty in identifying short speech sounds such as stop consonants and fricative consonants. These consonants especially within the same class are often difficult to differentiate and are more vulnerable to signal degradations hence, it is desirable to strengthen the available acoustic cues to make consonant contrasts more distinct and potentially more robust to subsequent noise degradations. It is also widely known that consonants are less intelligible than vowels, as they are weaker in strength and shorter in duration compared to vowels [24].

1 METHOD

1.1Subjects and Target stimuli system

Three male and two females in the age group of 16-45 years with normal hearing, participated in the listening experiments. None of the subjects were experienced with perceptual experiments; subjects went through a speech token familiarization training session before the experiment started.

Nonsense syllables with consonant-vowel (CV) structure were chosen for investigation, The idea was to maximize the contribution of acoustic factors; minimize the impact of adjacent vowels, and to avoid the confounding effects of linguistic cues. The experimentation focused on most common and least intelligible,6fricative consonants, in CV context with cardinal vowels; comprising of /f θ s v ð z/ with primary cardinal vowels /a, i, u/. The fricatives span three different places of articulation, labiodentals (/f, v/), dentals (/ θ, ð/), and alveolar (/s, z/); which can be further classified as voiced {/v, ð , z/} voiceless {/f, θ, s /}.

1.2 Experimental setups

The experiment was analyzed under two setups: test CV9 and test CV6. Test CV9 was designed to test the effect of CVR modification (CVRM) on 'voiceless/voiced syllable' recognitions. The test- syllables were divided into two subsets of 9 syllables, voiceless subset: /fa, fi, fu, θa, θi, θu, sa, si, su/ and voiced subset: /va, vi, vu, ða, ði, ðu, za, zi, zu/ . Test CV6 was designed to test the effect of CVRM on production-based categories of 'place of articulation' and 'voicing'. The stimuli included voiced and voiceless consonants under a single vowel context forming three subsets of 6 syllables, context-/a/ :/fa, θa, sa, va, ða, za/; context /i/ :/fi, θi, si, vi, ði, zi/; and context /u/ : /fu, θu, su, vu, ðu, zu/.

1.3 Speech signal processing

In the first stage of signal processing, we recorded the natural speech tokens and subjected them to resynthesis. The natural stimuli were recorded in a quiet room, sampled at 44.1K Hz, using a Praat monosound recorder. The best utterance out of 20 utterances of the first author (middle aged, female) was selected based on the phonetic clarity. The speech tokens were subjected to resynthesis using the procedure of LPC (linear prediction) analysis-synthesis as provided in PRAAT [25]. The idea behind the resynthesis was two-fold; firstly, the synthetic copy renders efficient and independent manipulation of the spectral, temporal and intensity characteristics; secondly, synthetic speech is as similar as possible to a human utterance. 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.

In the second stage of signal processing, consonant-vowel intensity (CVR) modifications were carried out on the baseline syllables set. CVR is referred as the difference in decibels between either the power/energy of the consonant and that of the adjoining vowel. CVR modifications (CVRM) can be achieved either by reducing the intensity of vowel or by increasing the intensity of the preceding consonant. The latter method is reported to be more efficient over the former [12, 13, 16, 17, 18, 20, and 26]. We manipulated the baseline syllables under five CVRMs: 0, +3, +6, +9, +12 dB, where 0 dB refers to the unmanipulated set (natural). In the process, the consonant and vowel segments were identified on simultaneous consultation with timing & spectrogram waveforms with repeated visual & auditory monitoring. The intensity of the vowel segment was fixed while that of the consonant segment was adjusted to the required CVR level in deriving CVR modified stimuli sets. CVR modification was restricted to +12 dB so as to avoid the possibility of weak-vowel cue [20]

The third stage of processing was designed to simulate the hearing impairment, by reducing the acoustic dynamic range. The masking noise responsible for the threshold elevation is believed to be predominantly of cochlear origin [27]. As reported in literature, the reduction in the hearing threshold can be approximately simulated by addition of white noise [27, 28]. Some researchers have employed multi-talker babble instead of white noise [29, 30, and 31]. However, due to its non-stationary nature, the effective masking it may provide during stimulus presentation is unpredictable. Hence, we decided to use white noise masker to model the hearing loss to a good approximation.

The CVR modified tokens from the previous processing stage were additively mixed with the synthesized noise at three noise conditions, i.e., no-masking noise, +12 dB and +6dB SNRs. The noise free (natural) tokens were considered to be as those at no-masking noise. The average power level of the speech token was fixed while that of the noise was adjusted for fixing the SNR to the required value. The SNR refers to the ratio of the average power in CV token to the average power of the noise token in decibels. The noise synthesis and the process of mixing were accomplished by PRAAT scripts [32]. The mixing algorithm summed up the sounds by point-to-point values, preserving real time across the time domains. Finally, the stimuli corpus holds 270 test tokens spanning across 18 baseline stimuli with 5 CVR modifications and 3 SNR conditions.

1.4 Experimental measures

Speech discrimination test results were summarized as the percentage of correct responses for many experimental runs. The sum of the diagonal elements in the stimulus-response confusion matrix gives the empirical probability of correct responses, known as recognition score RS (or articulation score). Though computation of RS is simple, it obscures the detailed and important information on the distribution of errors among the off-diagonal cells [33]; also it is sensitive to the subject's bias or chance scoring (an artificially high score). We adopted the information transmission analysis approach [20, 22, 26, and 33], which provides a measure of covariance between stimuli and responses, and takes into account the pattern of errors and the score in a probabilistic manner. The covariance measure of intelligibility can be applied to the sub matrices derived from the original matrix by grouping the stimuli in accordance with certain desired features [20, 22 and 33]. The information measures of the input stimulus X and output response Y are defined in terms of the mean logarithmic probability MLP, given by,

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

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

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

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

1.5 Experimental sessions

The perception tests were automated using a MATLAB code with graphic user interface. Stimuli were presented using a computerized testing procedure at the most comfortable listening level of 75 to 85 dB SPL for the listeners. The set of 270 tokens were divided into five blocks of 90 syllables with each syllable in a block assigned a CVR level between 0 to +12 dB in 3-dB steps, spanning three noise levels. Subjects were played tokens with ten randomized replications of each token; they were prompted to choose from the set of choices displayed on the computer screen. Each run lasted for 20-25 minutes, spanning a period of nearly 6-8 hours for the entire experimentation per listener. Results were cast into three groups of six by six confusion matrices (CM) per run; sub matrices (3*3) were derived for analyzing the effect on the production-based categories [33] such as consonant-voicing, place of articulation.

2 RESULTS AND DISCUSSION

The confusion matrices obtained were analyzed and quantified with perceptual (information transmission analysis) and statistical (two-tailed t-test) measures. The results of CV9 and CV6 test setups are presented below.

2.1 Test CV9

The results of the perceptual analysis (Table 1), statistical analysis (Table 2) are analyzed as follows. The scores in table 1were those averaged across the three vowel contexts /a, i ,u/ for individual subject; the last rows in the table indicates their means and standard deviations. The results suggest that there is an improvement in scores with respect to increasing CVRMs corresponding to voiceless and voiced syllables. The maximum benefit of +17%points (SNR +12 dB, CVRM +12 dB, voiceless fricatives), +21%points (SNR +6 dB, CVRM+12 dB, voiceless fricatives) indicated that the voiceless syllables exhibited better improvement than their voiced counterparts.

The statistical test (Table 2) presents the mean percent-correct recognition data, standard deviations (SD), probability value (p) and the corresponding statistical significance status of the perception test (Table 1). The processing factor examined the intelligibility benefit between the unprocessed speech and the processed speech, a benefit was treated significant at 0.05 levels; $0.01<p<0.05$ was accepted as indicative of moderate significance and $p<=0.01$ as high significance. Based on the statistical analysis, +12 dB SNR presentations reported significant intelligibility benefit ($p<=0.01$) corresponding to three CVRMs (+6, +9, +12 dB) for voiceless fricatives; while +6 dB SNR presentations reported no significant benefit.

2.2 Test CV6

The results of perceptual analysis (Table 3), statistical analysis (Table 4) under test CV6 are analyzed as follows. Table 3 scores represent the relative information transmitted pattern for consonant-voicing and place of articulation, being averaged across five listeners for individual vowel context. The scores exhibited close dependency between the vowel, place of articulation, and voicing. The maximum benefit of +24%points (/i/ context, SNR= +12 dB, consonant-voicing recognition), and +38%points (/i/ context, SNR +6 dB, consonant- place recognition), represent that perception of place of articulation were emphasized than their voicing counterparts.

Based on the statistical analysis (Table 4), +12 dB SNR presentations reported highly significant intelligibility benefit ($p<0.01$) for consonant place of articulation recognition corresponding to vowel context /i/ (+3, +6, +9, +12 dB CVRMs); similarly +6 dB SNR presentations reported highly significant intelligibility benefit ($p<0.01$) for place of articulation recognition corresponding to vowel context /i/ (+3, +6, +9, +12 dB CVRMs). The consonant voicing recognition reported no significant intelligibility benefit for all three vowel contexts. These results substantiate the fact that the place of articulation decisions have shown good improvement compared to voicing decisions in the presence of noise. CVR enhancement levels of +6, to +12 dBs have reported consistent intelligibility advantage corresponding to +12 dB SNR presentations, for a majority of the tested contexts. The voiceless fricatives reported better performance improvement compared to voiced fricatives; the vowel context /i/ reported good performance improvement compared to /a/ and /u/ contexts; the perception of consonant place of articulation was better compared to consonant voicing.

Table 1, Test CV9- Perceptual Analysis

Test Stimuli	Listener	Relative information transmitted(%)														
		No-noise					SNR=12 dB					SNR=6 dB				
		CVR (dB)					CVR (dB)					CVR (dB)				
		0	3	6	9	12	0	3	6	9	12	0	3	6	9	12
Voiceless Fricative-vowels	L1	97	97	100	100	100	88	100	100	100	100	84	100	100	100	100
	L2	100	100	100	100	100	91	100	100	100	100	62	85	78	97	97
	L3	100	100	100	97	100	78	100	88	94	100	66	97	93	93	100
	L4	100	95	93	100	97	77	97	100	96	96	100	100	95	97	100
	L5	99	97	100	97	100	73	63	89	98	94	71	59	76	68	90
	AVG	99	98	99	99	99	81	92	96	98	98	76	88	88	91	97
	SD	1	2	3	2	2	8	16	6	3	3	16	18	11	13	4
Voiced Fricative-vowels	L1	100	100	100	100	100	100	97	100	100	100	100	100	97	97	100
	L2	99	100	100	100	100	94	93	100	97	93	72	84	82	90	100
	L3	100	100	100	100	100	76	90	87	87	100	87	87	100	100	100
	L4	100	100	97	100	97	97	100	100	100	97	99	100	100	100	100
	L5	100	100	100	100	100	71	80	92	96	97	75	70	93	85	98
	AVG	100	100	99	100	99	87	92	96	96	97	87	88	94	94	100
	SD	0	0	2	0	2	13	8	6	5	3	13	12	8	7	1

Table 2, Test CV9-Statistical Analysis

Test	SNR (dB)	CVR (dB)	Mean	SD	Two Tailed t Test of Difference		
					t	p	Result
Voiceless Fricative-vowels	No-noise	0	99	1			
		3	98	2	-0.894	0.4424	NS
		6	99	3	0	1	NS
		9	99	2	0	1	NS
		12	99	2	0	1	NS
	12	0	81	8			
		3	92	16	1.375	0.2064	NS
		6	96	6	3.354	0.01	S
		9	98	3	4.449	0.0021	S
		12	98	3	4.449	0.0021	S
	6	0	76	16			
		3	88	18	1.114	0.2976	NS
		6	88	11	1.382	0.2043	NS
		9	91	13	1.627	0.1424	NS
		12	97	4	2.847	0.0216	S
Voiced Fricative-vowels	No-noise	0	100	0			
		3	100	0	NaN	NaN	ND
		6	99	2	NaN	NaN	ND
		9	100	0	NaN	NaN	ND
		12	99	2	NaN	NaN	ND
	12	0	87	13			
		3	92	8	0.732	0.4848	NS
		6	96	6	1.406	0.1975	NS
		9	96	5	1.445	0.1865	NS
		12	97	3	1.676	0.1323	NS
	6	0	87	13			
		3	88	12	0.126	0.9025	NS
		6	94	8	1.025	0.3352	NS
		9	94	7	1.06	0.32	NS
		12	100	1	2.229	0.0563	NS

(N=5, S=Significant, NS=Not significant, ND= Not defined)

Table 3, Test CV6- Perceptual Analysis

Vowel context	Feature		Relative information transmitted (%)														
			No masking noise					SNR=12 dB					SNR=6 dB				
			CVR (dB)					CVR (dB)					CVR (dB)				
			0	3	6	9	12	0	3	6	9	12	0	3	6	9	12
/a/	Voicing	Mean	98	96	93	93	100	74	74	80	86	98	73	80	78	83	91
		SD	5	9	10	15	0	24	26	28	20	5	30	26	32	27	21
	Place	Mean	98	99	99	98	98	84	88	95	94	96	84	84	92	95	95
		SD	4	2	2	4	3	23	17	8	6	4	16	16	13	9	11
/i/	Voicing	Mean	98	96	97	96	98	67	67	76	78	91	79	66	74	84	81
		SD	4	9	7	7	4	31	28	28	31	13	26	31	27	23	21
	Place	Mean	99	98	96	98	100	67	90	92	95	99	61	88	87	88	99
		SD	2	3	5	3	0	17	10	9	7	3	18	17	11	14	3
/u/	Voicing	Mean	97	97	92	93	100	72	78	81	76	77	69	69	69	70	81
		SD	7	7	18	9	0	26	16	12	26	18	25	21	23	23	21
	Place	Mean	98	98	99	98	99	80	82	76	94	92	75	88	87	86	95
		SD	4	4	2	3	1	11	13	18	9	6	21	14	9	6	8

Table 4 (a), Test CV6-Statistical Analysis for Consonant- Voicing recognition

FEATURE	SNR (dB)	CVR (dB)	Context-/a/					Context-/i/					Context-/u/				
			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test		
					t	p	Result			t	p	Result			t	p	Result
VOICING	No Noise	0	98	5				98	4				97	7			
		3	96	9	-0.434	0.676	NS	96	9	-0.454	0.6618	NS	97	7	0	1	NS
		6	93	10	-1	0.347	NS	97	7	-0.277	0.7885	NS	92	18	-0.579	0.579	NS
		9	93	15	-0.707	0.5	NS	96	7	-0.555	0.5943	NS	93	9	-0.784	0.455	NS
		12	100	0	NaN	NaN	ND	98	4	0	1	NS	100	0	NaN	NaN	ND
	12	0	74	24				67	31				72	26			
		3	74	26	0	1	NS	67	28	0	1	NS	78	16	0.439	0.672	NS
		6	80	28	0.364	0.725	NS	76	28	0.482	0.6429	NS	81	12	0.703	0.502	NS
		9	86	20	0.859	0.415	NS	78	31	0.561	0.5901	NS	76	26	0.243	0.814	NS
		12	98	5	2.189	0.06	NS	91	13	1.596	0.1491	NS	77	18	0.354	0.733	NS
	6	0	73	30				79	26				69	25			
		3	80	26	0.394	0.704	NS	66	31	-0.718	0.4929	NS	69	21	0	1	NS
		6	78	32	0.255	0.805	NS	74	27	-0.298	0.7731	NS	69	23	0	1	NS
		9	83	27	0.554	0.595	NS	84	23	0.322	0.7556	NS	70	23	0.066	0.949	NS
		12	91	21	1.099	0.304	NS	81	27	0.119	0.908	NS	81	21	0.822	0.435	NS

Table 4 (b), Test CV6-Statistical Analysis for Consonant- place of articulation recognition

FEATURE	SNR (dB)	CVR (dB)	Context-/a/					Context-/i/					Context-/u/				
			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test		
					t	p	Result			t	p	Result			t	p	Result
PLACE	No Noise	0	98	4				99	2				98	4			
		3	99	2	0.5	0.631	NS	98	3	-0.62	0.5524	NS	98	4	0	1	NS
		6	99	2	0.5	0.631	NS	96	5	-1.246	0.2481	NS	99	2	0.5	0.631	NS
		9	98	4	0	1	NS	98	3	-0.62	0.5524	NS	98	3	0	1	NS
		12	98	3	0	1	NS	100	0	NaN	NaN	ND	99	1	0.542	0.602	NS
	12	0	84	23				67	17				80	11			
		3	88	17	0.313	0.763	NS	90	10	2.608	0.0312	S	82	13	0.263	0.8	NS
		6	95	8	1.01	0.342	NS	92	9	2.906	0.0197	S	76	18	-0.424	0.683	NS
		9	94	6	0.941	0.374	NS	95	7	3.406	0.0093	S	94	9	2.203	0.059	NS
		12	96	4	1.149	0.284	NS	99	3	4.145	0.0032	S	92	6	2.141	0.065	NS
	6	0	84	16				61	18				75	21			
		3	84	16	0	1	NS	88	17	2.438	0.0407	S	88	14	1.152	0.283	NS
		6	92	13	0.868	0.411	NS	87	11	2.756	0.0248	S	87	9	1.174	0.274	NS
		9	95	9	1.34	0.217	NS	88	14	2.648	0.0294	S	86	6	1.126	0.293	NS
		12	95	11	1.267	0.241	NS	99	3	4.656	0.0016	S	95	8	1.99	0.082	NS

(N=5, S=Significant, NS=Not significant, ND= Not defined)

3 SUMMARY AND CONCLUSION

The CV9 and CV6 results indicated that variations in C-V ratio explained a great deal of the variation in the intelligibility of fricative consonants. The CVR enhancements can be implemented to improve the salience of perceptually important consonant portions, for the benefit of hearing-impaired listeners. Consistent with the previous work [22] on stop-vowels, the present work found a consistent clear speech advantage corresponding to +6 dB to +12 dB CVR levels in simulated low level sensorineural loss. Thus the findings in general are encouraging for the eventual incorporation of C/V processing into digital hearing aids.

In summary, the test reviewed above has reported improvement of consonant recognition, albeit by different amounts. It is noteworthy that the CVR enhancement paradigm leads to improved intelligibility; hence plays an important role in surmounting some of the speech recognition difficulties of hearing impaired listeners. The research has shown a much promising direction for our future investigations.

REFERENCES

[1] J C Krause & L D Braida, Investigating alternative forms of clear speech:the effects of speaking rate and speaking mode on intelligibility, J. Accoust. Soc. Am., vol 112, 2165-2172, 2002.

[2] A R Bradlow & T Bent, The clear speech effect for non-native listeners, J. Accoust. Soc. Am., vol 112, 272-284, 2002.

[3] S H Ferguson & D Kewley-port, Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners, J. Accoust. Soc. Am., vol 112. 259-271, 2002.

[4] A R Bradlow, N Kraus. & E Hayes, Speaking clearly for children with learning disabilities: Sentence perception in noise, *J. Speech Lang. Hear. Res.* vol 46, 80-97, 2003.

[5] C Liu & D Kewley-port, Vowel formant discrimination for high-fidelity speech, J. Accost. Soc. Am., vol 116, 1224-1233, 2004.

[6] J Gagne, C Qarengeer, P Folkeard, K Munhall, & V Mastern, Auditory, visual and audiovisual speech intelligibility for sentence-length stimuli: An investigation of

conversational and clear speech, *The Volta review,* vol 97, 33-51, 1995.

[7] L Payton, Karen, M Uchanski, Rosalie, & L D Braida, Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing, J. Accoust. Soc. Am., vol 95, 1581-1592, 1995.

[8] D Schum, Intelligibility of conversational and clear speech of young and elderly talkers, J. A. Acad. Audiol., vol 7, 212-18, 1996.

[9] R M Uchanski, S Choi, L D Braida, C M Reed, & N I Durlach, Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rates, *J.Speech and Hearing Res.,* vol 39, 494-509, 1996.

[10] K S Helfer, Auditory and auditory-vosual perception of clear and conversational speech, J. Speech Lang. Hear. Res.,vol 40. 432-443, 1997.

[11] F Chen, Acoustic characteristics of clear and conversational speech at segmental level, MIT, Cambridge, MA, 1980.

[12] M I Picheny, N I, Dulach, & L D Briada, Speaking clearly for the hard of hearing I: Intelligibility Differences between clear and conversational speech, J. Speech Lang. & Hear. Res., vol 28., 96-103, 1985.

[13] M I Picheny, N I, Dulach, & L D Briada, Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech, *J. Speech Speech Lang. & Hear. Res..* vol 29, 434-446, 1986.

[14] J C Krause, L D Braida, Accoustic properties of naturally produced clear speech at normal speaking rates, J. Accoust. Soc. Am., vol 115 (1), 362-378, 2004.

[15]N Schiavetti, R L Whitehead, & D E Metz, The effects of simultaneous communication on production and perception of speech. J. of deaf studies and deaf education, vol 3, 2004.

[16] R L Freyman, G P Nerbonne, The importance of Consonant vowel intensity ratio in the intelligibility of voiceless consonants, J. Accost. Speech Hear. Res. vol 32, 524-535,1989.

[17] R L Freyman, G P Nerbonne, & H A Cote, Effect of C-V ratio modification on amplitude envelope cues for consonant recognition, *J. Accost. Soc, Am.,* vol 34, 415-426,1991.

[18] E Kennedy, H Levitt, Neumann, & Weiss. C_V intensity ratios for maximizing consonant recognition by hearing impaired listeners, J. Accost. Soc, Am., vol 103 (2), 1098-1114,1998.

[19] A A Montgomery & R A Edge, Evaluation of Two speech enhancement techniques to improve intelligibility for hearing-impaired adults. J. speech and Hearing Res., vol 31, 386-393,1988.

[20] T G Thomas, Experimental Evaluation of improvement in speech perception with consonantal intensity and Duration modification, Ph.D. Thesis. Department of Electrical Engg. Indian Institute of Technology, Bombay, 1996.

[21] Hazan V, Simpson A, & Huckvale M, Proc.Intl. Conf. of speech and language processing, Sydney,Australia, Vol 5, 2163-2167, 1998.

[22]N H Shobha, T G Thomas, & K Subbarao, Influence of Consonant-Vowel Intensity Ratio on Speech Perception for Hearing Impaired Listeners, Proc. 2008 congress in Image and Signal processing , Hainan, CHINA, pp 342-346, 2008.

[23] Gordon-Salant (1986), Recognition of natural and time interval altered CV's by young and elderly subjects with normal- hearing, , J. Accost. Soc, Am.,, 80 (6), 1599-1607, Dec 1986.

[24] P Ladefoged, Vowels and Consonants: An introduction to the sounds of language, *Oxford: Blackwell Publishing, 2001*.

[25] P Boersma, & D Weenik, Praat: Doing Phonetics by computer *(Version 4.4.22), 2005.*

[26] R C Bilger, & M D Wang, Consonant confusions in patients with sensorineural hearing loss. J. speech and Hearing Res , vol 19, 718-748, 1976..

[27] S DeGennaro, L D Braida, & N I Dulach, Study of multi-band compression with simulated sensorineural hearing loss. J. Accost. Soc, Am., vol 69, S16, 1981.

[28] Oxenham, A.J. and Moore, B.C.J., Modeling the effects of peripheral nonlinearity in listeners with normal and impaired hearing', In W. Jesteadt (Ed.), Modeling sensorineural hearing loss Mahwah, NJ: Lawrence Earlbaum Associates) pp 273-288, 1987.

[29]Flecher, H, The perception of sounds by deafened persons, J. Acoust. Soc. Am.., 24. pp 490-497, 1952.

[30] S A Simpson, & M Cooke: Consonant identification in N-talker babble in a non monotonic function of N, J. Acoust. Soc. Am.., 118(5), pp 2775-2778, 2005.

[31] H D Lewis, V A Benignus, K E Muller, C M Malott, Babble and random noise masking of

speech in high and low content cue identifications, J. speech and Hearing Res , vol 31, 108-114, 1988.

[32]<http://www.lifesci.Sussex.ac.uk/home/ Chris_Darwin/praatscripts/>

[33]G A Miller, & P E Nicely, An analysis of perceptual confusions among some English consonants, J. Accost. Soc, Am.,vol 27(2), 338-352, 1955