Effect of Spectro-Temporal Enhancement on Speech Perception in Individuals with Cochlear Hearing Loss

Madhuri Sharma[1] & Animesh Barman[2]

Abstract

Current amplification devices for cochlear hearing loss have not proved to be beneficial in improving speech perception in noisy conditions as they fail to restore normal physiology of the auditory system and hence, the speech perception ability. Signal enhancement strategies might help them to improve their speech perception if the signals are presented at a comfortable level. Hence, the current study was taken up with a primary objective of comparing two strategies (companding & consonant enhancement) in the same population. The study consisted of normal hearing participants who served as control group (N=14) and individuals with cochlear hearing loss who served as clinical group (N=16). They were given a task of consonant identification for 19 consonants in the context of vowel /a/ which were presented in 3 conditions- unprocessed, companded and consonant enhanced at 5 SNRs (0, +5, +10, +15 and quiet). A significant improvement with consonant enhancement was seen at 15 dB SNR and 0 dB SNR for the control and clinical group, respectively. At lower SNRs, both the groups showed a significant improvement with increase in SNR. However, across SNR, control group performed like in quiet situations at 10 dB SNR itself whereas the clinical group required further reduction (15 dB SNR) in noise to obtain such results. Sequential Information Feature Analysis (SINFA) for CHL with flat and sloping configuration revealed maximum information transmission for manner cues followed by place and voicing cues in both groups. Consonant enhancement increased the spectral contrast in speech and hence proved beneficial in individuals with CHL in adverse listening conditions. Thus, it can be used as a rehabilitation technique in amplification devices.

Key words: *Speech perception, companding, consonant enhancement, cochlear hearing loss*

Introduction

Communication plays a key role in one's daily living. For a signal to be heard and understood, a sequence of events take place wherein, the ear transforms the sound waves into electrical signals and sends these nerve impulses (electrical signal) to the brain where they are processed and interpreted as sound. When the stimulus to be processed is speech, the auditory system must be capable of extracting certain cues that are essential for perception of any speech sound. They are, faster oscillations called temporal fine structure and the relatively slow varying envelope. It has been hypothesised that coding of this fine structure across the basilar membrane is by place coding, and that of the envelope is through phase locking by neural fibres (Rose, Brugge, Anderson & Hind, 1967; Joris & Yin, 1992).

Any alteration in the structure of the auditory system and its physiology will result in hearing loss which could be of conductive, sensorineural or mixed in nature. Sensorineural hearing loss (cochlear hearing loss, CHL), which is the most common type of pathology seen, is usually a consequence of loss or damage of outer hair cells (OHCs). OHCs are known to sharpen the auditory filters which in turn help in finer frequency discrimination. Abnormal structure or functioning of these cells would result in widening of auditory filters, as a result of which, many kinds of perceptual consequences can arise. They include, impaired frequency resolution, temporal resolution, and reduced sensitivity to low level sounds (Glasberg & Moore, 1986; Tyler, Wood & Fernandes, 1982; Thibodeau & Van Tasell, 1987; Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006). Owing to the same, their most often cited complaint is failure to comprehend speech, especially in noisy or reverberant conditions, or when more than one person speaks. The magnitude of problem is likely to increase with severity and change in configuration of hearing loss (Hornsby, Johnson & Picou, 2011).

While decoding speech, the auditory system represents its spectral shape as excitation pattern which resembles a slightly smoothed version of the input spectrum. Since individuals with cochlear pathology exhibit broader tuning curves, they produce a highly smoothened representation of the spectrum which makes it difficult to perceive fine acoustic information of speech. The reduced frequency resolution in cochlear pathology results in impaired discrimination of formant frequencies of vowels and consonants. On the other hand, reduced temporal resolution fails to quantify the subtle duration differences and hence results in poor discrimination of speech sounds (Schorn & Zwicker, 1990; Lorenzi et al., 2006). This problem worsens in the presence of background noise. Hearing impaired individuals require a higher signal to noise ratio (SNR) to understand speech when compared to normal hearing listeners (Glasberg & Moore, 1989). SNR value usually ranges from 2.5 dB for mild to up to 7 dB for moderate to severe degree of Sensorineural hearing loss. They also fail to make use of spectral dips in fluctuating background noise unlike normal hearing listeners, who are able to perform better

1. madhurisharma923@gmail.com
2. barmananimesh197@gmail.com

in steady state noise (Festen & Plomp, 1990).

To improve speech perception, most often, amplification devices are provided for individuals with CHL. However, they primarily compensate for loss of audibility. It was noted that, even when hearing aid sufficiently amplified speech well above the threshold for detection, speech perception did not improve significantly (Plomp, 1986). This could be attributed problems like reduced frequency and temporal resolution as seen in individuals with cochlear hearing loss.

Damage to the cochlea results in broadening of auditory filters whose consequences are reduced audibility, loss of temporal and frequency resolution (Moore, 1995). This in turn affects the spectral representation in the cochlea in individuals with cochlear hearing loss leading to poor speech perception. Baer, Moore and Gatehouse (1993) suggested that by pre-processing the signal to enhance spectral contrasts, the problem of reduced frequency and temporal selectivity can partially be overcome as it enhances those portions of the spectrum where the signal-to-noise ratio is highest (the peaks) and suppresses those where it is lowest (the valleys).

Consonant enhancement is one such signal enhancement technique which increases the spectral contrast (Guelke, 1987). Outcome measures of this technique were carried out in individuals with CHL of varying degree and configuration of hearing loss. They indicated an improved subjectively measured quality and intelligibility for sentences embedded in continuous background noise having the same long-term-average spectrum (Baer, Moore & Gatehouse, 1993). The technique also showed promising results in the similar population when CVCs were used as stimulus in quiet condition (Hazan, Simpson & Huckvale, 1998; Franck et al., 1999). Various other researchers (Summerfield, Foster, Tyler & Bailey, 1985; Simpson, Moore & Glasberg, 1990) examined spectral enhancement by narrowing bandwidths and digital signal processing and showed a slight improvement in intelligibility of speech in individuals with CHL.

Most of these studies used sentences as their stimuli in both quiet and noise. Sentences, being more redundant in nature also involve cognitive processing and do not represent the benefit at cochlear level in isolation. Hence, it is important to check the benefit of consonant enhancement with a less redundant stimulus like CVs which would give a clear picture of improvement from the technique at the level of cochlea alone. Although a few previous authors used CVCs, they evaluated in quiet condition alone, which does not give a measure of the improvement obtained in a noisy situation which occurs more commonly in daily life. As there is a dearth of information regarding use of CV as stimulus in quiet and noise with consonant enhancement techniques in individuals with CHL, the current study was taken up.

Another novel technique is companding (Turicchia & Sarpeshkar, 2005) which uses the concept of combination of two-tone suppression and dynamic gain control to increase spectral contrast. Investigators (Oxenham, Simonson, Turicchia, & Sarpeshkar, 2007; Bhattacharya & Zeng, 2007) examined the advantage of companded speech stimuli in cochlear implant listeners using simulation studies and observed a significant improvement in phoneme and sentence recognition tasks in the presence of steady-state noise. This technique also contributed to an improvement in both sentence and consonant identification tasks in quiet and 15dB SNRs in individuals with Auditory Neuropathy Spectrum Disorder (ANSD) (Narne, Barman, Deepthi & Shachi, 2014). A significant improvement in consonant and sentence recognition was found in persons with mild to moderately severe cochlear hearing loss (CHL), at lower SNRs for companded speech stimulus (Deepthi, 2012).

Spectro-temporal enhancement techniques mainly modify the regions of the signal that contain acoustic cues in order to make it more resistant to subsequent degradation. However, there has been only one previous study (Deepthi, 2012) which has examined the benefit obtained with companding in people with CHL. Hence, more number of studies are required to ensure consistent benefit in improving speech intelligibility in these individuals. Also, the current experiment involves only CV syllables (consonants in context of vowel /a/) in quiet and four other SNRs from 0 to 15dB with 6-talker babble as background noise. Whereas, various other studies used a combination amongst words and sentences which are more redundant in nature. Listening environments with noise at multiple intensity levels as compared to previous studies would be instrumental in comparing the amount of benefit in conditions ranging from most adverse to quiet situations. The use of 6 talker babble as competing signal is more close to noise encountered in a natural situation as opposed to prior studies which used steady state noise or speech shaped noise. Hence, these parameters were aimed at being tackled in this study.

Although companding and consonant enhancement differ in the method of enhancement, their ultimate goal is to enhance the cues available for identification of speech. Hence, it is also important to compare the benefit across the two in the same set of population. As there is a dearth of studies regarding this issue, there is a need to test the robustness of improvement provided by companding and consonant enhancement algorithms and also compare the same across different levels of noise in the same population to check which would be more beneficial for individuals with CHL. It is also required to identify the parameters that are critical in improving speech intelligibility. A further modification of those parameters can help to contribute to the technological advancement dedicated to improving speech perception in individuals with CHL.

Thus, the aim of the study was to evaluate the effect of spectro-temporal enhancement, using companding and consonant enhancement strategy, on speech perception in quiet and noise at different SNRs in individuals with normal hearing and cochlear hearing loss.

Thus, the present study was taken up with the following objectives.

- To evaluate the effect of spectrally and/or temporally enhanced speech stimulus using companding and consonant enhancement on speech perception across various listening environments in individuals with cochlear hearing loss and with normal hearing individuals as control group as well as between subgroups of individuals with CHL having flat and sloping configuration.
- To evaluate the relative benefit of processed signal at each listening condition between the groups.
- To compare the consonant identification scores at different SNRs within each signal condition and within each group.
- To compare consonant identification scores for different signal conditions within each SNR and within each group.
- To analyze the error patters of the consonantal phonetic features in terms of voicing, manner and place of articulation cues perceived in the two subgroups of CHL.

Method

Participants

The participants selected were divided into two groups. Those with cochlear hearing loss belonged to the clinical group while those with a normal auditory system were included under control group. They were all native speakers of Kannada.

Clinical group (Individuals with cochlear hearing loss): Fourteen Adults (16 ears) who were diagnosed as having post lingual acquired cochlear hearing loss were selected as participants for this group. The age range of the participants was from 23 to 55 years (mean age= 39.87). Their degree of hearing loss ranged from mild to moderately severe sensorineural hearing loss with either flat or gradually sloping configuration. They had speech identification scores that were in proportion to their pure tone thresholds.

All participants in the group had 'A' type tympanogram with ipsilateral and contralateral reflexes present elevated or absent depending on degree of hearing loss. Auditory brainstem response patterns were as expected with the severity of hearing loss and transient evoked otoacoustic emissions were absent indicating outer hair cell dysfunction, in conjunction with tympanogram results. None of them had any history or presence of neurological or middle ear related problems as reported.

The clinical group was further subdivided in to two subgroups based on their configuration- Subgroup I (individuals with flat hearing loss) and Subgroup II (individuals with sloping hearing loss).

Subgroup I (individuals with flat hearing loss): This subgroup consisted of ten adults (11 ears) in the age range of 23 to 45 years, having a mean age of 35.72. The degree of hearing loss ranged from mild to moderately severe sensorineural hearing loss with a flat configuration, i.e., the difference between thresholds of octave frequencies did not exceed 5-10 dB (Johnson, 1966; Davis, 1998).

Subgroup II (individuals with sloping hearing loss): The participants of this subgroup included four adults (5 ears) in the age range of 28 to 55 years and with a mean age of 49 years. The degree of hearing loss ranged from mild to moderately severe sensorineural hearing loss with gradually sloping configuration (Stephen & Rintelmann, 1978). The demographic and audiological details of all the participants of cochlear hearing loss group are shown in Table 1.

Table 1: Demographic and audiological details of subjects with cochlear hearing loss

Subjects (Test ear)	Age/Gender	Pure tone average	Configuration of HL	SIS in quiet (%)	Tympano metry	Acoustic reflexes	OAE
CHL1 (Right)	24y/M	40	Flat	88%	'A'	Absent	Absent
CHL2 (Left)	35y/M	45	Flat	92%	'A'	Absent	Absent
CHL3 (Right)	30y/F	56.2	Flat	84%	'A'	Present	Absent
CHL4 (Right)	36y/F	45	Flat	92%	'A'	Present	Absent
CHL5 (Left)	35y/M	36.25	Flat	84%	'As'	Present	Absent
CHL6 (Left)	23y/M	52.5	Flat	100%	'A'	Present	Absent
CHL7 (Left)	38y/F	40	Flat	96%	'A'	Present	Absent
CHL8 (Right)	42y/F	30	Flat	96%	'A'	Present	Absent
CHL9 (Left)	42y/F	32.5	Flat	92%	'A'	Present	Absent
CHL10 (Right)	45y/M	34.75	Flat	100%	'A'	Present	Absent
CHL11 (Right)	43y/M	42.5	Flat	96%	'A'	Absent	Absent
CHL12 (Right)	28y/M	52.5	Sloping	88%	'As'	Absent	Absent
CHL13 (Left)	52y/M	36.5	Sloping	92%	'A'	Absent	Absent
CHL14 (Right)	55y/M	51.25	Sloping	92%	'As'	Absent	Absent
CHL15 (Left)	55y/M	53.75	Sloping	88%	'As'	Absent	Absent
CHL16 (Right)	55y/M	37.5	Sloping	88%	'As'	Absent	Absent

Control group (Individuals with normal hearing sensitivity): Fourteen individuals (left=7, right=7 ears) were considered for the study. All the subjects had hearing sensitivity less than or equal to 15 dB HL (four frequency average pure tone threshold, 500 Hz, 1000 Hz, 2000 Hz & 4000 Hz), with no history or complaint of difficulty in understanding speech in noise. All of their SPIN scores were 60% and above at 0 dB SNR. Additionally, all the participants had 'A' type tympanogram with ipsi and contralateral reflex thresholds within 95dBHL, normal auditory brainstem responses at 90 dBnHL and presence of otoacoustic emissions in both ears. Through a structured interview, it was ensured that, none of them had any history or presence of neurological, otological or any other associated problems.

Stimulus generation

A set of 19 consonant- vowel (CV) non sense syllables in the context of the vowel /a/ (t*f*, £, r, n, m, v, j, l, 7, s, *f*, t, p, b, t, d, d, k, g) were digitally recorded thrice from a native Kannada speaker adult male. It was done in a sound treated room using a data acquisition system with 44.1 kHz sampling frequency and a 16 bit analogue to digital converter. Following this, test of goodness was carried out by giving the samples to 5 native Kannada adult listeners who rated the samples on a 5 point rating scale. One set of recorded unprocessed 19 CV syllables with maximum relative intelligibility scores were then considered for the study.

To serve as background noise, a six talker speech babble developed by Jain, Konadath, Vimal and Suresh (2014) was used. The selected syllable samples were mixed with this speech babble such that the target syllable was temporally aligned to the centre of the babble. Using MATLAB- 7.8, the target syllables were mixed to obtain signal to noise ratios (SNRs) of 0, +5, +10 and +15 dB. Thus, these files were labelled as 0SNR, +5SNR, +10SNR, +15SNR and quiet, respectively.

The next step was to spectro-temporally enhance the stimuli using companding and consonant enhancement technique. Spectro-temporal enhancement was chosen to be done after mixing of stimulus to duplicate real life situations where a signal reaching at the level of the ear would already be embedded in the surrounding noise.

Signal processing

Companding

The algorithm followed to carry out the spectro-temporal enhancement was based on that given by Turicchia and Sarpeshkar (2005). It was done using MATLAB- 7.8 (The Math Works, Natick, USA) software. The following is a description of the working of the algorithm.

The incoming signal was first passed through 50 independent frequency channels by a bank of relatively broad band filters. Every channel of the companding architecture had a relatively broad prefilter, a compression block; a relatively narrow band post filter and an expansion block. Following the initial filtering, it was subjected to envelope detection (ED), whose output along with compression index (n1) having a value of 0.3 determined the amount of amplitude compression the signal underwent at this second stage. The EDs consisted of a full wave rectifier with a first order low pass filter. This compressed signal was then expanded after being passed through a relatively narrow band-pass ûlter. The gain of the expansion block depended on the corresponding ED output and the ratio of (n 2- n 1)/ n 1. The n2 parameter of the algorithm is the expansion index and had a value of 1. The outputs from all the channels were then non-linearly summed to obtain the processed signal.

Consonant enhancement

The procedure was carried out using the method adapted by Guelke in 1987 as reference. UCL enhance software was used to process the incoming signal with consonant enhancement technique. The procedure consisted of an algorithm that automatically identified the location of vowels, nasals, fricatives and gaps based on broad class phonetic recognition system. The algorithm then increased the amplitude of the selected portion of the syllable up to the specified level of the vowel in normal speech. The following options, as represented in Table 2 were selected for the enhancement of the stimuli.

Table 2: *Table showing details of options chosen to enhance the consonant part of various consonants in the context of /a/*

Syllable	Options chosen (among Burs Fricative, Nasal and transitioı
Stops	Burst + Transition
Fricatives and Affricates	Fricative + Transition
Nasals	Nasal + Transition
Glides	Transition

For all the syllables, RMS amplitude gain was selected with an amplitude compression degree of 10 as recommended by the software. The RMS amplitude was used as it maintained an overall average of the non-silent portions of the signal which would not vary with additions of gaps due to variables like noise. This option was combined with amplitude compression to make sure that the increase in intelligibility is due to enhancement and not due to a general increase in signal to noise. Figure 1 gives a bird's eye view of the sound /s/ for unprocessed, enhanced and companded signal condition, in the context of vowel /a/.

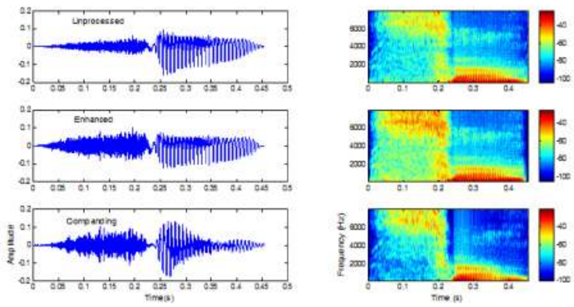

Figure 1: Figure showing the spectrum and spectrogram of the syllable /sa/ for all the three conditions without noise.

Following this, they were normalized along with the unprocessed stimuli in order to avoid any intensity differences amongst them serving as an unrequired variable. This was done using RMS amplitude normalization at -15dB in Adobe Audition software v5.

Data acquisition

The participants were comfortably seated in an air conditioned, double room situation with ambient noise levels within permissible limits (ANSI S-3, 1991). The listeners received the stimuli from headphones (TDH-50) which was routed through a calibrated diagnostic audiometer (MA-53) from a laptop. The responses were noted using MATLAB 7.8 in the same laptop.

Consonant identification task

Consonant identification scores were collected for different stimulus at different SNR conditions to test the objectives of the study. Nineteen CV stimuli were randomized by the software at each SNR. Thus, making it blind-folded for the tester in order to avoid tester bias. The CV identification was performed at five different listening conditions (0 dB, +5 dB, +10 dB, +15 dB SNR and quiet) for unprocessed, companded and consonant enhanced stimuli. All the listeners received the signal at the most comfortable level and they were instructed to repeat the consonant that was heard. Given below is the response screen that appeared. The tester clicked on the sound that the participant repeated.

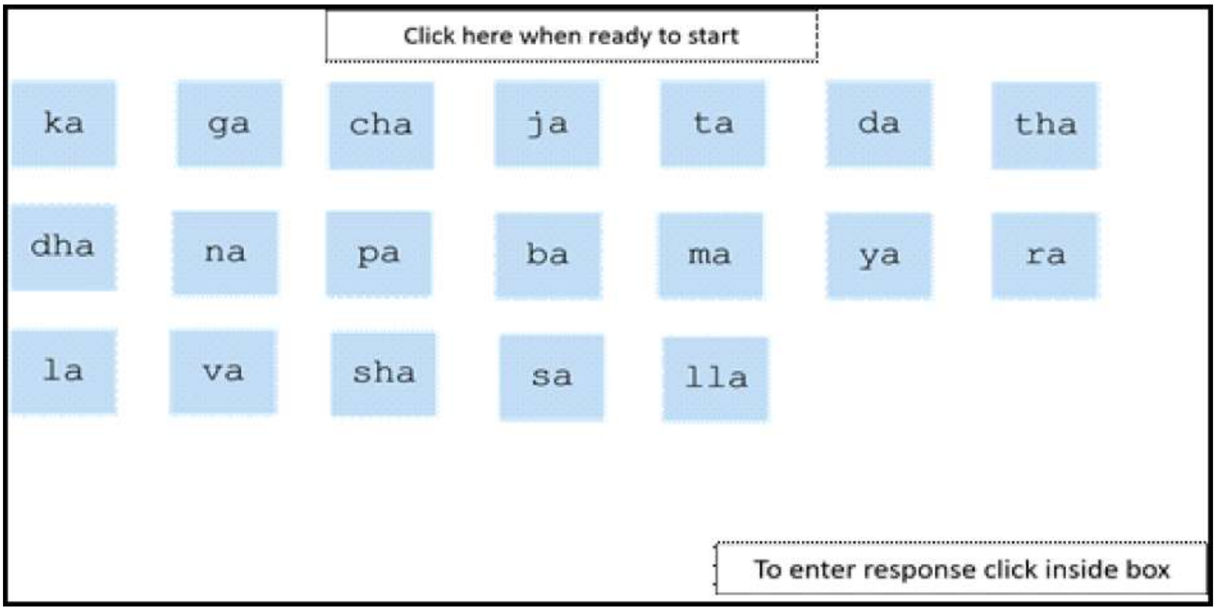

Figure 2: Response screen showing the arrangement of syllables considered for testing

Results

The consonant identification scores were tabulated and analyzed using Statistical Package for Social Sciences (SPSS, version 16.0). Non-parametric analysis was done as the data did not follow normal distribution. The following is a summary of the statistical analyses that was carried out.

Descriptive analysis, as represented in Figure 3, was done to obtain the median and standard deviation values separately for each of the three groups considered. Although both normal hearing individuals and those with cochlear hearing loss had maximum scores in quiet and minimum at 0 dB SNR condition, the corresponding values were higher in that of the normal hearing group. On an average, the clinical group had a range of 33% to 87% from 0 dB SNR to quiet condition while, the control group had a range of 67% to 99% indicating a better performance. The consonant identification scores deteriorated from quiet condition to 0dB SNR by 54% and 32% for the clinical and control group respectively. The effect of noise was more detrimental in the clinical group. Their scores at 0 SNR or maximum noise condition was 34% lesser than that of their normal hearing counterparts.

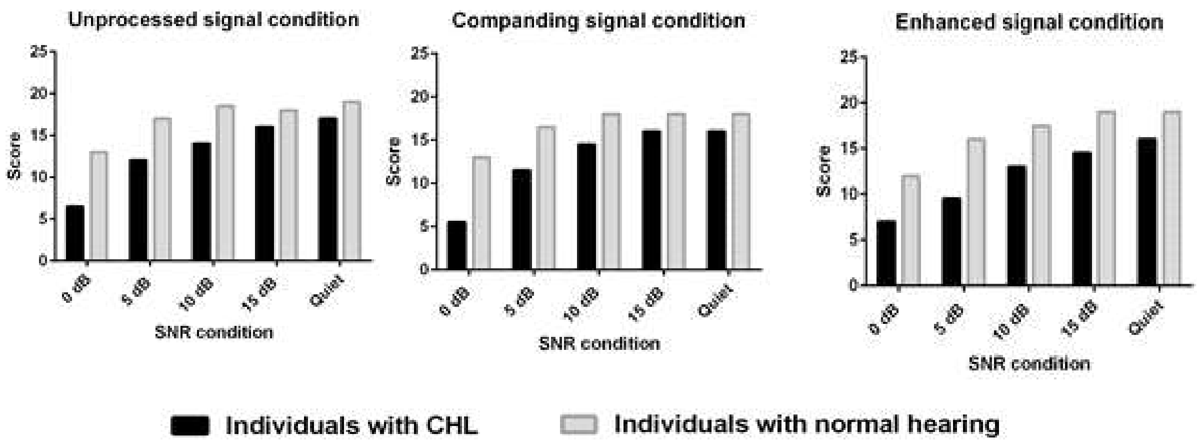

Figure 3: *Median values for consonant identification scores across SNRs in each signal condition for clinical and control group*

Mann-Whitney U test was administered to check if there existed a significant difference between the performances of groups. A significant difference was seen at all SNRs between the clinical and control group but not between the two subgroups of CHL. Hence, the subgroup data were thereafter combined and referred to as the clinical group collectively.

Comparison of consonant identification scores obtained at different SNRs within signal condition and group

Friedman test was administered for clinical and control groups separately which showed a significant difference in speech scores across SNRs within each condition for both groups (Table 3).

Table 3: *Results of Friedman test with X^2 and significance levels across all SNRs at each stimulus condition for individuals with normal hearing and CHL*

Population	Normal hearing		CHL	
Signal condition	$\chi^2_{(4)}$	*p* value	$\chi^2_{(4)}$	*p* value
Unprocessed	29.21	0.00	51.09	0.00
Consonant enhanced	39.39	0.00	55.29	0.00
Companding	36.67	0.00	53.54	0.00

Note: p<0.001, 2-tailed

Hence, Wilcoxon signed rank test was chosen to further evaluate which of the ten SNR pairs had a significant difference in the consonant identification scores. In the control group, the results are as indicated in the table below (Table 4). It can be noticed that as the SNR value increased, they were not significantly different from their adjacent SNRs. Unprocessed signal had the least amount of significant differences amongst higher SNR pairs, followed by enhanced condition, which showed significant differences in all SNR pairs except, 10 vs. quiet and 15 vs. quiet conditions.

In the clinical group, Wilcoxon signed rank test showed a significant difference (p<0.05) for all SNR conditions except quiet vs. 15 dB SNR in all three signal conditions with the higher SNR having better scores when compared to the lower SNR in the pair. The cells shaded darker indicate no significance in both the population.

Table 4: *Results of Wilcoxon signed ranked test showing significant differences for SNR pairs for each signal condition obtained in normal hearing individuals and CHL*

SNR pairs/ Signal conditions	Unprocessed	Companding	Enhanced
0 vs. 5	p < 0.05	p < 0.05	p < 0.05
0 vs. 10	p < 0.05	p < 0.05	p < 0.05
0 vs. 15	p < 0.05	p < 0.05	p < 0.05
0 vs. quiet	p < 0.05	p < 0.05	p < 0.05
5 vs. 10	p < 0.05	p < 0.05	p < 0.05
5 vs. 15	p < 0.05	p < 0.05	p < 0.05
5 vs. quiet	p > 0.05*	p < 0.05	p < 0.05
10 vs. 15	p > 0.05	p > 0.05	p < 0.05
10 vs. quiet	p > 0.05	p > 0.05	p > 0.05
15 vs. quiet	p > 0.05	p > 0.05	p > 0.05

*Note: p<0.001, 2-tailed. *p=0.054 indicates partial significance*

Comparison of consonant identification scores obtained across different signal conditions within each SNR and group

Friedman test for comparison of consonant identification scores across stimulus conditions at various SNRs for the control group revealed that a significant difference existed across signal conditions for only 15dB SNR and no other SNR (See Table 5). Further, Wilcoxon signed rank test showed that speech scores differed significantly (p<0.05) across unprocessed-enhanced and companded-enhanced signal condition pairs with consonant enhanced condition resulting in significantly higher scores in both pairs (See Table 6). On the other hand, Friedman test for the clinical group showed a significant difference across signal conditions at only 0 dB SNR (See Table 5). Wilcoxon signed rank test for the same indicated a significant difference (p<0.05) only between companded and enhanced signal conditions with consonant enhanced condition having significantly higher scores than companding (See Table 6).

Table 5: *Results of Friedman test with (df) and significance level across signal conditions at all SNRs for individuals with normal hearing and CHL*

	Normal hearing		*CHL*	
SNR (dB)	χ^2 (2)	*p*-value	χ^2 (2)	*p*-value
Quiet	3.80	0.150	3.800	0.150
15	11.619	0.003	0.122	0.941
10	2.47	0.290	4.66	0.097
5	3.17	0.205	2.33	0.311
0	1.73	0.420	6.87	0.032

Table 6: *Results of Wilcoxon signed ranked test for stimulus condition pairs in individuals with normal hearing at 15 dB SNR and at 0 dB SNR for individuals with CHL*

	Normal hearing			**CHL**	
15 dB SNR	Unprocessed	Companding	0dB SNR	Unprocessed	Companding
Unprocessed	-	-	Unprocessed	-	-
Companding	Not significant ($p>0.05$)	-	Companding	Not significant ($p>0.05$)	-
Enhancement	Significant ($p<0.05$)	Significant ($p<0.05$)	Enhancement	Not Significant ($p>0.05$)	Significant ($p<0.05$)

Sequential Information Transfer Analysis (SINFA)

This analysis was carried out to assess amount of information transmitted for each of the defined phonetic features, independently in subgroups of CHL. Patient responses which were obtained in the form of confusion matrices (See Table 7) were added for each individual using MATLAB 7.8 for each SNR across stimulus conditions. These matrices were then subjected to SINFA using the software Feature Information Xfer (FIX) (developed by University College of London, Department of Linguistics).

Table 7: *Example of a stimulus response matrix showing the results obtained for consonant enhanced condition at 0 dB SNR for 5 CHL participants with sloping configuration. The correct responses have been highlighted in the diagonal axis*

	b	tʃ	ḍ	d̲	g	k	L	l̲	m	N	p	r	s	ʃ	ṭ	t̲	J	ʤ	v
b	**2**										2								1
tʃ			2		1								1					1	
ḍ	1		**2**				1				1								
d̲			1	**3**			1												
g	3		1				1												
k				1	3	**1**													
l							**2**	1											2
l̲							2			2							1		
m	1								**3**										1
n			1	1			1		1	**1**									
p	1									1	**3**								
r	1		1	1				1											1
s		1									1		**2**	1					
ʃ		2												**3**					
ṭ	1		1												**1**	1		1	
t̲			1	1											2	**1**			
j																	**5**		
ʤ	1	1																**3**	
v	1																		**4**

The confusion matrices were subjected to SNIFA for assessing information transmitted for place, manner and voicing across conditions and SNR. The feature matrix was constructed with the phonetic features of each of the nineteen syllable considered. The same is represented in table 8.

Table 8: *Feature matrix of the 19 syllables considered*

	b	ḍ	G	ʤ	k	l̠	l	m	N	p	R	s	t̠	v	j	tʃ	d̠	ʃ	t̠
Voicing	+	+	+	+	-	+	+	+	+	-	+	-	-	+	+	-	+	-	-
Place	b	a	v	p	v	p	a	b	A	b	A	a	a	l	p	P	d	p	d
Manner	p	p	p	a	p	l	l	n	N	p	L	f	p	g	g	A	p	f	p

Note: Voicing: +=voiced, -=voiceless

Place: b=bilabial, a=alveolar, v=velar, p=palatal, l=labial, d=dental

Manner: p=plosives, a=affricates, l=laterals, n=nasals, f=fricatives, g=glides

The total information transmitted in this experiment, ranged from a minimum of 0 to a maximum value of 4.24. However, for each of the individual components like voicing, place of articulation (POA) and manner of articulation (MOA), the information transmitted ranged from a minimum of 0 and maximum of 1. The following Figure 4 represents the information transmitted for all the parameters in the two subgroups considered.

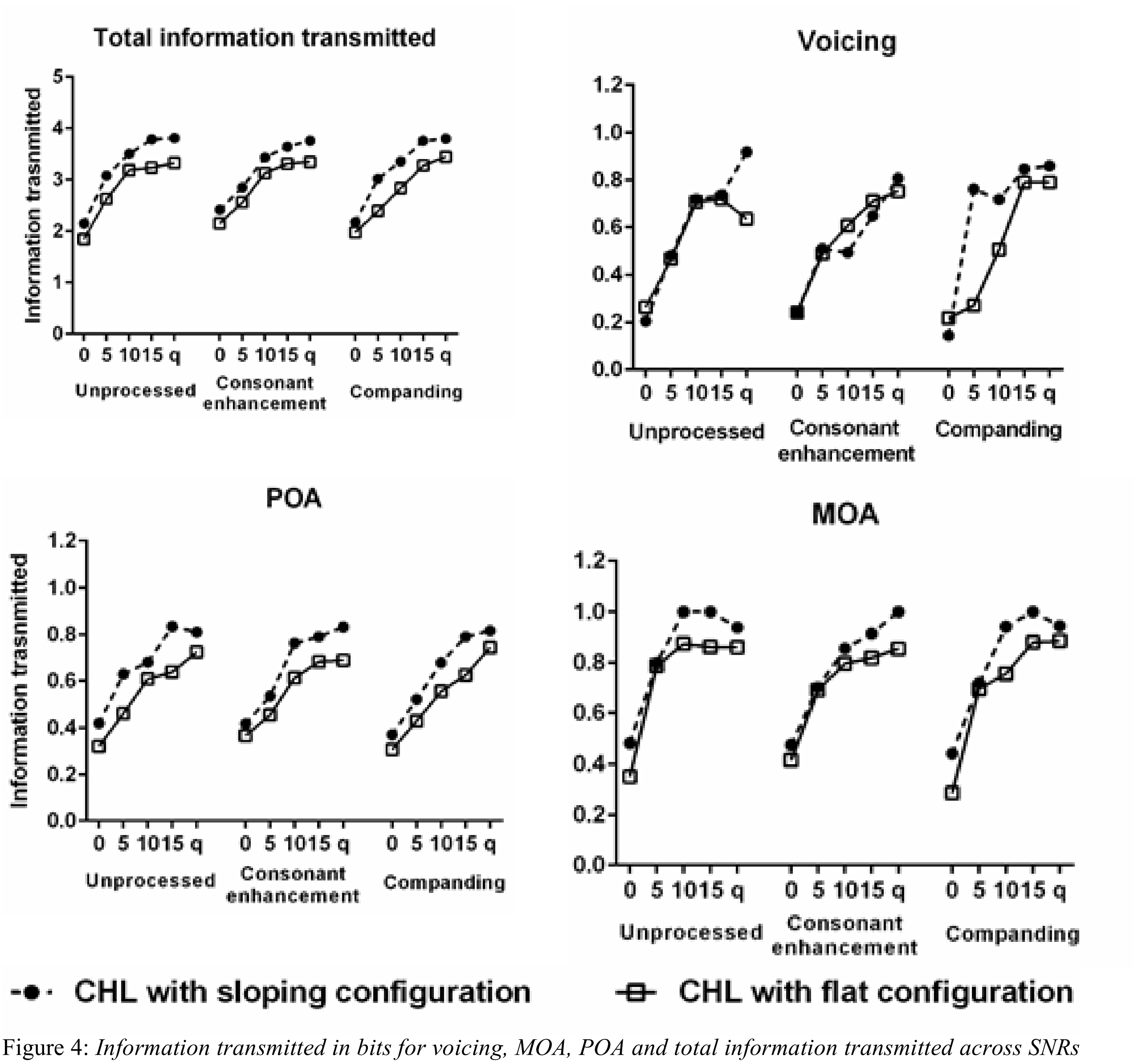

Figure 4: *Information transmitted in bits for voicing, MOA, POA and total information transmitted across SNRs and signal conditions for the two subgroups of CHL*

It can be inferred from Figure 4 that, the maximum information transmitted was for MOA cue followed by voicing and POA which were equally transmitted. The information transmitted across voicing and POA was similar within both individuals with flat and sloping configuration. Individuals with flat hearing loss, however, showed a higher information transmission for MOA cues with consonant enhancement condition only at 0 dB SNR. In rest of the SNRs, the information transmitted for MOA cues was similar within the subgroups of CHL.

Discussion

Effect of noise on consonant identification scores

This detrimental effect of addition of noise on speech perception has been well established in literature (Nabelek, Letowski, & Tucker, 1989; Loizou, Dorman & Tu, 1999). It is believed that the addition of a background noise reduces the distance between the peaks and troughs, thereby reducing the available spectral cues in order to identify speech. Hence, speech scores are poorer in the presence of noise (Baer & Moore, 1993). However, In spite of both the groups following the trend, the corresponding scores at each SNR were lesser for clinical group than the control group. This could be attributed to the classical features of cochlear pathology like, reduced audibility, reduced frequency selectivity and temporal resolution which has been well supported in literature as well (Eisenberg et al., 1995; Pekkarinen et al., 1990).

Descriptive statistics showed a significant difference in the consonant identification scores between clinical and control group, but not between the subgroups of CHL probably because all the participants in the group had a gradual slope. This is in accordance with findings in literature (Dubno, Dirks & Schaefer, 1987). Speech perception in sloping hearing loss was poorer than flat when the subjects considered had steeply sloping hearing loss or more. The subject performances on speech identification did not otherwise vary to a large extent up to moderate slope in configuration of hearing loss (Stephens & Rintelmann, 1978).

Comparison of consonant identification scores obtained across SNRs within each stimulus condition and group

In individuals with normal hearing a significant difference in consonant identification scores was noticed for lower SNR pairs but not for higher pairs. This improvement seen with increasing SNR can be correlated with the study done by Beattie, Barr and Roup (1997). They noted an improvement in the monosyllabic word identification scores as the level of multitalker background decreased from 5 dB SNR to quiet condition in individuals with normal hearing and CHL. There are several studies that showed that normal hearing individuals were able to extract spectral and temporal cues better than hearing impaired population even in noisy situations (Beattie et al., 1997; Pekkarison et al., 1990). This is attributed to the normal physiology in these individuals which is capable of differentiating the speech signal from competing background noise. Several mechanisms like medial olivocochlear bundle (MOCB) mediated suppression, two tone suppression and other nonlinearities of normal cochlea could aid in this process (Kumar & Vanaja, 2004). In companding condition, the lost spectral and temporal cues are made available to the listeners through processing of speech stimulus (Turicchia & Sarpeshkar, 2005). Hence, at 10 dB itself, these individuals perform almost like in quiet conditions.

However, when consonant enhancement was used, a possible distortion caused by the processing of stimulus would have led to significantly poorer consonant identification scores than in quiet even as noise reduced from 10 to 15 dB SNR. The possibility of addition of spurious artefacts due to processing of speech stimulus has been documented in literature as well (Lim, 1983). This distortion in companding signal affected normal hearing listeners to a lesser extent as companding restores both spectral and temporal cues as opposed to consonant enhancement that only enhances spectral cues. Therefore, due to the above reasons, normal hearing individuals could extract cues and perceive speech even in noise. Their consonant identification scores at 10 dB were similar to the scores in quiet situation in all stimulus conditions.

In individuals with CHL all the three conditions showed a significant difference in consonant identification scores in all SNR pairs except 15 dB SNR vs. quiet. Individuals with cochlear hearing loss have greater effects of noise than normal hearing listeners (Dubno & Schaefer, 1995; Pekkarinen, et al., 1990). While normal hearing listeners might be able to extract speech cues and understand speech like in quiet situations even at a noise level of up to 10 dB NR, individuals with CHL would still suffer poor perception of speech because of widened auditory filters and reduced temporal resolution. Literature reports that, speech perception in normal hearing individuals was not significantly affected until 0 dB SNR whereas, individuals with CHL required the SNR to be improved by 4-12 dB in order to obtain scores that are comparable to normal hearing listeners (Crandell & Smaldino, 2000). Hence, at SNRs that were equal to 15 dB SNR or greater, individuals with CHL were able to perform as well as in quiet situations.

Comparison of consonant identification scores obtained across different stimulus conditions within each SNR and group

In individuals with normal hearing there was a significant difference seen in consonant identification scores only at 15 dB SNR across enhanced- unprocessed and companding- enhanced signal conditions with consonant enhanced condition providing more benefit in both condition pairs. Previous studies have shown minimal or no benefit for individuals with normal hearing in consonant enhanced condition (Bunnell, 1990; Summerfield et al., 1985; Stone & Moore, 1992). The normally functioning auditory system is already capable of extracting spectral and temporal cues even in the presence of noise. An additional enhancement of these cues therefore doesn't always significantly improve speech perception in these individuals.

In individuals with CHL there was a significant difference seen in consonant identification scores only at 0 dB SNR across companding- enhanced signal conditions with better scores obtained in consonant enhanced condition. These results are an indication of higher benefit from consonant enhancement technique than companding. They are in agreement with previously existing literature which has shown a significant improvement in speech identification scores with envelope enhanced signal, in the presence of a competing signal (Apoux, Tribut, Debruille & Lorenzi, 2004; Baer et al., 1993; Bunnel, 1990; Clarkson & Bahgat, 1991; Franck, et al, 1999; Lyzenga, Festen & Houtgast, 2002). The benefit from consonant enhancement technique being more than companding can be attributed to the spectral enhancement provided by consonant enhancement. On the other hand, companding improves both spectral and temporal aspects of the signal. It has been reported in the literature that widened auditory filters mainly cause deterioration of spectral cues to an extent dependent on the amount of cochlear damage (Baer & Moore, 1993). Hence, a strategy, like consonant enhancement, that would compensate for this by making the spectral peaks and contrasts more available would benefit these individuals (Bunnell, 1990; Summerfield et al., 1985; Stone & Moore, 1992). When a strategy like companding is used, since the processing of signal is more complex with a series of compression and expansion, the process could have altered the spectral and temporal cues more than required for these individuals. Hence, it could have added distortion to the signal for individuals with CHL. Due to the above reasons; there was a significant difference in scores at 0 dB SNR between companding- consonant enhanced signal conditions with better scores in consonant enhanced condition.

Sequential Information Transfer Analysis (SINFA)

The information transmitted in both the subgroups of CHL (CHL with flat and sloping configuration) was maximum for manner of articulation (MOA) followed by voicing and place of articulation (POA) which were equally transmitted. There was no benefit seen with processed signal for the information transmitted with respect to voicing and POA. Individuals with flat hearing loss showed benefit for MOA cues with consonant enhancement condition at 0 dB SNR. These results are discussed under the following subheadings.

Place of articulation (POA)

The major cues for POA are formant transition (<50 ms) and spectrum of burst (Liberman, Delattre & Cooper, 1952). Paycho-acoustical studies have consistently demonstrated that individuals with CHL have significant difficulty in following change in frequency (formant transition) (Buss, Hall & Grose, 2004). As formant transition cues were unavailable, spectrum of burst could have help in extracting POA cues. However, the reason for loss of burst spectrum cue could be upward spread of masking of the burst spectrum by either competing signal (Nabelek, Letowski & Tucker, 1989) or backward masking by the vowel that followed these consonants. This was because vowels are higher in energy as compared to consonants (Fletcher, 1953).

In the present study, consonants were enhanced using two signal processing strategies, namely, consonant enhancement and companding. Both companding and consonant enhancement and companding did not bring any improvement in POA across all SNRs. The probable reason for not seeing an improvement in companding could be because, majority of the participants who took part in the present study had mild to moderate degree of hearing loss. Hence, frequency resolution could not have been largely affected (Dubno, et al, 1987). Also, companding enhanced only spectral contrast which might not have been useful for these participants.

The consonant enhancement strategy improved the burst amplitude by 6 dB in the present study. Enhancing the specific consonantal region in the consonants by 6 dB also did not show any benefit. The probable reason could be that the amount of enhancement provided might not have been sufficient. Another possible reason could be these participants were largely dependent only on frequency transition for extracting place cue. Hence, enhancing burst region did not show benefit. Therefore, although the signal was enhanced using consonant enhancement and companding, this enhancement was not perceived for POA.

Voicing

The major cues for voicing are voicing bars, which are low in intensity. Also, its spectral concentration is at low frequency (Lisker, 1977). The probable reasons for difficulty in perceiving the voicing bar for individuals with cochlear hearing loss are, poor frequency selectivity, inability to perceive low amplitude of voicing bars and, either upward spread of masking or backward masking as discussed for POA cues.

Enhancing the signal using consonant enhancement or companding did not primarily improve the voicing bars. This is because the strategy mainly aimed at improving the spectral contrast by increasing the burst and transition amplitude. However, even if the strategies enhanced the voicing bars, a simultaneous enhancement in the competing signal could have easily masked this low frequency voicing cue. Therefore, there was no improvement seen in the information transmitted for voicing.

Manner of articulation (MOA)

MOA cues are predominantly duration cues like, duration of burst or frication which is least for stops, and maximum for fricatives with affricates having an in between value. Results of the study conducted by Buss et al. (2004) indicated no correlation between amplitude modulation (AM) discrimination and speech perception. They suggested that, a gross temporal feature of the stimulus envelope served as a cue to discrimination of AM rate. The extraction of envelope cues being relatively unharmed in individuals with CHL was supported by Rosen (1992). Hence, due to the above discussed reasons, MOA cues were maximally transmitted.

As MOA cues were easily perceived, a further enhancement with consonant enhancement strategy probably retained the advantage of better transmission of MOA cues. However, when companding was used, the signal was modified in terms of both spectral and temporal features. This could have caused loss of naturalness for MOA cues which are more duration based. Hence, information transmitted was more with consonant enhancement strategy.

This benefit was more pronounced in individuals with flat configuration (N=11) and not in sloping loss (N=5). It can be hypothesised that higher information transmitted for MOA could have been present in sloping loss as well. However, owing to variability across individuals with cochlear hearing loss and less number of subjects with sloping loss considered in the present study, the effect could have been more evident in individuals with flat hearing loss. This was also supported by Dubno, et al, (1987) who did not show a difference in speech perception between flat and sloping configuration unless the subjects considered had hearing loss with configuration of greater than or equal to steeply sloping.

Conclusions

From the above findings, it can be concluded that speech perception deteriorates with an increase in noise in both normal listeners had individuals with CHL. Effect of noise is more for individuals with CHL than normal hearing listeners. In individuals with cochlear hearing loss, consonant enhancement might prove beneficial in noisy situations, although, the amount of improvement in speech perception could be minimal. Manner of articulation cue is the least affected parameter in both groups in both quiet and noise.

Clinical implications

The results have brought to notice that consonant enhancement strategy in individuals with CHL has the potential to improve speech perception in adverse listening conditions. Hence, this can be used as a rehabilitation technique. However, further research may be carried out for its successful implementation in amplification devices.

References

1. American National Standards Institute (1991), *Maximum permissible* ambient *noise levels* for audiometric test rooms. *ANSI S3.1-1991*. New York: American National Standards Institute.
2. Apoux, F., Tribut, N., Debruille, X., & Lorenzi, C. (2004). Identification of envelope-expanded sentences in normal-hearing and hearing-impaired listeners. *Hearing research*, *189*(1), 13-24.
3. Baer, T., & Moore, B. C. (1993). Effects of spectral smearing on the intelligibility of sentences in noise. *The Journal of the Acoustical Society of America*, *94*(3), 1229-1241.
4. Baer, T., Moore, B. C., & Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality, and response times. *Journal of rehabilitation research and development*, *30*, 49-49.
5. Beattie, R. C., Barr, T., & Roup, C. (1997). Normal and hearing-impaired word recognition scores for monosyllabic words in quiet and noise. *British journal of audiology*, *31*(3), 153-164.
6. Bhattacharya, A., & Zeng, F. G (2007). Companding to improve cochlear-implant speech recognition in speech-shaped noisea). *The Journal of the Acoustical Society of America*, *122*(2), 1079-1089.
7. Bunnell, H. T. (1990). On enhancement of spectral contrast in speech for hearing-impaired listeners. *Journal of the Acoustical Society of America, 88,* 2546-2556.
8. Clarkson, P. M., & Bahgat, S. F. (1991). Envelope expansion methods for speech enhancement. *The Journal of the Acoustical Society of America*, *89*(3), 1378-1382.
9. Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, speech, and hearing services in schools*, *31*(4), 362-370.
10. Deepthi, M. (2012). *Perception of Spectrally Enhanced Speech through Companding in Individuals with Cochlear Hearing Loss*, unpublished masters dissertation, University of Mysore, Mysore.
11. Dubno, J. R., & Schaefer, A. B. (1995). Frequency selectivity and consonant recognition for hearing impaired and normal hearing listeners with equivalent masked thresholds. *The Journal of the Acoustical Society of America*, *97*(2), 1165-1174.

12. Dubno, J. R., Dirks, D. D., & Schaefer, A. B. (1987). Effects of hearing loss on utilization of short duration spectral cues in stop consonant recognition. *The Journal of the Acoustical Society of America*, *81*(6), 1940-1947.

13. Eisenberg, L. S., Dirks, D. D., & Bell, T. S. (1995). Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *Journal of Speech, Language, and Hearing Research*, *38*(1), 222-233.

14. Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, *88*(4), 1725-1736.

15. Franck, B. A., Van Kreveld-Bos, C. S., Dreschler, W. A., & Verschure, H. (1999). Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. *Journal of the Acoustical Society of America, 106,* 1452-1464.

16. Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, *79*(4), 1020-1033.

17. Glasberg, B. R., & Moore, B. C. J. (1989). Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. *Scandinavian Audiology. 32,* 1-25.

18. Guelke, R. W., (1987). Consonant burst enhancement: A possible means to improve intelligibility for the hard of hearing. *Journal of Rehabilitation Research and Development, 24,* 217-220.

19. Hazan, V., Simpson, A., & Huckvale, M. (1998). Enhancement techniques to improve the intelligibility of consonants in noise: speaker and listener effects. In *International Conference on Spoken Language Processing.*

20. Hornsby, B. W., Johnson, E. E., & Picou, E. (2011). Effects of degree and configuration of hearing loss on the contribution of high-and low-frequency speech information to bilateral speech understanding. *Ear and hearing*, *32*(5), 543.

21. Jain, C., Konadath, S., Vimal, B. M., & Suresh, V. (2014). Influence of native and non-native multitalker babble on speech recognition in noise. *Audiology Research*, *4*(1), 9-13.

22. Joris, P. X., & Yin, T. C. (1992). Responses to amplitude modulated tones in the auditory nerve of the cat. *The Journal of the Acoustical Society of America*,*91*(1), 215-232.

23. Kumar, U. A., & Vanaja, C. S. (2004). Functioning of olivocochlear bundle and speech perception in noise. *Ear and hearing*, *25*(2), 142-146.

24. Loizou, P. C., Dorman, M., & Tu, Z. (1999). On the number of channels needed to understand speech. *The Journal of the Acoustical Society of America*,*106*(4), 2097-2103.

25. Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences*, *103*(49), 18866-18869.

26. Lyzenga, J., Festen, J. M., & Houtgast, T. (2002). A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity. *The Journal of the Acoustical Society of America*, *112*(3), 1145-1157.

27. Moore, B. C. J. (1995). Perceptual Consequences of Cochlear Hearing Loss and their Implications for the Design of Hearing Aids. *Ear & Hearing, 17(2),* 133-161.

28. Nábìlek, A. K., Letowski, T. R., & Tucker, F. M. (1989). Reverberant overlap and self masking in consonant identification. *The Journal of the Acoustical Society of America*, *86*(4), 1259-1265.

29. Narne, V. K., Barman, A., Deepthi, M., & Shachi (2014). Effect of companding on speech recognition in quiet and noise for listeners with ANSD. *International Journal of Audiology*, *53*(2), 94-100.

30. Oxenham, A. J., Simonson, A. M., Turicchia, L., & Sarpeshkar, R. (2007). Evaluation of companding-based spectral enhancement using simulated cochlear-implant processing. *The Journal of the Acoustical Society of America*,*121*(3), 1709-1716.

31. Pekkarinen, E., Salmivalli, A., & Suonpää, J. (1990). Effect of noise on word discrimination by subjects with impaired hearing, compared with those with normal hearing. *Scandinavian audiology*, *19*(1), 31-36.

32. Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold for the hearing impaired. *Journal of Speech and Hearing Research, 29,* 146-55.

33. Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of neurophysiology*, *30*(4), 769-793.

34. Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *336*(1278), 367-373.

35. Schorn, K., & Zwicker, E. (1990). Frequency selectivity and temporal resolution in patients with various inner ear disorders. *International Journal of Audiology*,*29*(1), 8-20.

36. Stephens, M. M., Rintelmann, W. F., (1978). The Influence of Audiometric Configuration on Pure-Tone, Warble-Tone and Narrow Band Noise Thresholds of Adults with Sensorineural Hearing Loss. *Journal of American Audiological Society, 3,* 221-226.

37. Stone, M. A., & Moore B. C. J. (1992). Spectral feature enhancement for people with sensorineural hearing impairment: effects on speech intelligibility and quality. *Journal of Rehabilitation Research and*

Development, 29 (2), 39-56.

38. Summerfield, Q., Foster, J., Tyler, R., & Bailey, P. J. (1985). Influences of formant bandwidth and auditory frequency selectivity on identification of place of articulation in stop consonants. *Speech Communication*, *4*(1), 213-229.

39. Thibodeau, L. M., & Van Tasell, D. J. (1987). Tone detection and synthetic speech discrimination in band reject noise by hearing impaired listeners. *The Journal of the Acoustical Society of America*, *82*(3), 864-873.

40. Turicchia, L., & Sarpeshkar, R. (2005). A bio-inspired companding strategy for spectral enhancement. *Speech and Audio Processing, IEEE Transactions on*,*13*(2), 243-253.

41. Tyler, R. S., Wood, E. J., & Fernandes, M. A. (1982). Frequency resolution and hearing loss. *British Journal of Audiology, 16,* 45-63.