Perception of Spectrally Enhanced Speech through Companding in Individuals with Auditory Dys-synchrony

[1]Shachi Vasishta,[2]Animesh Barman

Abstract

The study was taken up with the aim to know whether spectrally enhanced speech through compounding could improve speech perception of individuals with Auditory dys-synchrony in quiet and at different signal to noise ratios. To achieve the objective, 15 ears of fifteen individuals with auditory dys-synchrony and 15 ears of fifteen normal hearing individuals (mean age 27 years) were taken. Speech identification abilities of these individuals were assessed in quiet and at three different SNRs (+15 dB, +10 dB, +0 dB) for VCV syllables with and without companding. The minimum SNR at which these individuals correctly identified 50% of the words in a sentence was also assessed using Quick SIN sentences test in Kannada with and without companding. Results showed that as the SNR increased, the number of syllables correctly identified increased for VCV syllables in both with and without companding conditions , in both the groups. Normal hearing individuals showed significantly better performance for companded stimuli compared to non companded VCV syllables at + 0 dB SNR. Individuals with auditory dys-synchrony demonstrated this trend in quiet. Also, they required significantly lower SNR (better) for companded sentences than non companded sentences to correctly identify 50% words in a sentence. For companded sentences, SNR required to correctly identify 50% of the words in a sentence was lower than that of the non companded sentences, in both the groups. However, the normal hearing individuals performed well at a significantly lower (better) SNR than the individuals with auditory dys-synchrony for both companded and non companded test materials.

Keywords: *Companding, Auditory Dys-synchrony, Spectral enhancement, speech perception*

Introduction

Auditory dys-synchrony (AD) is a distinct hearing disorder characterized by auditory nerve dysfunction in the presence of normal outer hair cell activity (Starr, Picton, Sininger, Hood & Berlin, 1996). Desynchronized discharges at the level of 8th nerve and brainstem have been proposed as one of the underlying pathophysiologic mechanisms (Zeng, Oba, Garde, Sininger & Starr 1999; Kraus et al., 2000; Kumar & Jayaram, 2005). Psychophysical studies indicated that the consequences of disrupted auditory nerve activity are reflected as a significant impairment in temporal processing and difficulty in understanding speech that is disproptionate to the degree of hearing loss measured by pure tone thresholds (Siniger, & Oba, 2001; Zeng et al., 1999).

Difficulty in understanding speech, particularly in noise, is found to be a consistent problem reported by individuals with AD. Studies have investigated speech perception in noise in individuals with AD and illustrate that the noise has more detrimental effect on speech perception than that observed for listeners with normal hearing and those with cochlear hearing loss (Rance et al., 2007; Zeng & Liu, 2006).

The psychoacoustical studies have demonstrated impaired frequency discrimination for low frequency sounds (500 Hz) and with near normal values at 4000 Hz (Zeng et al. 1999; Rance, Mckay & Grayden 2004). Consistent with these findings, it has been observed that individuals with AD show good identification for phonemes that lie in the high frequency range than those phonemes that lie in the low frequency range (Rance & Barker, 2008; Narne & Vanaja, 2008).

Individuals with AD show severely affected temporal processing abilities that seem to be the basis of their poor speech perception (Zeng et al., 1999). Psychophysical measures showed that the disrupted neural activity significantly impairs timing related perception, such as pitch discrimination at low frequencies, temporal integration, gap detection, temporal modulation detection, forward and backward masking, binaural beats, signal detection in noise, and sound localization using interaural time differences (Zeng, Kong, Michalewski & Starr, 2005). Listeners with AD typically required silent periods of 20 ms or more to detect a gap compared to less than 5 ms in normal listeners (Rance et al, 2008). Also, speech signals are often degraded by noise, in real life situations. While normal-hearing listeners are capable of extracting the critical information from noisy speech, this ability is affected in individuals with AD.

Management of AD continues to be difficult and challenging. Persons with cochlear hearing loss derive significant benefit from hearing aids which employ nonlinear compression circuits. All these hearing aids assume abnormal functioning of outer hair cells (Berlin, Hood, Hurley & Wen 1996). Hence, these aids are of not much use for individuals with auditory dys-synchrony who

[1]Email: shachi.vasishtha@gmail.com.
[2]Reader in Audiology, Email: nishiprerna@yahoo.com

have normal outer hair cell functioning as supported by Rance, Cone-Wesson, Wunderlich and Dowell, (2002). Several other management strategies which can be used with individuals with auditory dys-synchrony include FM systems, cochlear implants, perceptual training, speech reading and cued speech (Kraus, 2001). Many studies have emphasized cochlear implants as the treatment of choice for AN (Rance & Barker, 2008). However, the invasive nature of cochlear implants and their doubtful efficacy points to a need for research on alternative strategies to improve speech intelligibility. This is particularly true for individuals with mild AD (Zeng & Liu, 2006).

Turicchia and Sarpeshkar (2005) have proposed a novel spectral enhancement scheme, companding, which combines two-tone suppression and dynamic gain control to increase the spectral contrast. Studies have shown that companding is also present along the auditory pathway. Both cochlea and the cochlear nucleus perform logarithmic compression on the input signals, while the brain performs exponential expansion (Zeng & Shannon, 1999). Considering this hypothesis, a signal processing strategy has been developed where certain signal will be compressed and certain frequencies will be enhanced which is termed as companding. Implementing the companding strategy, Bhattacharya and Zeng (2007) showed significant improvement in both phoneme and sentence perception in noise, in the cochlear implants users. However, the usefulness of this option has not been investigated in individuals with AD.

It is thus essential to study if listeners with AD can benefit from spectral enhancement of speech through companding. A systematic comparison with normal hearing listeners at various SNRs will be more appropriate to determine if the effect is level dependent or not. Keeping all this in mind the present study aimed to a) know whether companding of speech stimuli helps to improve speech intelligibility in individuals with normal hearing and those with AD, b) identify benefit of spectrally enhanced speech through companding at different SNRs in both the groups and c) determine how individuals with normal differ from those with AD in their performance for companded stimuli at quiet and at different SNRs.

Method

Participants

The participants were divided into clinical group consisting of those with AD and control group consisting of individuals with normal hearing.

Clinical group: consisted of 15 ears from 10 participants, in the age range of 15 to 42 years (mean age 27 years), fulfilling the criteria of AD in both ears. The demographic data and audiological test findings of 10 participants are given in Table 1. All the participants in the clinical group had degree of hearing loss ranging from mild to moderately severe sensorineural hearing loss. They had acquired, post lingual hearing loss and had disproportionately poor speech identification scores in relation to their pure tone threshold or poor SIS at 0 dB SNR. Further, they had absent auditory brainstem responses beyond that can be expected from thei degree of hearing loss, but had present otoacoustic emissions indicating normal OHC function. They had type 'A' tympanogram with no ipsi or contralateral reflexes.

Table 1: The Audiological test findings in individuals with AD

Participants	Age/ Gender	Ear	Severity of Hearing loss	SIS in quiet	SIS at 0 dB SNR	Pure tone configuration
AD1	42/ F	Right	Mild	34%	0%	Reverse slope
		Left	Mild	32%	0%	Reverse slope
AD2	32/M	Right	Moderately severe	88%	0%	Reverse slope
		Left	Moderately Severe	60%	0%	Reverse slope
AD3	25/F	Right	Minimal	92%	24%	Flat
		Left	Minimal	86%	28%	Flat
AD4	15/F	Right	Mild	84%	36%	Sloping
		Left	Mild	88%	44%	Flat
AD5	19/F	Right	Mild	88%	32%	Reverse slope
		Left	Mild	88%	28%	Flat
AD6	29/M	Right	Mild	76%	25%	Trough shape
		Left	Mild	76%	25%	Trough shape
AD7	26/M	Right	Mild	40%	0%	Reverse slope
		Left	Minimal	36%	0%	Reverse slope
AD8	27/ F	Right	Moderate	68%	0%	Reverse slope
		Left	Mild	92%	0%	Reverse slope
AD9	38/M	Right	Mild	88%	32%	Flat
		Left	Mild	100%	36%	Flat
AD10	19/F	Right	Moderately Severe	40%	0%	Flat
		Left	Moderately Severe	44%	0%	Flat

No other neurological symptoms were reported by these participants. Peripheral neuropathy or space-occupying lesion was ruled out by a neurologist. Any other otological disorder including middle ear infections were ruled out by an Otologist. All participants were fluent in Kannada and had no speech and language problems.

Test Environment

All the tests were carried out in a sound treated room. Noise levels in the test room were within permissible limits as per ANSI S3.1-1991.

Test equipment

A calibrated double channel diagnostic audiometer GSI-61 with TDH-50 P headphones were used for pure tone and speech audiometry. A calibrated immittance meter GSI tympstar was used to confirm the normal middle ear function through tympanometry and acoustic reflex measurement. Intelligent Hearing system Evoked potential instrument was used to record Cochlear microphonics (CM) and ABR. A calibrated ILO V6 instrument was used to measure DPOAEs.

A PC with Matlab version (2009) and Adobe Audition version 3 software was used for companding the speech stimuli. The speech stimuli, method used for companding and the procedure used to obtain SIS in different conditions are discussed below.

Test stimuli

In the present study two types of stimuli were used; Sentences and VCV nonsense syllables.

Sentences: Two lists of sentences were taken from quick SIN sentence test in Kannada developed by Methi, Avinash and Kumar (2009). Each list contains 7 sentences and each sentence has 5 key words. The sentences were spoken by a male native speaker of Kannada and was digitally recorded in an acoustically treated room using a unidirectional microphone kept at a distance of 10 cm from the speaker's mouth. Adobe audition software (version 3) was used to record the stimuli. The recorded sentences were normalized so that all the words in a sentence had equal intensity.

Speech shaped noise was used to generate sentences with different SNRs. Speech shaped noise was used as it was made to have the same long term average spectrum as sentences had. It was generated from the whole set of sentences at a sampling frequency of 44.1-kHz by estimating the long-term power spectrum of recorded test sentences. This was done by randomizing the phase of the Fourier spectrum of concatenated words of original signals using MATLAB (The Math Works, Natick, MA, USA) software (version 2009). It had a spectrum which approximates the long term average spectrum of the target sentences spoken by an adult male with a secondary peak presented around 100 Hz. Different SNRs were generated using MATLAB. In each list, first sentence were recorded without noise, second sentence was recorded at +15 dB SNR, third sentence at+10 dB SNR, fourth sentence at +5 dB SNR, fifth sentence at 0 dB SNR, sixth sentence at -5 dB SNR, and last sentence was recorded at -10 dB SNR. This was done as the sentences were used to obtain minimum SNR at which 50% word correctly identified in a sentence. The rms level of all these noises was adjusted according to the level of the target speech. The noise and speech was added prior to companding process.

VCV nonsense syllables: Twenty Vowel-Consonant-Vowel syllables (VCV) comprising of the kannada consonants /k, g, ch, t, d, th, dh, n, p, b, m, j, r, v, s, sh, y, h, l, l./ in the context of the vowel /a/ were used. These syllables were spoken by a female native speaker of Kannada and digitally recorded. The data acquisition system had a sampling frequency of 44.1 kHz and 32 bit analogue-to-digital converter.

The 20 VCV syllables were randomized to form 4 lists. Speech shaped noise was used to generate VCV syllables with different SNRs. List 1 was recorded without any noise. List 2 was recorded at +15 dB SNR. List 3 was recorded at +10 dB SNR. List 4 was recorded at 0 dB SNR.

The intelligibility of these recorded stimuli was established by obtaining a speech intelligibility rating on 10 normal hearing young Kannada speaking adults. Only those recorded stimuli judged as having good intelligibility were considered for the study.

These sentences and VCV syllables were spectrally enhanced using companding both in quiet and at different SNRs. The companding architecture was implemented in MATLAB. These companded test stimuli were named as modified sentences and VCV syllables. The companding was done after the mixing of speech noise and speech stimuli because, if this is found to be beneficial for individuals with AD, then the same can be suggested to be incorporated in amplification devices. The devices with this type of strategy incorporated would process noise and speech together as it would receive both simultaneously.

The strategy used a non-coupled filter bank and compression-expansion blocks as shown in Figure. 1. Every channel in the companding architecture had a relatively broad prefilter, a compression block, a relatively narrowband postfilter, and an expansion block. The prefilter and postfilter in each channel had the same center frequency. The pre and postfilter banks had logarithmically spaced center frequencies that span the desired spectral range.

First, the incoming signal was divided into a number of

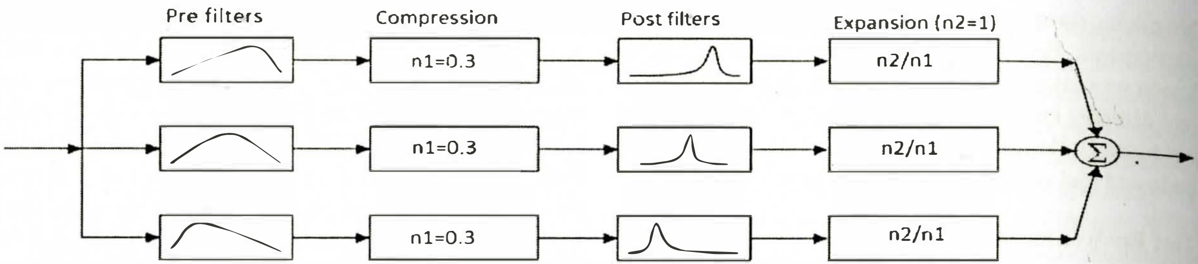

Figure 1: Block diagram of the companding architecture, showing the stimulus analyzed in a bank of broad band prefilters. The output of each prefilter was then subjected to compression, and output was filtered again using sharper postfilters, before it was expanded. The outputs from each channel are then summed to produce the processed broadband companded stimulus.

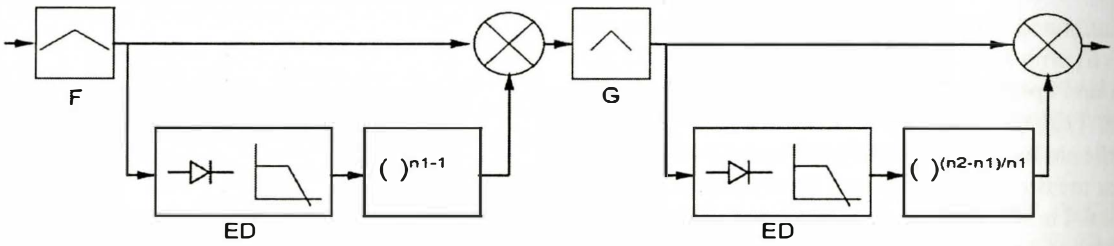

Figure 2: Detailed architecture of a single channel envelope detector.

frequency channels by a bank of relatively broad band-pass filters *F*. Figure. 2 shows the detailed architecture of a single channel companding pathway. The signal within each channel was then subjected to amplitude compression. The extent of compression was dependent on the output of the envelope detector, (ED), and the compression index, (n 1). The compression index $n1$ had a value of 0.3. The compressed signal was then passed through a relatively narrow bandpass filter *G* before being expanded. The gain of the expansion block depends on the corresponding ED output and the ratio of $n\ 2\ /n1$. The $n2$ parameter of the algorithm is expansion index and had a value of 1. The outputs from all the channels were summed to obtain the processed signal. The Adobe audition software (Version 3) was used to normalize the test stimuli to a level of -15dB. After implementing the companding process, there were altogether 8 VCV syllables lists (4 with & 4 without companding VCV syllables list) and 4 sentence test lists (2 with & 2 without companding process). These sentence lists and the VCV syllables lists were transferred digitally to a recordable compact disc. A calibration tone of 1 kHz with a level matched to the normalized level of the stimuli, was recorded prior to each list, using the 1-kHz calibration tone, VU meter on the audiometer was adjusted to read '0'.

Test Procedure

Speech recognition experiments were conducted in normal hearing listeners and individuals with Auditory Dys-synchrony. The test stimuli were played manually by a PC and were routed to a calibrated diagnostic audiometer (GSI-61) which was presented to the participants monaurally at their most comfortable level through the TDH 50 headphones. In the sentence recognition tests and the VCV syllable recognition tests, the participants were presented with a target sentence and the VCV syllables respectively. They were told that, they would first hear the stimulus without any background noise and then with a noise background. They were instructed to repeat or write the stimuli, after the stimuli were heard.

All the participants with normal hearing had to take a pretest consisting of syllables and sentences (with and without companding) at 0dB SNR and participants who scored 60% and above for sentences and 40 % or above for the VCV syllables were taken as participants in the study.

All VCV syllables lists (companded and without companding) at different SNRs were presented and numbers of syllable correctly identified were obtained for both the clinical as well as the control group. Without any modification, VCV syllables were presented in quiet and the number of syllables correctly identified was obtained in both the groups (clinical and control group). Modified (spectrally enhanced through companding) VCV syllables were presented in quiet and number of syllable correctly identified was obtained in both the groups. The procedure was repeated for different SNRs (+15, +10, +0 dB). For the sentence identification, the parameter was the minimum SNR that resulted in more than 50% of words being correctly identified (SNR-50).

Table 2: Mean and standard deviation (SD) of number of VCV syllables correctly identified (syllable identification scores) at 4 different SNRs (quiet, +15, +10, 0 dB), in two stimulus conditions (with and without companding) in the control group.

Stimulus	Mean number of syllables	SD
VCV at quiet without companding (VCVq)	20.00	0.00
VCV at quiet with companding(VCVqC)	20.00	0.00
VCV at +15 dB SNR without companding (VCV15)	19.53	0.74
VCV at +15 dB SNR with companding (VCV15C)	19.13	1.30
VCV at +10 dB SNR without companding (VCV10)	18.20	1.37
VCV at +10 dB SNR with companding (VCV10C)	18.00	1.51
VCV at 0 dB SNR without companding (VCV0)	10.47	2.13
VCV at +15 dB SNR with companding (VCV0C)	13.00	2.13

Out of the 2 companded sentence lists, one list was presented in quiet initially and number of word correctly identified was noted. If they had correctly identified at least 50% words, then the SNR was reduced to + 15 dB and a different set of sentences were presented once again. The same procedure was followed till they achieved the minimum SNR at which they could correctly identify 50% of the words in a sentence and that SNR value was noted. A similar procedure was used for the non companded list. Sentence lists were chosen randomly to minimize the list bias. The noise conditions in all the steps were presented in the order of increasing level of difficulty.

Results

There were basically two parameters of interest: a) the number of syllables correctly identified on presenting VCV syllables (companded and non companded) in quiet and noise in both the groups and b) the minimum SNR at which 50% of the words were correctly identified in a sentence (companded and non companded) in both the groups. Descriptive and inferential statistics were carried out over both these parameters in the control and the clinical groups which were then compared between the two groups.

Control group

A total of 15 ears with normal hearing sensitivity were included. The mean and standard deviation for the number of VCV syllables correctly identified across the two stimulus conditions (with and without companding) at four different SNRs were calculated and the details are given in the Table 2.

It is evident from Table 2 that the syllable identification improved as the SNR improved for both companded and non companded stimuli. The control group performance reached the ceiling in quiet. At +15 dB SNR and +10 dB SNR, non companded stimuli show slightly bet-

Table 3: Bonferroni adjusted multiple comparison test results for syllable identification scores obtained between any two SNRs (quiet, +15, +10, 0 dB) in the control group

(I) SNR	(J) SNR	Mean Difference (I-J)	Sig.
Quiet	+15dB SNR	0.667	0.14
	+10dB SNR	1.900	0.00
	+0dB SNR	8.267	0.00
+15dB SNR	+10dB SNR	1.233	0.01
	+0dB SNR	7.600	0.00
+10dB SNR	+0dB SNR	6.367	0.00

Table 4: t-value, degree of freedom and level of significance for pair wise comparison of syllable identification scores between stimulus conditions (with and without companding) and between different SNRs in the control group

Comparison	t-value	df	Sig
VCVq - VCV15	2.43	14	.029
VCVq - VC10	5.08	14	.000
VCVq - VCV0	17.31	14	.000
VCV15 - VCV10	4.40	14	.001
VCV15 - VCV0	15.18	14	.000
VCV0 - VCV0	12.31	14	.000
VCVqC - VCV15C	2.58	14	.022
VCVqC - VCV10C	5.12	14	.000
VCVqC - VCV0C	11.64	14	.000
VCV15C- VCV10C	3.01	14	.009
VCV15C - VCV0C	8.99	14	.000
VCV10C - VCV0C	8.79	14	.000
VCV15 - VCV15C	2.45	14	.028
VCV10 - VCV10C	.642	14	.531
VCV0 - VCV0C	5.429	14	.000

Table 5: Mean and standard deviation (SD) of minimum SNR at which in minimum 50% of the words correctly identified in sentence test materials, with and without companding in the control group

Conditions	Mean (SNR)	N	SD
Minimum SNR without companding	-2.67	15	2.58
Minimum SNR with companding	-5.00	15	0.00

Table 6: Mean and standard deviation (SD) of number of syllables correctly in the two conditions (with and without companding) in quiet and +15 dB SNR in the clinical group

	VCVq	VCVqC	VCV15	VCV15C
Mean(No. of syl)	10.67	12.47	3.33	4.33
SD	6.275	6.413	4.451	5.912

ter scores while at 0 dB SNR, companded stimuli appear to be better.

To know whether there was any significant difference in performance of the control group at different SNRs and in the two stimulus conditions (with and without companding), two way repeated measure ANOVA was done. A significant main effect on syllable identification was seen at different SNRs [$F(3, 42) = 143.72$, $p < 0.05$]. A significant main effect was also observed for companded and non companded conditions [$F(1, 14) = 8.83$, $p < 0.05$]. A significant interaction between SNRs and conditions [$F(3, 42) = 24.11$, $p < 0.05$] was present. Bonferroni's adjusted multiple comparison test was used as the post-hoc test and the details are shown in Table 3. The Bonferroni adjusted multiple comparison test revealed a significant difference in syllable identification between all SNR conditions except between quiet and +15 dB SNR condition. Paired sample t-test was used to test significant difference in syllable identification between the conditions and also between the SNRs. Details of the paired sample t- test is shown in Table 4. Paired sample t-test results showed indicated a significant difference in all the conditions expect at +10 dB SNR where the control group performed equally well in with and without companding.

The mean and standard deviation of the SNR-50 for sentences with and without companding in the control group are given in Table 5. It is evident from Table 5 that the SNR-50 for sentences was lower (better) for companded condition compared to non companding condition. To test for significance, a paired sample t-test was done. A significant difference was seen [$t(14) = 3.50$, $p < 0.05$] between the two conditions indicating that performance was better with companding than without companding in the control group.

Table 7: Z-value and level of significance for pairwise comparison of syllable identification between two stimulus conditions (with and without companding) and 2 different SNRs (quiet, +15 dB) obtained in the clinical group

	VCVqC -VCVq	VCV15C -VCV15	VCV15 -VCVq	VCV15C -VCVqC
Z-value	-3.09	-1.90	-3.19	-3.18
Sig.	0.00	0.06	0.00	0.00

Clinical group (Auditory Dys-synchrony)

A total of 15 ears (10 individuals) having auditory dys-synchrony comprised the clinical group. The clinical group participants could not identify any of the syllables at +10 dB SNR and 0 dB SNR. The mean and standard deviation for the syllable identification scores across the two stimulus conditions (with and without companding) in quiet and at +15 dB SNR are given in Table 6. The number of syllables identified was more for companded stimuli than the non companded stimuli, both in quiet as well as at +15 dB SNR. Due to a high standard deviation, non parametric tests were used for the clinical group to assess significance. To know whether the performance of the participants in the clinical group differed significantly across the two stimulus conditions and between the two SNRs, Wilcoxon Signed Rank Test was carried out. Table 7 details the test results. As evident from Table 7, Wilcoxon Signed Rank Test showed a significant difference in syllable identification in all the conditions except at +15 dB SNR, where the difference between companded and non companded syllable identification was not significant ($p=0.06$).

Table 8: Mean and standard deviation (SD) of minimum SNR at which 50% of the word correctly identified in sentence test materials, in two conditions (with and without companding) in the clinical group (n=13)

	Mean(SNR)	SD
Minimum SNR without companding	4.62	2.47
Minimum SNR with companding	0.77	3.44

For the sentence test material, only 13 ears were tested as data from two ears could not be obtained for the task. The mean and the standard deviation for this data are shown in Table 8. The data appear to indicate better performance in the companded condition. To test for significance, Wilcoxon Signed rank test was carried out. A significant difference was found between the two conditions [$Z = -2.89$, $p < 0.05$] indicating that the clini-

Table 9: Z- value and level of significance for comparison of VCV syllable identification scores between the clinical and control groups obtained at 2 SNRs (quiet, +15 dB) and in two conditions (with and without companding)

	VCVq	VCVqC	VCV15	VCV15C
Z- value	-4.99	-4.99	-4.83	-4.83
Sig.	0.000	0.000	0.000	0.000

Table 10: Z- value and level of significance for comparison of minimum SNR that resulted in 50% of the words correctly identified in sentence test materials, in two conditions (with and without companding) between the clinical and the control groups

	Minimum SNR without companding	Minimum SNR with companding
Z- value	-4.40	-4.73
Sig.	.00	.00

cal group indeed performed better with companded sentences when compared to non companded sentences.

Between Group Comparisons

Comparison of number of syllables correctly identified across the clinical and control group at different SNRs in two conditions (with and without companding): Mann Whitney test was used to compare between the two groups across the SNRs and between companded and non companded syllables. The details can be seen in Table 9. The results of Mann-Whitney Test showed significant difference in performance between control and clinical group for syllable identification at both the SNRs and also in between the two conditions (with and without companding). Thus, the control group performed significantly better than clinical group in VCV identification task at all the SNRs in both stimulus conditions.

Comparison of the SNR-50 for sentences between the clinical and the control group with and without companding : Mann Whitney test (Table 10) revealed a significant difference between the two groups with and without companding. The control group, thus performed better at low SNRs than the clinical group in both companded and non companded conditions.

Discussion

Findings in Control group

In the present study, the number of syllables correctly identified significantly reduced as the SNR reduced. The number of syllables identified was the least for the non companded signal at 0 dB SNR and maximum number of syllables correctly identified was seen in the quiet condition.

The results obtained of the present study are similar to those obtained by earlier investigators (Dorman, Loizou & Tu, 1998). Houtgast and Steeneken (1985) also reported that speech intelligibility is reduced in the presence of background noise. This is partly because the noise reduces the modulations of speech envelope. In addition, the decline in intelligibility may also result from the distortion of temporal fine structure and introduction of spurious envelope modulation, as these modulations obscure or mask the modulation pattern of speech, and obliterate some of the cues for identification (Drullman, 1995).

The individuals in the control group demonstrate a significant benefit from companded VCV syllables at 0 dB SNR. This shows that companding enhances the spectral peaks and listeners could take advantage of these enhanced peaks in adverse listening conditions. The result of the present study is in agreement with the results of previous investigators where they showed significant improvement in identification scores in the presence of background noise, when envelope enhanced stimuli were presented to individuals with normal hearing and cochlear hearing loss (Apoux et al., 2004; Baer et al., 1993; Bunnell, 1990; Clarkson & Bahgat, 1991; Franck et al., 1999; Lyzenga e t al., 2002). Turicchia and Sarpeshkar (2005) also showed that spectral contrast is an emergent property of the companding strategy and had speculated that this strategy has the potential to improve speech performance in noise.

Studies have found small improvement in the identification of stop consonants in quiet, using a contrast enhancement technique in which the envelope amplitude of each Fast Fourier Transform bin is enhanced in proportion to the difference in the original envelope amplitude and the average spectrum level (Bunnell, 1990; Franck et al., 1999; Lyzenga *et al.*, 2002). Baer *et al.* (1993) convolved the spectrum with a difference of Gaussian filter to provide spectral enhancement. They showed that their normal hearing participants preferred speech in noise with moderate enhancement in terms of quality and intelligibility. In the quiet condition, the performance of normal hearing individuals had already reached the ceiling without companding. So, the benefits of companding could not be seen. In situations with good SNR, companding is not expected to provide additional benefit since the individuals with normal hearing do not have any kind of spectral and temporal deficit. But companding will be effective in adverse listening conditions, which is shown by a better performance at 0 dB SNR in the companded condition.

The results of the normal hearing individuals averaged across 15 listeners showed that for companded sentences the minimum SNR that is required to correctly identify 50% of the words is achieved at significantly lower SNR (better) than without companded sentences. This shows better performance at lower SNR for with companded stimuli compared to without companding sentence test material.

These findings are consistent to earlier investigator's results who reported average improvement in sentence and word recognition in noise, in normal hearing participants using companding strategy (Bhattachrya & Zeng; Oxenham et al., 2007). Normal hearing individuals, in the adverse listening conditions utilized the enhanced spectral and temporal contrast of the companded speech stimuli. The improvement observed in the present study for the companded stimuli can be attributed to the increased spectral and temporal contrast provided by companding strategy.

Findings in Auditory Dys-synchrony

Individuals with AD also showed the trend of reduced number of syllable identification with poorer SNR. Participants in the clinical group could only perform at two SNRs, that is in quiet and at +15 dB SNR. This suggests that listeners with AD have difficulty in utilizing available information if the condition is even slightly worse. Studies have investigated speech perception in noise in individuals with AD and have reported that noise has a detrimental effect on speech perception (Rance et al., 2007; Zeng & Liu, 2006). The results in the present study are in accordance with the previous studies (Rance et al., 2007; Zeng & Liu, 2006). Adding noise to the speech signal leads to problem in perceiving the envelope of speech, because of reduction in modulation depth and addition of spurious modulation (Drullman, 1995). This explanation would explicate severe degradation in speech intelligibility in the presence of background noise for individual with AD.

The exact mechanism causing extreme difficulty in understanding speech in the presence of noise in individuals with AD is unclear. Zeng et al. (2005) reported an excessive masking effect for the detection of tones in the presence of noise. This excessive masking may be one of the factors in these individuals contributing to the extreme difficulty in understanding speech in the presence of noise.

The results of the present study unequivocally demonstrated benefit from companding. Companded VCV syllables significantly improved syllable identification in quiet. Companding increases the spectral and temporal contrast in speech (Bhattacharya & Zeng; 2007, Oxenham et al., 2007). Loizou (2005) implemented companding strategy in CI users and found a modest improvement in vowel recognition. Narne and Vanaja (2008) have shown that enhancing the envelope improved consonant identification at quiet in individuals with AD. In addition, Zeng and Liu (2006) have said that the participants with AD showed improved performance, in quiet when clear speech is presented and this improvement is attributed to enhanced envelopes in the clear speech. Thus, it can be said in the present study that, the enhanced spectral and temporal contrast for the companded stimuli might be the reason for the improved speech perception in individuals with AD.

There was a difference in performance between with companding (mean=4.33) and without companding (mean=3.33) VCV syllables at + 15 dB SNR, but it was not statistically significant. Lack of significant difference at + 15 dB SNR suggests that, individuals with AD cannot utilize enhanced information in the presence of noise which indicates that they have more of neural problem, which predominantly exhibit temporal deficit.

The results of the individuals with AD averaged across 15 ears showed that for companded sentences the minimum SNR that is required to correctly identify 50% of the words is achieved at significantly lower SNR (better) than without companded sentences. Hassan (2011) reported that temporal modification of speech is beneficial for participants with AD. These results are consistent with earlier investigators who reported an average improvement in sentence and word recognition in noise in individuals with AD using enhanced envelope cues (Zeng & Liu, 2006). Companding enhances the spectral peaks and listeners are able to take advantage of these enhanced peaks in adverse listening condition (Bhattachrya & Zeng, 2007; Oxenham, Simonson, Turicchia & Sarpeshkar, 2007). Bhattachrya and Zeng (2007) reported that companding apart from improving spectral contrast also enhances the temporal contrast. As the individual with AD exhibit temporal deficits enhancing the spectral and temporal contrast might have lead to significant improvement.

Findings in Between Group Comparisons

Syllable identical scores in the presence of noise were significantly more affected in individuals with AD when compared to listeners with normal hearing. Studies have reported that normal hearing listeners use fine structure in understanding speech in adverse listening conditions (Zeng & Liu, 2006). However, individuals with AD are impaired in extracting both envelope and fine structure cues from speech signal even in quiet, adding noise to the speech signal may exaggerate their problem in perceiving the envelope of speech, because of reduction in modulation depth and addition of spurious modulation (Drullman, 1995). In the present study, probably the impaired ability to extract the envelope and fine structure cue could be the reason for their poorer performance compared to normal hearing individuals.

Another possible reason for poorer performance in presence of noise in individuals with AD compared to normal hearing individuals might be due to the excessive masking effect in individuals with AD (Zeng et al., 2005). They could not utilize any information in the presence of noise which indicates they have more of neural problem, which predominately exhibits temporal difficulty.

There was improvement in both the group for companded VCV syllables perception, but the performance of the normal group was significantly better compared to clinical group for syllable identification at all the SNRs. The results suggest that individuals with auditory dys-synchrony were unable to fully utilize the temporal and spectral cues. Potential reasons behind this can be attributed to the poor temporal and spectral processing abilities (Rance et al., 2004; Zeng et al., 2005).

Secondly, spectral enhancement provided by the companding was across frequencies, including low frequencies. The low frequencies instead of enhancing speech perception might have caused upward spread of masking, thus the benefit which individuals with AD got from high frequency enhancement also reduced causing minimal improvement in their SIS compared to normal hearing individuals. Improvement in both the group, for the minimum SNR that is required to correctly identify 50% of the words in a sentence for the companded sentence test stimuli compared to without companded stimuli was observed, but the performance of the normal group was significantly better compared to clinical group. Apoux et al. (2004) have clearly shown that envelope enhancement enhances the consonant portion and comprises the vowel portion of the signal and improves the perception of speech in noise better than other signal processing strategies in individuals with normal hearing. In addition, Picheny, Durlach and Braida (1985, 1986) have said that advantage of clear speech over conversational speech in noise for cochlear hearing loss listeners may be due to the increased consonant to vowel ratio and enhanced envelopes. Thus, it can be said that the improvement observed in the present study may probably be due to enhanced spectral and temporal contrast.

A possible reason for the poorer performance of individuals with AD compared to normal hearing individuals is their inability to fully utilize the temporal and spectral cues. In AD, the neural temporal processing is disrupted, which affects the listener's ability to cope with the dynamic nature of speech signal. Severe disruption of timing cues could impair not only the ability to use amplitude envelope cues in speech but also to perceive rapidly changing spectral shapes in the flow of speech stimuli (Rance et al., 2004).

To conclude, speech perception is significantly impaired in individuals with AD. This is probably a reflection of excessive masking and diminished temporal processing abilities. Enhancing the spectral and temporal contrast through companding may improve speech perception in these individuals. Utilizing the companding strategy in hearing aids may provide benefit in individuals with AD.

Conclusion

Our behavioral data suggest that companding the speech signal enhances the spectral and temporal contrast and lead to better speech perception in individuals with auditory dys-synchrony. There are as many studies which have reported that hearing aids have failed to show beneficial effects in participants with auditory dys-strategies improve speech perception in auditory dys-synchrony. The present study provides indirect evidence that hearing aids incorporating companding strategy can enhance spectral and temporal contrast and be beneficial to persons with auditory dys-synchrony. Thus, it is suggested that companding of signal can be used as one of the signal processing strategies improve speech perception in auditory dys-synchrony.

References

American National Standards Institute. (1991). *Maximum Permissible Ambient Noise Levels for Audiometric Tests Rooms.* ANSI S3:1- (1991). New York: American National Standards Institute.

Apoux, F., Tribut, N., Debruille, X., et al. (2004). Identification of envelope expanded sentences in normal-hearing and hearing-impaired listeners. *Hearing Research, 189*, 13-24.

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,* 49-72.

Berlin, C. I., Hood, L. J., Hurley, A., & Wen, H. (1996). Hearing aids: only for hearing impaired patients with abnormal otoacoustic emissions. In C.I. Berlin (Ed.), *Hair cells and hearing aids.* (pp. 99-111). San Diego: Singular Publishing groups.

Bhattacharya, A., & Zeng, F. G. (2007). Companding to improve cochlear-implant speech recognition in speech-shaped noise. *Journal of the Acoustical Society of America, 122,* 1079-1089.

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.

Carhart, R., & Jerger, J.Â (1959).Preferred method for clinical determination of pure-tone thresholds.*Journal of Speech and Hearing Disorders, 24,* 330-345.

Clarkson, P., & Bahgat, F. (1991). Envelope expansion methods for speech enhancement. *Journal of the Acoustical Society of America, 89,* 1378-1382.

Dorman, M., Loizou, P., & Tu, Z. (1998). The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processor with 6-20 channels. *Journal of the Acoustical Society of America, 104*, 3583-3585.

Drullman, R. (1995). Speech intelligibility in noise: relative contribution of speech elements above and below the noise level. *Journal of the Acoustical Society of America, 98*, 1796-1798.

Franck, B. A. M., van Kreveld-Bos, C. S. G. M., Dreschler, W. A., & Verschuure, H. (1999). Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. *Journal of the Acoustical Society of America, 106*, 1452-1468.

Houtgast, T., & Steeneken, H. T. M. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibity in auditorium. *Journal of the Acoustical Society of America, 77*, 1069-1077.

Hassan, D. M. (2011). Perception of temporally modified speech in auditory neuropathy. *International Journal of Audiology, 50*, 41-49.

Kraus, N., Bradlow, A. R., Cheatham, M. A., Cunningham, J., King, C.D., & Koch, C.D. (2000). Consequences of neural asynchrony: A case of auditory neuropathy. *Journal of Association for Research in Otolaryngology, 1*, 33-45.

Kruase, J. C., & Braida, L. D. (2004). Acoustical properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America, 115*, 362-378.

Kumar, U. A., & Jayaram, M. (2005). Auditory processing in individuals with auditory neuropathy. *Behavioral and Brain Functions*,1-21. doi:10.1186/1174-9081-1-21.

Loizou, P. (2005). Evaluation of the companding and other strategies for noise reduction in cochlear implants. *Conference on Implantable Auditory Prosthesis*, Asilomar, Monterey, California.

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

Methi, R., Avinash, & Kumar, U. A. (2009). Development of sentence material for Quick Speech in Noise test (Quick SIN) in Kannada. *Journal of Indian speech and Hearing Association, 23*(1), 59-65.

Narne V. K. & Vanaja C. (2008). Speech identification and cortical potentials in individuals with auditory neuropathy. *Behavioral and Brain Function, 31*, 4 - 15.

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

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. *Journal of Speech Language and Hearing Research, 28*, 96-103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech Language and Hearing Research, 29*, 434-446.

Rance, G., Cone-Wesson, B., Wunderlich, J., & Dowell, R. (2002). Speech perception and cortical event related potentials in children with auditory neuropathy. *Ear and Hearing, 23*, 239-253.

Rance, G., McKay, C., & Grayden, D. (2004). Perceptual characterization of children with auditory neuropathy. *Ear and Hearing, 25*, 34-46.

Rance, G., Barker, E. J., Mok, M., Dowell, R., Ricon, A., & Garratt, R. (2007). Speech perception in noise for children with auditory neuropathy/ dys-synchrony type hearing loss. *Ear and Hearing, 28*, 351-360.

Rance, G., & Barker, E.J. (2008). Speech perception in children with auditory neuropathy/dyssynchrony managed with either hearing AIDS or cochlear implants. *Otology Neurotology, 29*, 179-182.

Rance, G., Fava,R., Baldock, H., Chong, A., Barker, E., Corben, L, & Delatycki, M. (2008). Speech perception in individuals with Fredeich ataxia. *Brain, 131*, 2002-2012.

Sininger, Y., & Oba, S. (2001). Patients with auditory neuropathy: Who are they and what can they hear? In Y. Sininger, & A. Starr (Eds.), *Auditory neuropathy: A new perspective on hearing disorder*, (pp. 15-36). Canada: Singular publishing group.

Starr, A., Picton, T.W., Sininger, Y., Hood, L., & Berlin, C.I. (1996). Auditory neuropathy. *Brain, 119*, 741-753.

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

Zeng, F. G., Oba, S., Garde,S., Sininger, Y., & Starr, A. (1999). Temporal and speech processing deficits in Auditory Neuropathy. *NeuroReport, 10*(16), 3429-3435.

Zeng, F. G., & Shannon, R. V. (1999). Psychophysical laws revealed by electric hearing. *NeuroReport, 10*, 1931-1935.

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology, 93*, 3050-3063.

Zeng, F. G., & Liu, S. (2006). Speech perception in auditory neuropathy participants. *Journal of Speech and Hearing Research, 42*(2), 367-380.