Relationship between Consonant Perception and Psychoacoustic Measures in Auditory Dys-Synchrony

[1]G. Nike Gnanateja & [2]Animesh Barman

Abstract

The speech perception deficit has been attributed mainly to the impaired supra-threshold temporal processing abilities in individuals with auditory dys-synchrony. Most studies on speech perception in auditory dys-synchrony focus on the speech identification scores and not the specific pattern of errors seen. Simulation studies and temporal modifications of speech, hint that there is some other aspect of sound signal processing which affects the speech perception in addition to the impaired temporal processing. The current study aimed at analyzing the frequency and temporal resolution abilities and their relation to perception of the place and manner of articulation and the voicing feature of speech in individuals with auditory dys-synchrony and the comparison of the results in individuals with auditory dys-synchrony with that of individuals with cochlear hearing loss and individuals with normal hearing sensitivity. The results of the study showed that the auditory filters in individuals with auditory dys-synchrony were broader than that of the normal hearing and cochlear hearing loss ears with age and audiometric configuration matched to the individuals with auditory dys-synchrony. The impairment was more so in the lower frequencies. The peak sensitivity and the bandwidth of the temporal modulation transfer function were reduced relative to the controls. The voicing errors were most prominent followed by errors in place and manner. The errors in manner and place correlated with the temporal resolution and the frequency resolution at 500 Hz and 1000 Hz. The voicing errors were correlated with the frequency resolution at 500 Hz and 1000 Hz.

Key words: *Speech perception, frequency resolution, temporal resolution, auditory dys-synchrony*

.Introduction

Auditory dys-synchrony is a hearing disorder characterized by normal cochlear amplifier and disordered afferent neural transmission (Starr, Picton, Sininger, Hood & Berlin, 1996; Zeng, Kong, Michalewski, & Starr, 2004). The most striking feature that defines auditory dys-synchrony is the absence of the action potentials in spite of normal oto-acoustic emissions or cochlear microphonics (Berlin et al., 2005). Histopathological studies in individuals with auditory dys-synchrony have shown that there is colossal reduction in the number of spiral ganglion cells and also nearly 30% loss in the number of outer hair cells (Starr et al., 1996; Varga et al., 2003).It is not clear as yet if the pathology in AN/AD lies at the level of the synapse of the inner hair cell and auditory nerve, the auditory nerve itself or at the level of the inner hair cell (Starr et al., 1996). Hence, it is now being referred to as Auditory neuropathy/dys-synchrony (AN/AD) in the recent studies (Knox, 2005; Kundu & Rout, 2010; Rance et al., 2010).

The hearing sensitivity in individuals with AN/AD ranges from normal hearing sensitivity to profound degree of sensori-neural hearing loss (Zeng, Oba, Garde, Sininger & Starr,1999; Starr, Sininger & Pratt, 2000; Kumar & Jayaram, 2005). The speech perception, however, are disproportionate to the degree and configuration of hearing loss (Zeng et al., 2004; Rance, 2005), which implies that the poor speech perception is not just because of the reducedaudibility in individuals with auditory dys-synchrony. The disproportionately poor speech perception scores have been attributed to the supra-threshold processing deficits as a consequence of disrupted nerve activity.

The prominently discussed deficit in AN/AD across studies is the temporal processing deficit as evidenced by the large gap detection thresholds and also the reduced peak sensitivity and bandwidth of the temporal modulation transfer function (Zeng et al., 1999; Zeng et al., 2004; Kumar & Jayaram, 2005; Rance, 2005). The poor speech perception abilities have been predominantly attributed to the temporal processing deficit (Zeng et al., 1999; Rance, Fava, Chong, Barker, Corben & Delatycki, 2008; Rance et al., 2010). The frequency resolution results on the other hand have been equivocal across different studies. The frequency difference limens in AN/AD have been reported to be poorer than that seen in normals and more so at the lower frequencies (Zeng et al., 2004; Rance, 2005; Barman, 2008). However, this is a measure of frequency discrimination and not frequency resolution, and Moore (1973) and Patterson and Moore (1986)

[1]E-mail: nikegnanateja@gmail.com; [2]Reader in Audiology, E-mail:nishiprerna@yahoo.com

reported no significant relationship between frequency discrimination and frequency resolution. Kraus et al., (2000) have shown that there is little change in threshold of the pure tone with change in notch width of a notched noise masker in individuals with AN/AD, which indicates at broadened auditory filters (Patterson & Moore, 1986). Result contrasting to this was obtained by Rance, McKay and Grayden (2004) where they found that the threshold shift from a no notched to notched condition in a broadband noise in individuals with AN/AD and the normal hearing controls were similar, whereas the individuals with cochlear hearing loss had smaller threshold shifts. Vinay and Moore (2007) measured the psycho-acoustical tuning curves in individuals with AN/AD and reported that the PTCs were normal in shape and Q10 dB values were lower than those reported by Kluk and Moore (2004). The major deficit seen in individuals with AN/AD is the deficit in temporal resolution with relatively spared frequency resolution compared to individuals with cochlear hearing loss who have a major deficit in frequency resolution and relatively spared temporal resolution (Rance et al., 2008).

The voiced/voiceless confusion is the predominant error followed by errors in the perception of place and manner of articulation in individuals with AN/AD (Rance et al., 2008; Narne & Vanaja, 2008; Rance et al., 2010). Rance et al. (2010) found that the errors in voicing perception were correlated with the reduced temporal resolution abilities and the high frequency consonant perception was correlated with the difference limen at high frequencies. Narne and Vanaja (2008) found that envelope enhancement improved perception of manner and place of articulation in individuals with AN/AD, however, the voicing perception showed no significant improvement and this was attributed to the poor low frequency processing abilities in this population which is necessary for the perception of the voicing bars which are very important for voicing perception.

Need for the study

Most of the studies on speech perception in individuals with auditory dys-synchrony have concentrated on analysing the correlation of percentage speech identification scores with the psychoacoustic test results. However, the scores as such do not give information on what phonetic features a person is able to or not able to perceive, and it does not give a correct picture about the speech perception abilities in individuals with auditory dys-synchrony. There are only a handful of studies (Narne & Vanaja, 2008; Rance et al., 2010) which have assessed the perception per se based on the phonetic features perceived in auditory dys-synchrony. However, psychoacoustic tests were not performed which could have helped to explain if the perception errors seen were because of a temporal or a spectral processing deficit.

Rance et al., (2010) analysed the open-set consonant perception in auditory dys-synchrony along with temporal modulation transfer function (TMTF) for modulation frequencies of 10, 50 and 150 Hz. However, the modulation frequencies below 10 Hz are the ones which significantly contribute to speech perception. Further, they also analysed the open set consonant perception results using information transfer analysis. But open set consonant identification task frequently gives rise to an irregular matrix and information transfer analysis results cannot be relied upon when performed on an irregular matrix (Wang & Bilger, 1973). To tap the ability to perceive the temporal envelope cues in speech it is necessary that the amplitude modulation detection ability be studied at more modulation frequencies within the range of 2 Hz to 50 Hz, as the temporal envelope of speech ranges from 2 Hz to 50 Hz (Rosen, 1992). Additionally data on modulation detection abilities in AN/AD suggests a relatively greater decrease in the modulation detection thresholds at frequencies lower than 10 Hz and greater than 64 Hz (Zeng et al., 2004; Kumar & Jayaram, 2005). Typical speech contains periodicity information in the range of 50 Hz to 500 Hz which suggests that individuals with AN/AD have difficulty processing the periodicity information too. These data suggest that, to analyse the temporal resolution in relation with speech perception in individuals with AN/AD, it is important to study the temporal resolution at more modulation frequencies between the range of 2 Hz to 50 Hz and also modulation frequencies above 50 Hz.

In addition to the above, there are only a handful of studies which have assessed the frequency resolution in auditory dys-synchrony. Vinay and Moore (2007) investigated frequency resolution using psychoacoustic tuning curves (PTCs) using a simultaneous narrow-band noise masking. But the simultaneous narrowband noise masking technique has its disadvantages i.e. the occurrence of beats and off-frequency listening. Glasberg and Moore (1986) have suggested the use of notched noise maskers to rule out the effects of beats and off-frequency listening. This suggests that further studies are needed to assess frequency resolution in individuals with AN/AD.

Hence, there is a need to study the closed set consonant identification and its relationship with the temporal resolution and frequency resolution measures obtained in a controlled and comprehensive manner. Also, these measures in AN/AD should be compared with those seen in individuals with cochlear hearing loss.

The basic aim of this study was to observe the relationship between the consonant perception and the frequency and temporal resolution measures in individuals with AN/AD. The study also aimed at comparing the speech perception and frequency and temporal resolution in individuals with AN/AD and individuals with cochlear hearing loss.

Method

Participants

A total of 25 participants were included in the study. They were divided into two clinical groups and one control group. All the participants were literate with mother tongue being Kannada and they could read the Kannada script without any difficulty.

Table 1: The Audiometric details of the participants with AN/AD

Ear	Age (years)	Hearing sensitivity	SIS quiet (%)	Audiometric pattern
AN1	27	Normal	70	Flat
AN2	27	Normal	70	Flat
AN3	19	Normal	64	Flat
AN4	19	Normal	68	Flat
AN6	34	Normal	44	Flat
AN5	34	Normal	40	Flat
AN7	18	Normal	20	Flat
AN8	18	Normal	16	Flat
AN9	21	Normal	80	Flat
AN10	21	Normal	80	Flat
AN11	13	Mild	36	Flat
AN12	13	Mild	36	Peaked
AN13	31	Mild	40	Peaked
AN14	31	Mild	36	Peaked
AN15	31	Mild	32	Flat
AN16	31	Mild	36	Flat
AN17	35	Mild	24	Flat
AN18	17	Mild	12	Flat

Note: SIS - percentage correct speech identification scores for phonemically balanced word list.

Clinical Group 1 - Auditory Dys-synchrony (AN/AD): Eighteen ears from eleven participants clinically diagnosed as having auditory dys-synchrony were included in Clinical Group I (AN/AD group). Participants with pure-tone thresholds ranging from normal to moderate sensorineural hearing loss were taken. The age-range of the participants was 13–34 years with a mean age of 24.4 years. None of the participants in this group had replicable auditory brainstem responses and acoustic stapedial reflexes, and all of them had normal click evoked oto-acoustic emissions. Additionally, all of them had disproportionately poor speech identification scores with respect to the audiometric thresholds or abnormal speech identification scores in the presence of noise in those with good speech identification in quiet. The audiometric details of the participants with AN/AD are shown in Table 1.

Clinical Group 2 - Cochlear hearing loss (CochHL): Seven ears from four participants clinically diagnosed as having cochlear hearing loss were taken up for the study. The participants in this group had audiometric configuration matched to the participants with auditory dys-synchrony who had hearing loss. Three ears with cochlear loss had mild hearing loss with peak at 2 kHz (CochHL-Peak) and the other 4 had mild flat cochlear hearing loss (CochHL-Flat). The age range of the participants in this group was 19–35 years with a mean age of 25 years.

All of them had ‘A’ type tympanogram with acoustic reflexes proportional to their degree of hearing loss. Auditory brainstem responses were proportional to the degree of hearing loss, which indicated no indication of retro-cochlear pathology. Click evoked oto-acoustic emissions were absent in all the participants. Speech identification scores were proportionate to their degree of hearing.

Control Group - Normal hearing group (Normal): Ten participants with an age range of 18 to 26 with a mean age of 21 years were considered for the study. All the participants had pure tone thresholds within 15 dB HL through octave frequencies from 250 Hz to 8000 Hz and speech identification scores greater than 60% at 0 dB SNR for PB word list. Additionally all the participants had A or As type tympanograms with acoustic stapedial reflex thresholds within 95 dBHL and normal click evoked oto-acoustic emissions.

Procedure

The main procedure used to obtain the data consisted was divided into three experiments. All the participants underwent three experiments. Experiment I involved the assessment of temporal resolution. Experiment II involved the assessment of the frequency resolution. Experiment III involved the assessment of consonant Perception.

Experiment I – Assessment of Temporal Resolution

Stimuli: Broadband noise was used as a carrier stimulus on which sinusoidal modulations were

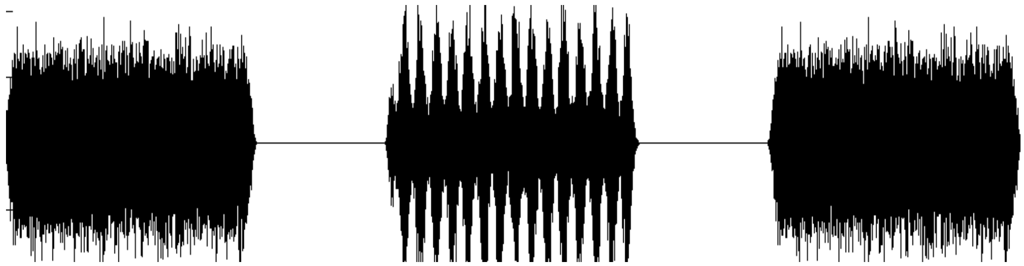

Figure 1: The unmodulated noise signals (left and right) and the modulated noise signal (middle)

imposed to test for the amplitude modulation detection threshold. The Noise carrier denoted as n(*t*) with a total duration (t) of 500 ms. Sine waves of frequencies equal to 2, 4, 8, 16, 32, 64, 128, 256 Hz and these were used to modulate the broad band noise carrier, as shown in equation 1.

Modulated noise=c[1 + *m* sin (2 $\pi f_m t + \acute{\O}_m$)] × n(t) {as in Viemeister 1979} .....*Equation 1*

Where, '*m*' is the modulation depth (0≤ *m*≤ 1), 'f_m' is the modulation frequency ($f_m$was 2, 4, 8, 16, 32, 64, 128, 256 and 512 Hz), and 'ϕ_m' is the starting phase of the modulation, randomized on each interval. The term '*c*' is a multiplicative compensation term (Viemeister, 1979) set such that the overall power will remain the same in all intervals. The intensity of the modulated noise was set to 60 dB SPL and the modulation depths considered here were in the logarithmic units of 0 to -40 dB{modulation depth (dB)=20 log(m)}. The waveforms of the modulated and unmodulated noise signals are shown in figure 1.

Experiment II – Assessment of the frequency resolution

Stimuli: Pure tones of frequencies (Fc) 500 Hz, 1000 Hz and 2000 Hz with durations of 300 msec were used as the test stimuli. Notched noises centered at frequencies (Fc) 500 Hz, 1000 Hz and 2000 Hz were used as the masking stimuli. To generate the notched noise, first, a random noise was band filtered by limiting the outer edges of the band on either side of Fc to $0.8 \times Fc$. Notches in this band pass noise were created by setting the amplitudes of all frequency components between Fc and $Fc \pm (g \times Fc)$ ('Fc' is the probe signal frequency and 'g' is the normalized notchwidth $\frac{|f - Fc|}{Fc}$, where *f* is the edge frequency of the notch on either side of Fc) to zero amplitudes. The intensity of the notched noise was set to 50 dBSPL across all the noise components, such that the resulting noise had a spectrum level of 50 dBSPL/Hz. The spectrum of notched noise centered at 1 kHz is shown in Figure 2.

Notched noises centered at 500 Hz, 1000 Hz and 2000 Hz were generated using the above mentioned method. The normalized notchwidths used were, g equal to 0.0, 0.1, 0.2, 0.3, 0.4, and 0.5. The notched noises were then mixed with the probe signals to constitute the variable signal with the pure tone placed temporally at the center of the masking stimuli. A 400 msec notched noise alone token constituted the standard signal. Figure 3 shows the spectrum of the notched noise along with the probes signal.

Procedure for Experiment I and Experiment II

A 3 Interval - 3 Alternative Forced Choice procedure was used in both the experiments using psychoacoustics toolbox in Matlab developed by Grassi and Soranzo (2009). The variable signal was varied adaptively using a 2 Down-1Up stepping rule to converge on to the 70.7% point on the psychometric function at each notch width for each probe frequency. The starting size in experiment I was 8 dB and after the first reversal the step size was reduced to 2 dB. The starting step size in experiment II was 8 dB and after the first reversal the step size was reduced to 5 dB and the following reversals had a step size of 2 dB. A total of six reversals were considered, and the average of the midpoints of the six reversals was considered as the threshold.

The participants were made to indicate verbally or by pressing the keys on the keyboard as to which of the three tokens presented had a pure tone or modulated noise in them. If the participant did not hear any pure tone or modulated noise, (s)he was asked to respond with a random guess.

Experiment III - Assessment of consonant perception

Stimuli: Twenty consonants | p t k b d g t^h d^h ʧ ʤ l r m n h j ṇ w ḷ s ʃ | were chosen to study the perception of all the consonants in Kannada by those with Auditory dys-synchrony. These consonants were recorded in the context of vowel |a| to obtain Vowel-Consonant-Vowel combinations (VCV). Each VCV was recorded by three adult male native speakers of Kannada in a

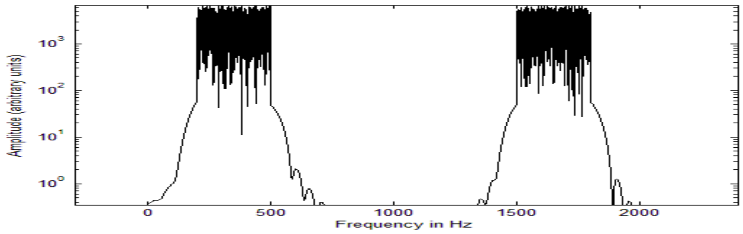

Figure 2: Spectrum of notched noise with Fc= 1000 Hz and g = 0.5.

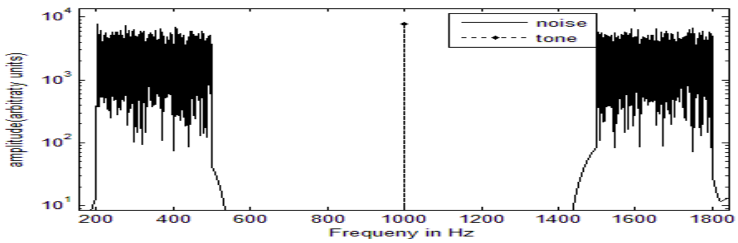

Figure 3: Spectrum of notched noise centered at 1000 Hz with g = 0.5 along with the 1000 Hz probe tone.

clearly articulated manner to obtain a set of sixty multitalker VCV stimuli. The recording was done using Adobe Audition (Version 1.5) software at a sampling frequency of 44100 Hz and resolution of 16 bits.

Procedure

The monoaural consonant perception testing was performed using APEX 3 software developed at ExpORL (Francart, van Wieringen, & Wouters, 2008). The presentation of the stimulus was in a random order and the each stimulus was presented 200 ms after the participant's response. All the speech stimuli were presented monaurally at the most comfortable levels.

Response pattern

Twenty VCV combinations were displayed on a computer screen in Kannada script using a graphical user interface on an APEX 3 platform. The participants were made to respond to the speech stimuli which (s)he heard by clicking on the appropriate VCV combination out of the twenty VCV combinations displayed on the computer screen.

Instrumentation

Matlab version 7 was used for the stimulus generation in Experiment 1 and Experiment 2, and all the stimuli were gated using a cosine ramp with 20 msec rise/fall time. In all the experiments the stimuli from the sound card of the PC were reproduced by the RP2.1 processor (TDT sys3) at 50 kHz sampling rate and attenuated using a PA5 programmable attenuator to achieve desired intensity levels. The PA5 output was then used to drive a pair of Sennheisser HDA200 headphones through an HB7 headphone buffer. All the experiments were done in a sound treated room and the ambient noise levels in the testing room were within the permissible limits as per ANSI S3.1 (1991).

Analyses

The AN/AD group were divided into three subgroups based on the degree and pattern of hearing loss. The first group consisted of 11 ears with normal hearing sensitivity (AN/AD normal), the second consisted of 3 ears with mild low frequency hearing loss with peak at 2 kHz (AN/AD-Peak) and the third group consisted of 4 ears with mild flat hearing loss (AN/AD-Flat). The CochHL group consisted of3ears with low frequency cochlear hearing loss and peak at 2 kHz (CochHL-Peak) and 4 ears with mild flat cochlear hearing loss (CochHL-Flat). The control group consisted of 18 ears with normal hearing sensitivity (Normal). All the comparisons were made both across groups and across sub-groups.

Temporal Resolution

Temporal modulation transfer function (TMTF) was obtained from the temporal modulation detection data by fitting the data to a lowpass butterworth filter of first order as in Zeng et al.(2004). The parameters peak sensitivity of the TMTF denoted by 'Pk' and bandwidth of the TMTF denoted by 'BW' were derived from the filter function. These two parameters

'Pk' and 'BW' were the measures of temporal resolution and were used for further analysis.

Frequency resolution

Auditory filter shapes were derived by fitting the notched noise data to a double rounded exponential function as used by Glasberg and Moore (1990). The roex parameters 'p', 'r' and 'ERB' are derived. The parameter 'p' denotes the slope of the pass band of the auditory filter, the parameter 'r' denotes the slope of the tail of the filter and the parameter ERB denotes the equivalent rectangular bandwidth of the derived auditory filters. The p, r and ERB parameters for 500 Hz, 1000 Hz and 2000 Hz are denoted as p500, r500 and ERB500; p1000, r1000 and ERB1000; p2000, r2000 and ERB 2000 respectively.

Consonant Perception

Consonant confusion matrices were drawn based on the responses to the VCV combinations. The confusion matrices were then subjected to Sequential Information Transfer Function Analysis (SINFA) using FIX to analyze the pattern of errors. The matrices were analyzed according to parameters of manner, place and voicing features. The conditional information is calculated as number of bits of information transmitted per each feature out of the total number of bits of information held by the available per feature.

Results and Discussion

The deficits in the frequency resolution and temporal resolution were compared across the groups using non-parametric Kruskal-Wallis test and pairwise comparison of the groups was done further using the Mann-Whitney U test. The Pearson's product-moment correlation between the parameters of frequency resolution and temporal resolution with that of the information transmitted for manner, place and voicing feature was evaluated. The parameters which yielded the best correlation were then used to fit a model through linear regression using a least squares design to investigate the relationship between each speech feature and the temporal resolution and frequency resolution in AN/AD.

Temporal resolution

The mean TMTF had a low pass shape in the Normal group and the CochHL-Flat sub-group, whereas, all the AN-AD subgroups and the CochHL-Peak had a TMTF which was band pass in shape. Figure 4 shows the TMTFs across the groups and sub-groups. The mean modulation detection thresholds were the best in the Normal group followed by the CochHL group and least in the AN/AD group. Comparison across the groups showed that the peak sensitivity varied significantly across the groups [$\chi^2(2, N)=32.012$, $p<0.000$]. The AN/AD group had significantly lower peak sensitivity (-10.78 dB) and bandwidth (23 Hz) compared to the Normal and CochHL group. There was no significant difference between the peak sensitivity of the Normal (-21.72 dB) and the CochHL group (-20.25 dB). However, the bandwidth of the TMTF was significantly lower in the CochHL group (54.91 Hz) compared to the Normal group (72.7 Hz). Comparison across sub-groups of AN/AD showed that there was no difference between the sub-groups of AN/AD for both peak sensitivity and bandwidth of the TMTF. This finding implies that the peak sensitivity and bandwidth of the TMTF in the AN/AD group is not dependant on the audibility factor, as the sub-groups of AN/AD had different audiometric patterns yet they had similar TMTFs. The TMTFs in this study were measured at a constant intensity level of 65 dB SPL and if there were any difference in the modulation sensitivity because of the effect of audibility it would have shown as a difference in peak sensitivity and bandwidth across the sub-groups of AN/AD.

The CochHL-Peak and CochHL-Flat sub-groups were not significantly different from that of the Normal group in terms of the peak sensitivity of the TMTF. However, the Bandwidth was significantly different across the sub-groups [$\chi^2(2, N)=31.844$, $p<0.000$]. The bandwidth in CochHL-Flat subgroup was significantly lower (BW=55.85) than that of the Normal group. But, the bandwidth in the CochHL-Peak (53.67) was not significantly different from the Normal group. The lower bandwidth in the CochHL-Flat group can be explained in two ways. The first being disrupted gross temporal resolution because of the cochlear damage and the second explanation being the effect of audibility. Bacon and Viemeister (1985) and Formby and Muir (1988) have shown that individuals with cochlear hearing loss had poorer sensitivity for detecting the modulations at higher modulation rates. The reasons they cited was the poorer audibility at the high frequency regions in the cochlea i.e. the higher frequency regions are more responsible for providing the accurate temporal envelope information as they have lesser inherent fluctuations. Moore (2007) also reported the same as above and concluded that the reduced temporal resolution in cochlear hearing loss is because of the reduced audible bandwidth, and when this was increased by increasing the sensation level, then individuals with cochlear hearing loss performed similar to that of the normal hearing listeners. This would give weight to the second explanation causing reduction in the bandwidth of the TMTF in the

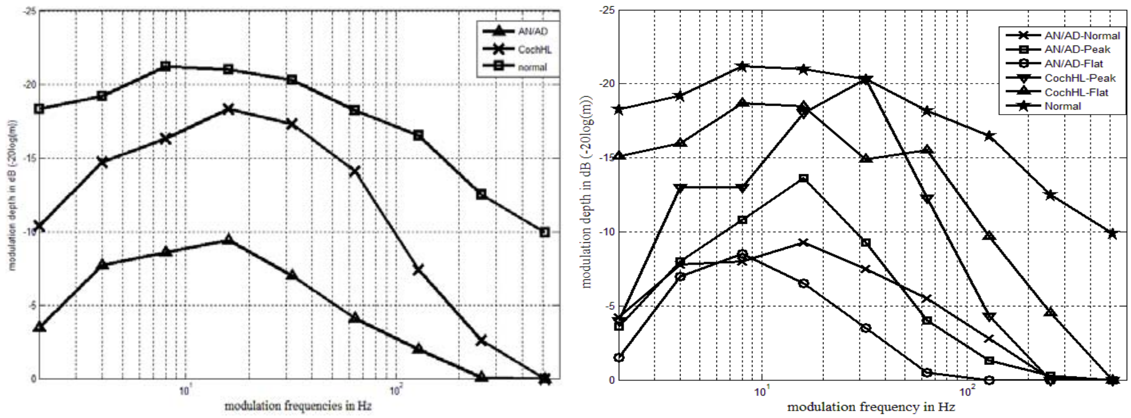

Figure 4: The TMTF across the groups (left panel) and the sub-groups (right panel).

CochHL group. The CochHL-Peak subgroup did not show any significant difference from that of the Normal group. It would be convenient to assume that this finding could have been because of the better hearing in the high frequencies in this sub-group, but it should be noted that the mean bandwidth in this sub-group was lower than that of the CochHL-Flat subgroup. This discrepancy can be attributed to the higher standard deviation in the CochHL-Peak which might have masked any difference and similarities in the TMTF between this sub-group and the Normal group.

The comparison of the AN/AD group with that of the cochlear group showed that the peak sensitivity and bandwidth of the TMTF are significantly lower in the AN/AD group. This is in accordance to Zeng et al., (1999, 2004), Kumar and Jayaram (2005) and Rance et al. (2008). This can be attributed to the well-established fact that AN/AD is a predominantly timing related disorder because of loss of neural synchrony which affect temporal resolution markedly (Starr et al., 1996; Zeng et al., 1999, 2004; Rance, 2005; Rance et al., 2008). Thus, temporal resolution is clearly different in cochlear hearing loss and auditory dys-synchrony.

The bandwidths of the TMTFs obtained in this study are in accordance with that of Formby and Muir (1988) and Eddins (1993) but do not go along with the bandwidths measured by Zeng et al., (1999). Zeng et al., (1999) obtained a mean bandwidth of 237 Hz in normal hearing listeners as opposed to a mean of 72 Hz in the current study. This might be because of the use of larger frequency intervals used by Zeng et al. (1999) which might have smoothened off the TMTF and overestimated the bandwidth.

It is evident from the TMTF that the low and the high modulation frequencies are more affected hence, giving the TMTF a band pass shape in those with AN/AD as opposed to the low pass shape in normal hearing listeners. All in all, individuals with AN/AD had poorer temporal resolution abilities irrespective of the audibility factor compared to cochlear hearing loss, in whom temporal resolution abilities are affected by the audibility factor.

Frequency resolution

The parameters r500, ERB500, ERB1000, p2000, and ERB2000 were the ones which were significantly different across the sub-groups. Table 2 gives the results of Kruskal-Wallis test for frequency resolution parameters across the sub-groups.

Table 2: Results of Kruskal-Wallis test across the sub-groups

roex parameter	Frequency resolution	χ^2	df	Sig.
	p500	4.237	2	.120
p	p1000	5.831	2	.054
	p2000	1.233	2	.540
	r500	22.15	2	.000
r	r1000	9.818	2	.007
	r2000	1.573	2	.455
	ERB500	32.62	2	.000
ERB	ERB1000	32.59	2	.000
	ERB2000	18.63	2	.000

Figure 5 shows the means and standard deviations for the parameters p and r for the three frequencies across the groups and sub-groups. The parameter r500 which decides the tails of the auditory filters, was significantly higher in the individuals with AN/AD which indicates at a steep tail of the auditory filter compared to the normal hearing individuals. However, the slope of the tail of the auditory filter at 500 Hz was not significantly different from that of the cochlear hearing loss group. The increased slope of the tail of the auditory filters in the AN/AD groups would lead to greater susceptibility to masking from spectral components even far away from the centre frequency. The mean equivalent rectangular bandwidth (ERB) at 500 Hz was nearly four times that of the normal hearing group and the cochlear loss with peak sub-were not the major reasons for the poorer frequency resolution abilities. The equivalent rectangular bandwidth at 1000 Hz was three and a half to four times greater in individuals with AN/AD compared to normal hearing listeners, which means that individuals with AN/AD have very broad auditory filters at 1000 Hz. But ERB at 1000 Hz in the AN/AD group was similar to that of the cochlear group. Additionally ERB at 1000 Hz in the cochlear hearing loss sub-group with mild flat hearing loss was greater than ERB seen in the normal group. This finding indicates at similarities in the frequency resolution between the cochlear hearing loss sub-group with a peaked audiogram and AN/AD group at 1000 Hz.

Findings similar to the raw notched noise data in the current study was reported by Kraus et al., (2000) where they found only a 3 dB change in threshold for change in notchwidth to 0.25 from 0.00. They attribute this factor to the over masking effect taking place in individuals with AN/AD as explained by Zeng et al., (2004). This small threshold shift with increase in notchwidth is what is seen even in individuals with cochlear hearing loss with broadened auditory filters i.e. individuals with broadened auditory filters show very small change in the masked threshold with changes in the notchwidth (Glasberg & Moore, 1986). Though, this was not cited as one of the possible reasons for the results reported by Kraus et al., (2000), the results of this study suggest that this small change in threshold seen in individuals with AN/AD at 1000 Hz could be because of the broadened auditory filters.

The parameter p2000 depicts the slope parameter 'p' which describes the slope of the pass band of the auditory filter (nearly the upper half of the auditory filter). This slope of the auditory filter pass band was significantly shallower in the subgroups of AN/AD with flat hearing loss when compared to the normal group. But the other sub-groups of AN/AD had 'p' similar to that of normal. This implies that the AN/AD subgroup with flat hearing loss had a shallower tip in their 2000 Hz auditory filter. This would also imply group. This was evidenced even in the raw notched noise data as no shift or 5 to 10 dB shift in threshold of detection of the tone, obtained in the notched noise even after varying the notchwidth in most of the individuals with AN/AD. The ERB at 500 Hz in the sub-group with flat cochlear hearing loss was higher than that of the Normal group and the sub-group of cochlear hearing loss with peak at 2 kHz. The bandwidth of the auditory filter in AN/AD was not significantly different across sub-groups. This implies that the pure-tone thresholds and the audiometric patterns

that the slope of the auditory filter at 2000 Hz was related to the absolute threshold at 2000 Hz in the AN/AD group. The slope of the auditory filter at 2000 Hz in the cochlear hearing loss group with peak at 2000 Hz was also shallower compared to the normal group. However, the same was not noticed in the cochlear hearing loss sub-group with flat hearing loss. This indicates that shallower auditory filter tips (pass band) are not necessarily related to the absolute thresholds in individuals with cochlear hearing loss.

The ERB at 2000 Hz in the AN/AD group was nearly one and half to two times that of the 2000 Hz ERB in the normal group. The ERB at 2000 Hz in AN/AD sub-group with flat hearing loss was significantly larger than ERB at 2000 Hz in the other AN/AD subgroups with lesser degrees of hearing loss. This shows a relation between the absolute threshold and the frequency resolution in AN/AD for 2000 Hz. And it was also seen that there was no difference in ERB at 2000 Hz for all AN/AD sub-groups compared to the cochlear hearing loss sub-group with flat audiometric pattern. This points towards similarities in the frequency resolution abilities between the cochlear hearing loss group and the AN/AD sub-groups for the higher frequencies.

The results of frequency resolution showed broader auditory filters at lower frequencies compared to the higher frequencies in individuals with AN/AD. This might also be the additional reason why the DLFs at lower frequencies are more impaired compared to the higher frequencies as seen in Zeng et al., (2004) and Barman (2008) apart from the explanation of place and temporal coding.

The results of the study showed poorer frequency resolution at 500 Hz through 2000 Hz in individuals with AN/AD. Though AN/AD is reported in literature to be a significantly timing related deficit, the results of the current study indicate that temporal resolution deficits coexist with frequency resolution in individuals with AN/AD. These results could be supported with the

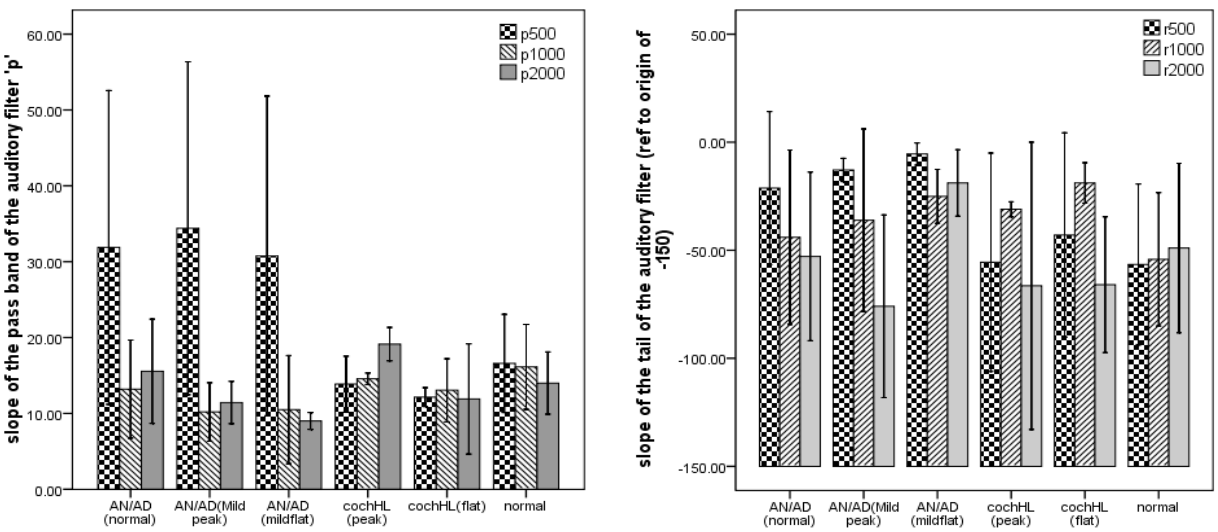

Figure 5: The means of parameter 'p' (left panel) and 'r' (right panel) across the groups/sub-groups. The error bars represent +/- 1 S.D.

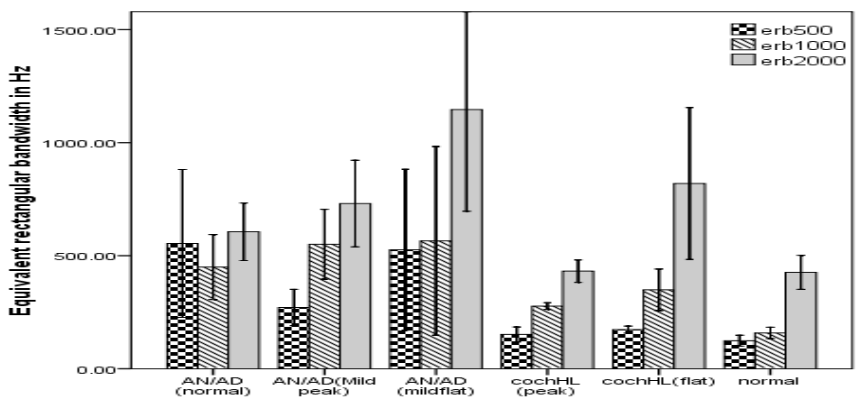

Figure 6: Equivalent rectangular bandwidth across the groups/sub-groups for the three frequencies. The error bars represent +/- 1 S.D.

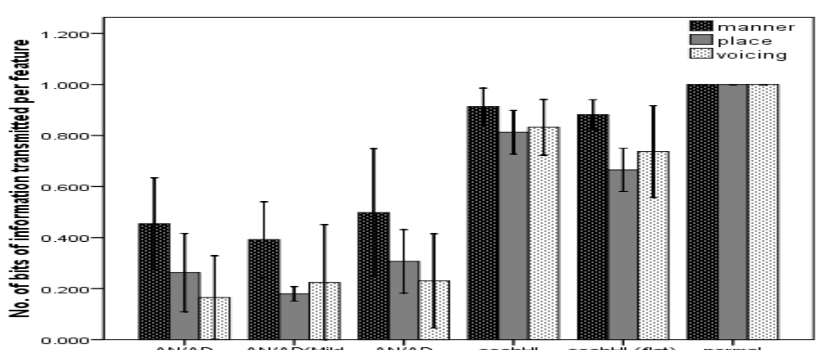

Figure 7: Conditional information transmitted for manner, place and voicing features across the subgroups. The error bars represent +/- 1 S.D.

findings of Starr et al., (2003) where they found 30% loss of outer hair cells in the apical turns of the cochlea in an individual with AN/AD. This suggests that individuals with AN/AD might have a dysfunction right at the level of the cochlea which might hamper the frequency selectivity. However, the results of the current study and that of Kraus et al., (2000) are not in accordance with the that of Rance et al., (2004) where they reported no difference in the threshold shift between a notched noise and a broadband noise across normal hearing listeners and individuals with AN/AD. This might be the consequence of the difference in procedure used, i.e. the masker stimuli were similar in the current study and Kraus et al., (2000), and however, Rance et al., (2004) used broadband noises (with and without notch) as maskers and also used a low intensity pink noise in their normal hearing listeners to simulate the effect of the elevated threshold.

Consonant perception

The results of consonant perception evaluation show that AN/AD group had the most difficulty in perception of manner, place and voicing compared to the cochlear hearing loss group. Voicing feature was perceived the least, followed by place feature. Manner was the best perceived feature. However, in the cochlear hearing loss group, manner feature was the best perceived, followed by voicing and place features. Manner was the best perceived feature in both AN/AD and the cochlear hearing loss groups. There was no significant difference in the perception of manner, place and voicing across sub-groups of AN/AD. This implies that there was no relationship between the audiometric pattern and speech perception in individuals with AN/AD. The consonant perception results are in accordance with reports by Rance et al., (2008), Narne and Vanaja (2008). They reported that the voicing errors are the predominant errors in AN/AD followed by place and manner errors. Figure 7 shows the conditional information transmitted for the three speech features assessed across the sub-groups.

Relationship between consonant perception and temporal and frequency resolution

The analysis of correlation between consonant features perceived and the temporal and frequency resolution parameters showed significant correlation between the consonant features perceived and the temporal and frequency resolution in individuals with AN/AD. Figure 8 shows the scatter plot of temporal resolution parameters as a function of the speech feature perceived. Figures 8 and 9 show the scatter plots of the frequency resolution parameters as a function of the speech features perceived. The perception of place and manner features correlated significantly with the peak sensitivity of the TMTF. This implies that individuals with AN/AD are not able to utilise the temporal cues which are important for the perception of manner and place cues. The temporal parameters which are most important for the perception of manner cues are the consonant duration and the temporal envelope. And any deficits in the temporal resolution would thus hamper the perception of the manner feature. The temporal parameters important for perception of place features are the rapid formant transitions which are again impaired because of the temporal resolution (especially the fine structure coding) deficit. The voicing feature however did not correlate with temporal resolution deficit. This however is in contrast to the results of Rance et al., (2010), wherein they reported good correlation between the perception of voicing feature and the gap detection threshold. This would mean that, the voice onset time might not have been an important cue, in the stimuli used in this study which were VCV combinations. This is in fair accordance with the results of Narne and Vanaja (2008) where they reported negligible improvement in the voicing perception after temporal envelope enhancement. They explain this by saying that the envelope enhancement changes the depth of the temporal envelope enhancing the voice onset time. They conclude that envelope enhancement does not bring about changes in the voicing bars and the first formant frequency and thus, the perception of the lower frequencies is important for voicing perception.

The errors in consonant feature perception for manner and place correlated with the frequency resolution parameter r1000 and ERB500. This implies that the errors in perception of the consonant features significantly correlated with the tail of the 1000 Hz auditory filter and the bandwidth of the 500 Hz auditory filter. The broader tail of the auditory filter at 1000 Hz might have led to increased susceptibility to the spread of masking, and poorer resolution of important spectral components of the speech sounds on the basilar membrane. The important frequency related parameters for perception of place feature are consonant spectrum, formant transition and formant frequencies of the preceding and succeeding vowels in a VCV combination. Affected place feature perception is also a consequence of significant frequency resolution deficit in the AN/AD group in the current study. This is in accordance to studies by Thibodeau and VanTasell (1987), Preminger and Wiley (1985) and Turner and Henn (1989) in individuals with cochlear hearing loss, and they attributed the poorer speech perception (place and manner) to affected frequency selectivity. The voicing feature correlated with the bandwidth of the auditory filter at 500 Hz and

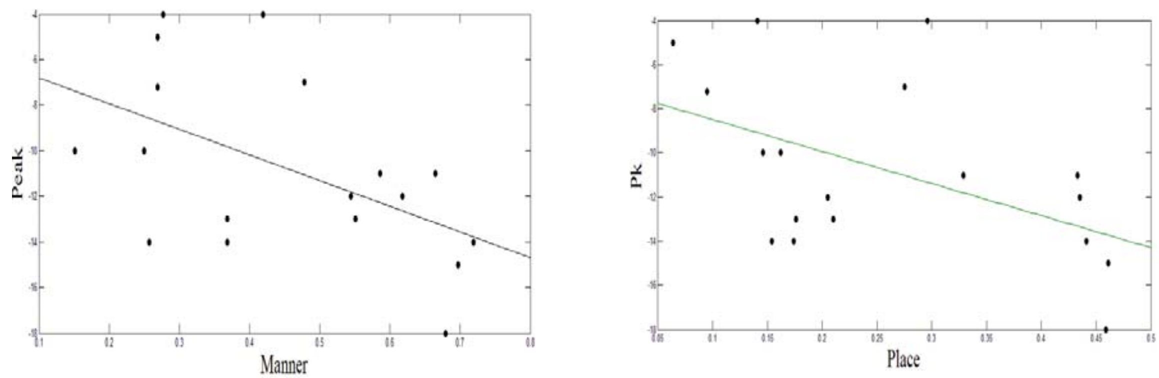

Figure 8: Scatter plot of the peak sensitivity against manner (left panel) and place features (right panel)

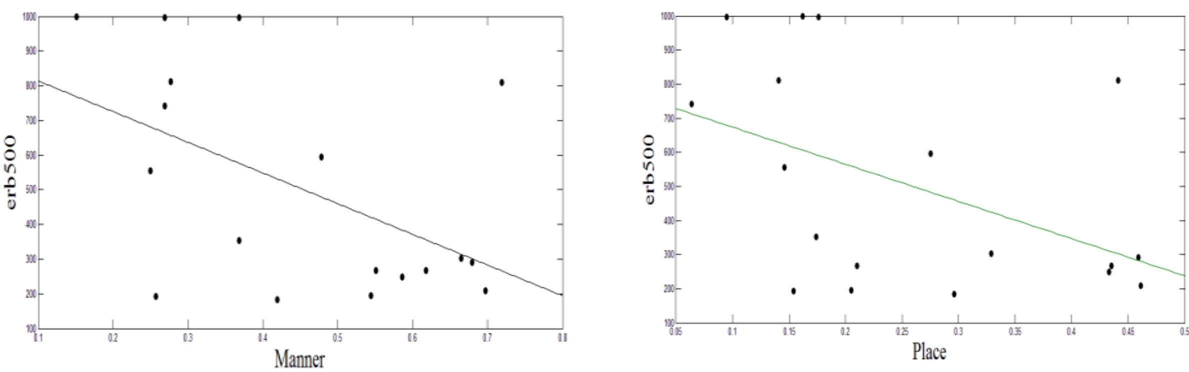

Figure 9: Scatter plot of erb500 and manner (left) and place features (right).

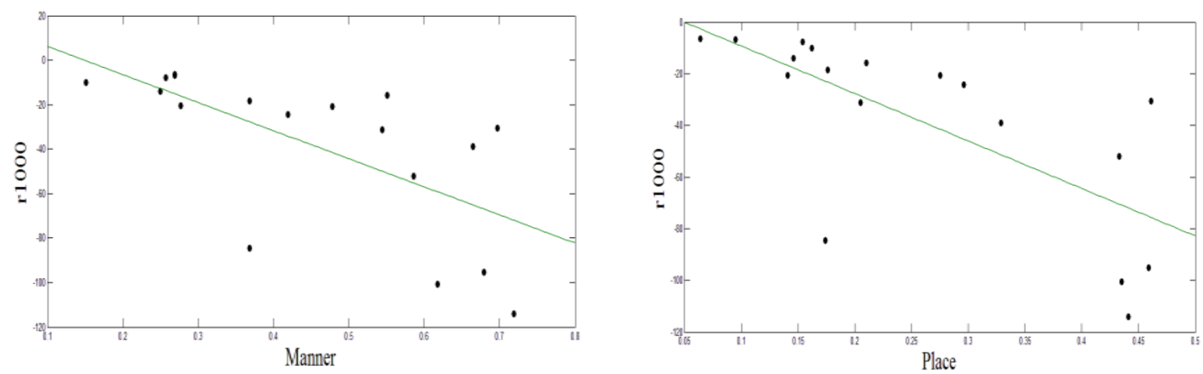

Figure 10: Scatter plot of r1000 and manner (left) and place features (right).

the slope of the tail of 1000 Hz auditory filter, and not the temporal resolution deficit. This could be because of the fact that the voicing murmur was probably the greater cue in the VCV stimuli used in our study compared to the voice onset time. Thus, this voicing murmur cue could have been significantly affected by the frequency resolution at low frequencies as the voicing murmur is a low frequency cue. This is in accordance to the findings of Narne and Vanaja (2008).

For the derivation of the model for manner and place, the ERB500 and Pk and r1000 were considered and for the derivation of the model for voicing, the ERB500 and the r1000 were considered. Only the above mentioned parameters were considered for the regression analysis as these were the parameters which showed significant correlation with the consonant perception measures. The regression analysis was carried out by fitting a range of linear equations (models) to the data, with the speech parameters as the dependent variables and the temporal and frequency resolution parameters as the independent variables. The models with the highest r-square value were chosen and have been given below.

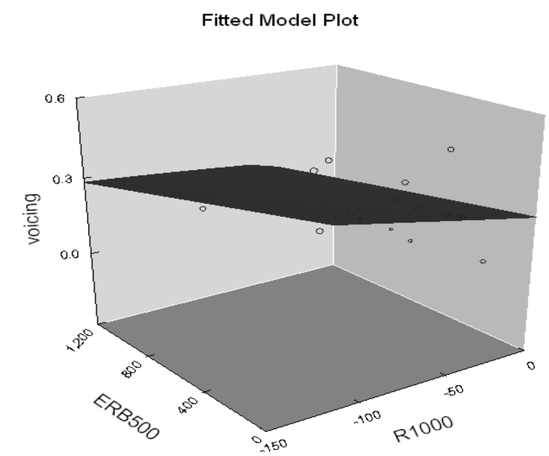

Figure 11: Fitted model plot of voicing based on ERB500 and r1000.

The model for predicting manner information transmitted in bits per feature based on ERB500 and Pk is as in equation 2. The model had an R-square value of 0.581 with F(3, 14)=6.462 at p=0.006.

Manner=0.39102-(0.0066 x Pk) – (0.00021 x erb500) – (0.0025 x r1000)

...Equation 2.

The place feature was also fitted with a model based on the ERB500, r1000 and Pk using a least squares method as shown in equation 3 to predict the place information transmitted in bits per feature. This model had an R-square value of 0.597 with F(3,14)=6.904 at p=0.004.

$$Place = 0.20892 - (0.00013 \times erb500) - (0.00225 \times r1000)$$

…Equation 3.

The model for predicting voicing feature based on ERB500 and r1000 is as in equation 4. The model had an R-square value of 0.691 with F (2, 15) = 16.808 at p = 0.0001.

$$voicing = 0.22890 - (0.00028 \times erb500) - (0.00260 \times r1000)$$

…Equation 4.

Figure 11 shows the fitted model plot for the voicing feature based on ERB500 and r1000. The dark shaded region shows the prediction region of the model for the voicing based on ERB500 and r1000. The dots represent the voicing feature transmitted across the ERB500 and r1000 for each subject.

Regression analysis of the relationship between manner and feature perception and frequency and temporal resolution was carried out to fit a model to investigate the extent of relationship between frequency and temporal resolution as shown in equations 2, 3 and 4. The model clearly gives lesser weightage to the bandwidth of the 500 Hz auditory filter and greater weightage to peak sensitivity of the TMTF and greatest weightage to the tail of the 1000 Hz auditory filter for the place and manner perception. The peak sensitivity had greatest weightage for the place perception, followed by r1000 and ERB 500. However, this model should be used with caution, as a linear model has been fitted for the data only for ease of calculation and this does not guarantee hundred per cent prediction accuracy.

Conclusions

It can be concluded from the current study that the frequency and temporal resolution are affected in individuals with AN/AD and this deficit is greater than what is seen in individuals with cochlear hearing loss. The errors in the perception of place and manner features in individuals with AN/AD are because of inefficient processing of the temporal cues in speech and the inefficient processing of the lower frequency components of speech. The voicing related errors in speech perception in individuals with AN/AD is primarily due to inefficient processing of the lower frequency information in the speech signal in the VCV combinations used in the study.

Acknowledgements

The authors would like to thank the Director, All India Institute of Speech and hearing for providing the infrastructure needed for the study. The authors also like to thank Dr. Vijayakumar Narne for his help throughout the study.

References

American National Standards Institute. (1991). Maximum permissible ambient noise levels for audiometric test rooms. ANSI S3.1-1991, New York: American National Standards Institute.

Barman, A. (2008). Psycho-acoustical profile in normals and individuals with auditory dys-synchrony. Unpublished doctoral thesis submitted to the University of Mysore, Mysore,India.

Berlin, C. I., Hood, L. J., Morlet, T., Wilensky, D., St.John, P., Montgomery, E., et al. (2005). Absent or elevated middle ear muscle reflexes in the presence of normal otoacoustic emissions: A universal finding in 136 cases of auditory neuropathy/dys-synchrony. *Journal of the American Academy of Audiology, 16*, 546-553.

Eddins, D. (1993). Amplitude modulation detection of narrow-band noise:Effects of absolute bandwidth and frequency region. *Journal of the Acoustical Society of America, 93*, 470-- 479.

Formby, C., & Muir, K. (1988). Modulation and gap detection for broadband and filtered noise signals. *Journal of the Acoustical Society of America, 82*, 1933-1943.

Francart, T., VanWieringen, A., & Wouters, J. (2008). APEX-3: A multi-purpose test platform for auditory psychophysical experiments. *Journal of Neuroscience Methods, 172*(2), 283-293.

Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *Journal of the Acoustical Society of America*, *79*(4) 1020 - 1030.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched noise data. *Hearing Research, 47*, 103-138.

Grassi, M., & Soranzo, A. (2009). MLP: A matlab toolbox for rapid and reliable auditory threshold estimations. *Behavior Research Methods, 41*(1), 20-28.

Kluk, K., & Moore, B. C. J. (2004). Factors affecting psychophysical tuning curves for normally hearing subjects. *Hearing Research, 194*(1-2), 118-134.

Knox, E. N. (2005). Diagnosis and management of auditory neuropathy/auditory dys-synchrony. *A doctoral thesis submitted to the Washington University School of Medicine.*

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

Kumar, A. U., & Jayaram, M. (2005). Auditory processing in individuals with auditory neuropathy. *Behavioral and Brain Functions, 1*(21), 1-8.

Kundu, P. & Rout, N. (2010). The impact of high gain conventional hearing aid on OAEs in a case of auditory neuropathy/dys-synchrony. *Eastern Journal of Medicine, 15,* 26-30.

Moore, B. C.J. (1973). Frequency difference limens for short duration tones. *Journal of the Acoustical Society of America, 54*, 610-619.

Moore, B. C. J. (2007). Temporal resolution in people with cochlear damage. In B. C. J. Moore, *Cochlear hearing loss: physiological, psychological and technical issues, second edition* (pp. 128-135). Chichestor, England: Wiley-Interscience.

Narne, V. K., & Vanaja, C. S. (2008). Effect of envelope enhancement on speech perception in individuals with auditory neuropathy. *Ear and Hearing, 29*, 45-53.

Patterson, R. D., & Moore, B. C. J. (1986). Auditory filters and excitation patterns as representations of frequency resolution. In B. C. J. Moore, *Frequency selectivity in hearing* (pp. 123-177). London: Academic Press.

Preminger, J., & Wiley, T. L. (1985). Frequency selectivity and consonant intelligibility sensorineural hearing loss. *Journal of Speech and Hearing Research, 28*, 197-206.

Rance, G. (2005). Auditory neuropathy/dys-synchrony and its perceptual consequences. *Trends in Amplification, 9*, 1-43.

Rance, G., Corben, L., Barker, E., Carew, P., Chisari, D., Rogers, M., et al. (2010). Auditory perception in individuals with friedreich's ataxia. *Audiology and Neurotology, 15*, 229-240.

Rance, G., Fava, R., Chong, A., Barker, E., Corben, L., & Delatycki, M. B. (2008). Speech Perception ability in individuals with friedreich ataxia. *Brain, 131*, 2002-2012.

Rance, G., McKay, C., & Grayden, D. (2004). Perceptual characterization of children with auditory neuropathy. *Ear and Hearing, 25*(1), 34-46.

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

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

Starr, A., Sininger, Y., & Pratt, H. (2000). The varieties of auditory neuropathy. *Journal of Basic Clinical Physiology and Pharmacology, 11*(3), 215-230.

Starr, A., Michalewski, H. J., Zeng, F. G., Fujikawa-Brooks, S., Linthicum, F., Kim, C. S., et al. (2003). Pathology and physiology of auditory neuropathy with a novel mutation in the MPZ gene (Tyr145-Ser). *Brain, 126*, 1604-1619.

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

Turner, C., & Henn, C. (1989). The relation between vowel recognition and measures of frequency resolution. *Journal of Speech and Hearing Research*, *32*, 49–58.

Varga, R., Kelley, P. M., Keats, B. J., Starr, A., Leal, S. M., Cohn, E., et al. (2003). Non-syndromic recessive auditory neuropathy is the result of mutations in the otoferlin (OTOF) gene. *Journal of Medical Genetics, 40*, 45-50.

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

Vinay, S. N., & Moore, B. C. J. (2007). Ten(HL)-test results and psychophysical tuning curves for subjects with auditory neuropathy. *International Journal of Audiology, 46*, 39-46.

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

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., & Starr, A. (2004). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology, 93*, 3050-3063.

Zeng, F. G., Oba, S., Garde, S., Sininger, Y., & Starr, A. (1999). Temporal and speech processing deficits in auditory neuropathy. *NeuroReport, 10*, 5429-5435.