Effect of Spectro-Temporal Enhancement on Speech Perception in Individuals with AD

Bhuvana S

13AUD006

This Dissertation is submitted as part fulfilment
for the Degree of Master of Science in Audiology
University of Mysore, Mysore

May 2015

Certificate

This is to certify that this dissertation entitled **"Effect of Spectro-Temporal Enhancement on Speech Perception in Individuals with AD"** is a bonafide work in part fulfillment for the degree of Master of Science (Audiology) of the student with Registration No. 13AUD006. This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore,

May 2015

Dr. S. R. Savithri

Director

All India Institute of Speech and Hearing

Manasagangothri,

Mysore – 570006

Certificate

This is to certify that this dissertation entitled **"Effect of Spectro-Temporal Enhancement on Speech Perception in Individuals with AD"** has been prepared under my supervision and guidance. It is also certified this has not been submitted earlier in other University for the award of any other Diploma or Degree.

Mysore,

May, 2015

Prof. Animesh Barman

Guide

Professor of Audiology

Department of Audiology

All India Institute of. Speech and Hearing

Manasagangothri, Mysore – 570006

Declaration

This dissertation entitled **"Effect of Spectro-Temporal Enhancement on Speech Perception in Individuals with AD"** is the result of my own study under the guidance of Prof. Animesh Barman, Professor in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore, **Register No**: **13AUD006**

May, 2015

Dedicated to

The first synchronous neural firing, a wonderful conception of 'The CREATOR'

Acknowledgement

"Feeling gratitude and not expressing it is like wrapping a present and not giving it"- William Arthur Ward

I would like to take a moment here to thank everyone in my life for helping me reach to this stage of life. And seek Thou blessings to help move ahead in the beautiful life I am gifted with.

They say we can't choose our family but can choose our friends. But I say appa, amma and kitty, God tagged me along with you to get me a family but I got a bonus of three great friends for life. Cannot find any more words to describe a mother's care, a father's love or a brother's concern. Thanks for being what you are in my life.

I would like to thank Animesh sir, for he is one of the first inspirations for me to continue in the field of Audiology. I consider myself a privileged person to get this opportunity to do the dissertation under your supervision. Every interaction with you, how you clear the doubts, give analogies to explain the concepts, I should say, you continue to inspire me sir. Thank you sir.

Vijay sir, I don't know if thank you can suffice for the help and knowledge you imparted throughout the process of this dissertation. Be it a silly doubt or a complex MATLAB code, the answer we knew was you sir. Your 'always ready to help' attitude is something I respect a lot and wish to imbibe that in me too.

I extend my gratitude to our director, Prof S. R. Savithri and HOD-Audiology, Ajith sir for giving me the opportunity to carry out this research.

I would like to thank all the clinical staffs and lecturers who have molded me to the clinician I am today and giving me the priceless experience of working in the clinics.

I would like to express my heartfelt thanks to Vasanthlakshmi ma'am for introducing a beautiful branch of Mathematics to us. Your words- "Don't approach the test directly, understand the data first", I shall remember that my whole life ma'am!

Madhuri and PG, the moment we all said psychophysics as area of interest when sir asked, I thought to myself: aren't they the right partners I can get? With the journey from being classmates to co-pros to a mad team of 'Panda-Hippo-Julien', this process couldn't be any more fun. And then, I changed my thought, you weren't the right partners, you were the best partners! Thanks for being there ALWAYS!

Dear mafia family, I thank everyone individually for the great part of my life you guys form. Did I just mentally recall everyone in the attendance order?! Hmmm, maybe I will cherish and nurture that as a beautiful circle of trust in brain..! Lastly, I hope you all continue to follow the code of mafia family- "once entered, cannot get out". I thank all my BSc batch mates and the MSc batch mates for being a wonderful bunch in making the college life more memorable.

I would like to thank all my seniors and juniors who helped me in one or the other way in the process of this dissertation. Thank you!

I shall fail if I don't thank AIISH for what it has given me in the last 6 years. What I can think of is only gains without a single loss. AIISH, it will run into pages if I explain. In simple terms, you taught me what life is. Thank you. A special thanks to the library and the staffs there. Thanks for being considerate in extending the closing time always.

My sincere acknowledgement to the participants of the study for their selfless participation which helped this study get a meaning.

PS: I feel am on a safer side not mentioning names, my friends call it, I guess "name retrieval deficit".

Abstract

Auditory dys-synchrony (AD) is a unique disorder which presents itself with impaired speech understanding ability in the presence of relatively spared audibility. The management options for individuals with AD are hearing aids and/or a cochlear implant. Although the listening devices helps these individuals to communicate, the improvements observed are not consistent across individuals. There is a need to see if there are other management options like speech enhancement that can improve speech perception in individuals with AD. Thus in the present study, two types of enhancement strategies (companding and consonant enhancement) were taken to see if they bring about any change in speech intelligibility. To analyse the same, consonant identification scores were obtained from two groups of individuals (10 individuals with AD and 10 individuals with normal hearing sensitivity) in different signal to noise ratio. Individuals with AD showed a significant improvement with companding when compared to unprocessed condition only in quiet situation. However with consonant enhancement, the improvement noticed was not statistically significant. In contrast, normals did not show any advantage with processed stimuli when compared to unprocessed condition across different signal to noise ratios. Sequential Information Transfer Analysis for AD revealed that the information transmitted was most for manner feature followed by place and voicing features. The improvement following enhancement are probably due to the increase in the spectral contrast of the speech stimuli. The improvement in the contrast might have helped the individuals with AD to extract the acoustical cues available and aided speech perception.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Hearing forms an integral part of effective communication. The physical process of hearing is the ability of the auditory system to detect/perceive a sound by sensing the vibrations in the environment. Interference in this process due to any abnormality along the auditory pathway lead to loss of audibility and perception of sounds. One such interference in the auditory system is auditory dys-synchrony which is recognised as a discrete disorder triggered mainly by disruption in the interpretation of timing information by the auditory nerve.

Auditory dys-synchrony (AD) usually presents itself with unique features and impact of the same perceptually is more severe. The typical characteristic of AD is the presence of normal outer hair cell functioning; and dysfunction of the auditory nerve making it a retro-outer hair cell disorder (Starr, Picton, Sininger, Hood & Berlin, 1996). The underlying pathophysiology that has been reported in literature varies from being at the level of the inner hair cells, synapse between inner hair cells and the auditory nerve or the discharges at the level of auditory nerve or the brainstem (Zeng & Shannon, 1999; Kraus et al, 2000). As a consequence of this desynchronised firing, the conduction of timing information is disturbed. The hearing sensitivity in individuals with AD with respect to pure tone detection threshold can range from normal hearing sensitivity to profound hearing loss. However almost 60-70% of them reported to have disproportionate speech identification scores which cannot be predicted from their pure tone thresholds (Sininger & Oba, 2001; Zeng, Kong, Michalewski, & Starr, 2005). Thus, speech identification ability and therefore communication stands out as a major problem faced by these individuals which aggravates in the presence of noise (Rance, Barker, Mok, Dowell, Ricton & Garratt., 2007). These discrepancies between speech identification and tone identification ability in individuals with AD is attributed to distorted temporal cues at suprathreshold level (Zeng, Oba, Grade, Sininger, & Starr, 2001; Kumar & Jayaram, 2006).

Outcome of a series of research have shown that temporal perception is affected mainly in individuals with AD (Zeng & Shannon, 1999; Zeng et al., 2001; Zeng et al., 2005 & Kraus et al., 2000). It has been shown that there is a good correlation between speech identification ability and temporal modulation detection function in individuals with AD (Zeng & Shannon, 1999). Kumar and Jayaram, (2005) reported that there is predominantly affected temporal processing which directly affects the speech perception. They also concluded saying audibility is not the factor, impairing the speech perception in individuals with AD. The impact of temporal processing on speech perception showed that temporal smearing of the speech reduces the contrast between consonant and vowels reducing the consonant to vowel ratio (Drullman, 1995; Zeng & Shannon, 1999). This makes it difficult to extract the cues for identification of consonants and vowels and thus impairing speech perception. In addition to impaired temporal processing, the frequency resolution has also shown to be compromised in this population (Vinay & Moore, 2007). It can be said that speech perception deficit observed in AD is due to interaction between reduced frequency selectivity and impaired temporal perception.

There are various management options that have been attempted in order to rehabilitate individuals with AD. The most conventional rehabilitation strategy of hearing aids are usually tried to improve the hearing ability. A person with sensory hearing loss obtains a good benefit in terms of audibility and speech perception from hearing aid which incorporates a non-linear compression. However, many studies conclude that the advantage of a hearing aid to a person with auditory dys-synchrony is minimal as their primary complaint would be difficult to understand speech rather than hearing it (Berlin, Hood, Morlet, Rose, & Brashears, 2003). The results yield no significant improvement especially in speech intelligibility (Berlin et al., 2003). Due to limited success with hearing aids, cochlear implants are emerging as the choice of treatment for AD. Along with cochlear implants, FM systems, speech reading and cued speech have also been tried to rehabilitate individuals with auditory dys-synchrony (Kraus, 2001). Literature suggests that this population can be benefitted from cochlear implantation (Trautwein, Sininger & Nelson, 2000; Sininger & Trautwein, 2002 & Budenz, et al., 2013).

However, the recent studies indicate that not all children with auditory dys-synchrony show benefit from it (Gibson, & Sanli, 2007; Range- Samuelson, Drake, & Wackym, 2008). Other limitations of CI like cost and affordability, pure tone threshold indicating normal to mild-moderate loss might contradict the consideration of CI as an efficient option. These discrepancies in the results create a need to explore alternative strategies that may help to perceive speech by individuals with auditory dys-synchrony.

A variety of psychophysical techniques have also been tried to check for improvement in speech perception. Kraus et al. (2000) demonstrated that the individuals with auditory dys-synchrony fail to utilize a fast rate of spectro-temporal changes. This is one of the important acoustic cue for perception of speech, especially the stop consonants. Thus, it makes it difficult to perceive the stop consonants in individuals with AD.

Baer, Moore and Gatehouse (1993) suggest that processing of signal to enhance spectral contrasts increases the prominences of the peaks. They observed that a moderate enhancement of speech contrasts improves intelligibility and quality rating of speech. When it was incorporated to study the speech perception scores, they obtained better scores in enhanced condition than the unprocessed condition. This was designed considering the fact that the individuals with sensorineural hearing loss has reduced frequency resolution which causes an inability to resolve the spectral shape. This basis can also be generalised for individuals with AD as they exhibit a similar problem of frequency processing. A study in this domain (Narne & Vanaja, 2008a) has shown an improvement in speech perception when envelop was enhanced in individuals with AD.

Alternative means of improving speech intelligibility is to ‘speak clearly’ (Picheny, Durlach, & Braida, 1985, 1986, 1989). This means that when a speaker is asked to speak clearly, he/she usually produce more intelligible speech than they would when they are casually conversing. The higher intelligibility in clear speech than in conversational speech results is due to the differences of acoustic and phonetic aspects of speech between the two types. The clear speech emphasises on reduced speaking rate, increased energy in the 1000–3000Hz range, enhanced temporal modulations, expanded voice pitch range and vowel space (Ferguson & Kewley-Port, 2002; Krause & Braida, 2004 & Liu, Del Rio, Bradlow, & Zeng, 2004). These advantages of clear speech is found to be advantageous in various population with auditory neuropathy, and individuals with cochlear implants (Kraus et al., 2000; Zeng & Liu, 2004) reported that there is a higher advantage of clear speech when given through electric stimulation when compared to acoustic stimulation.

Another such technology is burst enhancement which incorporates enhancement of burst portion of the consonant modifying the consonant to vowel ratio. This would bring about an amplification of the spectral contrast. Guelke (1987) suggested that, increasing the amplitude of the burst of a stop consonant up to the level of the vowel in normal speech could result in increased intelligibility by perceiving information that would otherwise be inaudible to a person with hearing impairment. The increase in intensity following the gap would also help indicate the presence of a stop consonant. They suggested that this technique might aid in improving the perception of speech. Hazan, Simpson and Huckvale (1998) also studied the effect of consonant enhancement on speech perception in listeners with normal hearing sensitivity. They obtained an improvement with processed condition when compared to unprocessed stimulus condition. Other studies (Simpson, Moore & Glasberg, 1990; Summerfield, Foster, Tyler & Bailey, 1985) examined spectral enhancement by narrowing bandwidths and digital signal processing and showed a slight improvement in intelligibility of speech.

Companding is one of the novel ideas which is a combination of the process "compressing" and "expanding." It is a technique which is developed based on relatively broadband compression followed by more frequency selective expansion. Turicchia and Sarpeshkar (2005) proposed this strategy for time domain spectral enhancement. This approach tries to mimic certain properties shared by the peripheral auditory system. This technique enhances the spectral peaks in a stimulus, relative to nearby spectral valleys. It is observed that the spectral enhancement techniques mainly modify the regions of the signal that contain acoustic cues for consonant identification to aid the perception. This retains the consonant cues makes it more resistant to further degradation of the stimuli.

Investigators have used these technique in cochlear implant subjects (Oxenham, Simonson, Turicchia, & Sarpeshkar, 2007; Bhattacharya & Zeng, 2007). They demonstrated some improvement in speech intelligibility for cochlear implant users and for normal-hearing subjects.

Speech perception with lengthened transition duration was studied by Kumar and Jayaram (2011) in individuals with AD. The CVs with enhanced transition duration, have shown an improvement in speech perception both with respect to placement and voicing feature. However the mechanism underlying this improvement remains unclear.

Due to lack of benefit with the conventional amplification system, some alternate management options like enhancement of the signals are tried to evaluate the usefulness in perception of speech in individuals with auditory dys-synchrony. There is still scope for fine tuning these strategies as many techniques that have been developed, either improve the signal-to noise ratio (SNR) without any improvement in intelligibility, or have added artifacts by introducing spurious sounds and enhancing random spectral peaks. Thus, the aim should be to choose an appropriate technology that would increase the intelligibility of speech in favourable and unfavourable listening environments in individuals with auditory dys-synchrony which the current hearing aids do not do efficiently.

1.1 Need for the study:

Multiple studies are carried out concerning the issue of management of AD which poses a huge challenge. Each study which incorporates the application of psychophysical techniques have seen a considerable improvement in speech perception (Bhattacharya & Zeng, 2007; Guelke, 1987; Narne & Vanaja, 2008a). The same technology was also used to study the speech perception in individuals with AD (Narne, Barman, Deepthi & Shachi, 2014). They considered 10 participants with AD to check the perception of consonants and sentences in quiet and at different SNRs (0, 5, 10 and 15 dB). There was a significant improvement observed in terms of increment in consonant identification scores in individuals with AD especially in quiet and 15dB SNR.

Consonant enhancement for AD individuals are tried in various studies. They have either considered enhancement with respect to bandwidth or duration of burst or transition. These are done owing to the perceptual consequence of the pathophysiology of AD where temporal perception is the severely affected. However, there is a need to see if a spectrally enhanced stimuli (amplitude of consonant enhanced) can bring about any change in the perception.

Also, there are no comparison made across strategies in the same study which helps to compare the improvement with each strategy. Also, the studies of speech perception have either taken words or sentences as the stimuli which possess some redundancy within them. This makes it difficult to attribute the differences in perception to the different techniques. To overcome the same the present study considered consonant vowel combinations which has very less inherent redundancy.

It is established that speech perception has greater detrimental effects especially in the presence of noise in AD when compared to normal hearing counterparts. Hence, this poses a need to test the robustness of the improvement provided by the companding and consonant enhancement algorithms and compare the same across different levels of noise. Furthermore, there are no controlled studies which report of comparison between the strategies to compare the improvement across different signal to noise ratios. It is also required to identify the spectro-temporal cues that are critical in providing the improvement in speech intelligibility that can help us further rehabilitation procedure required for individuals with auditory dys-synchrony.

1.2. Aim of the study:

To evaluate the effect of companded and consonant enhanced speech on speech perception in quiet and in noise among normal hearing individuals and individuals with auditory dys-synchrony.

1.3. Objective of the study:

Thus the present study has been taken up with the following objectives.

1. To evaluate the effect of spectral-temporal enhancement using companding and consonant enhancement strategy on speech perception in quiet and noise among individuals with normal hearing sensitivity and auditory dys-synchrony.
2. To evaluate the relative benefit of processed signal at each listening condition across the two groups.
3. To compare the speech perception scores obtained using companding and consonant enhancement and to compare the amount of improvement seen with processed stimulus across the two groups and two subgroups.
4. To analyse the consonantal error patters in terms of voicing, manner and place of articulation cues perceived in the AD group.

Chapter 2

Literature Review

Auditory neuropathy now referred as auditory dys-synchrony (Berlin, Hood & Rose, Hood & Ross, 2001) is a distinct disorder initially described by Starr et al in 1991. It is a hearing impairment mainly caused due to dysfunction in the auditory nerve in the presence of normal functioning outer hair cells (Star, Picton, Hood & Berlin, 1996). According to Sininger & Oba (2001), to diagnose as auditory dys-synchrony one must be having the following clinical symptoms:

a) Poor auditory functioning which is shown as a poor speech perception scores disproportionate to the pure tone thresholds.

b) Affected neural functioning shown as abnormal or absent auditory brainstem responses and elevated/absent auditory brainstem reflexes such as stapedial reflex, and

c) normal outer hair cell function evidenced by presence of cochlear microphonics or otoacoustic emissions.

Interest in the field of auditory dys-synchrony has persisted ever since it was described by Starr about the typical findings about it in around mid-1990s. Initially researchers were more concerned about the causative factor of AD. Slowly, the focus shifted towards the identification of different psychoacoustic parameter that affects their speech perception ability. Currently, the concern is more towards management of the condition by researching on various strategies available.

2.1 Prevalence of Auditory dys-synchrony:

The prevalence reports across the world varies largely and is yet not clear. The prevalence is reported to be more in infants in western countries whereas it is more observed in later stage of childhood and early adulthood in the Indian scenario. In the western population, the reported prevalence is approximately 1-3 infants per 10,000 births have AD (Dolphin, 2004). In school going children, 87 in 1000 children are having AD according to Berlin et al (2000). In the Indian context, the prevalence reported is around 1 in 183 individuals with sensorineural hearing loss (Kumar & Jayaram, 2006). They also reported a 2:1 female to male ratio for AD.

2.2 Pathophysiology of Auditory dys-synchrony:

Etiology of auditory dys-synchrony appears to be associated with variety of factors and are multiple. Owing to limited access to the auditory system and not as much of invasive studies, it is difficult to pin point the exact cause of auditory dys-synchrony. However, there are many theories postulated which tries to explain the cause based on the audiological profile of these individuals.

Some possible sites of lesion cited in research includes the cochlear inner hair cells, the synapse between the inner hair cells and type 1 auditory nerve fibers, and the auditory nerve itself (Starr et al., 1996; Rance et al., 1999). But, the most common pathophysiology for a cochlear afferent disorder, particularly in adults, is a neuropathy affecting the afferent nerve fibers. It can be caused by a primary demyelination or an axonal disease. A primary demyelination can cause slowing of conduction velocity whereas an axonal disease is characterised by normal conduction velocity and reduced amplitude of compound action potential. Other than these pathophysiology, a genetic dysfunction involving disruption of the otoferlin (OTOF) protein, can present with the auditory dys-synchrony result pattern (Varga et al., 2003).

These anatomical variations can significantly affect the functioning of the auditory system which in turn disturb the spectro-temporal representation of the signal and as a result, speech perception scores are reduced. Some patients can also have associated generalised neuropathy affecting other cranial nerves or peripheral nervous system (Starr et al., 1996).

2.3 Clinical findings:

Individuals with AD exhibit divergent audiological findings owing to vast etiological factors contributing to the signs and symptoms. This is more so with the pure tone audiometry and speech perception scores. Audiological profile of these individuals are given below.

2.3.1: Pure tone thresholds:

The pure tone thresholds are normally reported as fluctuating type of hearing loss that is obtained on a pure tone audiogram in individuals with auditory dys-synchrony. Sininger & Oba (2001) reported that individual with AD can have any degree of hearing loss (mild to profound) or can have normal hearing sensitivity. Starr et al. (2000) reported that the average hearing sensitivity ranged less than 35 dB HL in 31% of the ears with AD, a range between 35-70 dB HL in 39% of the ears and greater than 70 dB HL in 30% of the ears with AD. Individuals with AD most often report of problems in both ears (bilateral expression) and are usually affected in a symmetrical fashion (Sininger & Oba, 2001; Sininger & Starr, 2001; Kumar & Jayaram, 2006).

As with the thresholds, the configuration of hearing loss in individuals with AD are also found to be variable. The reported configurations in these individuals are usually peaked, rising, flat, saucer shaped and sloping types. However, it was noted that the peaked audiogram configuration was most predominant with peak at 2000Hz (Kumar & Jayaram, 2006). These findings were consistent to the results obtained in the retrospective study by Jijo and Yathiraj (2012). Also they found a significant association of duration of the problem did exist with audiometric configurations which tended to be flat if the duration of the pathology was above 10 years.

2.3.2: Speech audiometry:

The speech perception ability can range between no measurable speech identification scores to 100% scores in quiet situation. The speech scores generally do not correlate with the hearing thresholds obtained using pure tones (Zeng & Shannon, 1999; Kraus et al, 2000; Rance et al, 2004; Zeng et al, 2004). Kumar and Jayaram (2006) obtained a correlation between speech perception scores and pattern of hearing loss where the best speech identification scores were found in individuals with peaked audiograms. However, this correlation cannot be generalised owing to small sample size as reported by the authors.

A retrospective study by Jijo and Yathiraj (2012) aimed to see if any correlation exists between duration of the disorder to the audiological characteristics. But, both the pure tone thresholds and speech identification scores did not relate to duration of the disorder.

2.3.3: Physiological and electrophysiological test findings:

There are various physiological and electrophysiological tests carried out in individuals with AD to get a better understanding of the problem.

- The acoustic reflex testing usually show an absent response (Kraus, 2001; Sininger and Oba, 2001; Starr et al, 2001).
- Otoacoustic emissions are described to be by and large present in most cases with AD. Sininger & Oba (2001) reported that 80% of individuals with AD had good OAEs. However, over a course of time an absence of OAEs can be observed which can indicate progressive nature of the condition.
- The auditory brainstem response usually are absent even at maximum presentation level and a low rate of stimulation. These findings also

does not correlate with behavioural pure tone thresholds. ABR if present will exhibit poor morphology (Starr et al., 2000).On the other hand, a clear cochlear microphonics indicative of outer hair cell functioning can be observed in these individuals which is typically long ringing (Starr et al., 1996; Starr et al., 2000).

Consequently, obtaining a combination of these findings help to diagnose AD.

2.4: Psychophysical profile:

Along with elevation of thresholds, any damage to the inner ear and the auditory nerve leads abnormal loudness, pitch and temporal processing (Moore & Oxenham 1998). In the following section, brief information on the psychophysical abilities in individuals with AD are given.

2.4.1: Frequency discrimination:

The frequency discrimination ability studied in individuals with AD has shown significantly poor performance when compared to normal hearing counterparts, especially in the low frequency (Rance et al., 2004; Zeng et al., 2005). Zeng et al. (2005) compared the frequency discrimination in 12 AD individuals where they found a significantly poor frequency discrimination in low frequencies when compared to normal. But this difference was not observed in high frequencies.

In a study where just noticeable difference was measured, it was found that the AD subjects required more difference in the frequency to differentiate as two different stimuli. AD subject required 172 HZ and 235 HZ between the two tones for left and right ear respectively, in contrast to which normal hearing subject required 2 to 17 HZ for left and right ear respectively (Starr et al, 1996).

Zeng et al (2001) found that the AD subjects demonstrated non-monotic increment in the just noticeable difference (JND) with increase in frequency; with maximum effect at low frequencies and near normal at high frequency (8 KHz); whereas for normal it was found to be monotic growth with increase in frequency. Similar findings were observed by Rance et al (2004), where it was found that in AD individual difference limen at 4 KHz was 4.5 time higher than normal and at 500 HZ it was 11 time higher than the normal hearing cohort. Similarly when investigated with frequency modulated tone, it was found that children with AD were less able to use phase locking compared to normal hearing children (Zeng et al., 2005)

It can be concluded that subjects with AD had greater problem with frequency discrimination, which was most dominant in lower frequency range compared to high frequency. At higher frequencies, the response was comparable to normal hearing individual.

2.4.2. Intensity Processing

Initial study in intensity discrimination was done by Starr et al in 1991, in which he compared JND for 1000 Hz tone of 11 year old girl with dys-synchrony with five normal hearing children. The study concluded that subject with AD required approximately double the JND (10 dB) compared to normal (4 dB). Psychoacoustic intensity discrimination task was also carried out in two AD cases, where the client had to discriminate one different stimulus among the triad of stimuli. It was found that normal hearing required less than 1dB to discriminate, whereas the AD cases required 3 to 6 dB to discriminate (Starr et al, 1996).

Similarly Zeng et al (2001) investigated loudness growth using magnitude estimation and loudness scaling technique. The result indicated more compressive loudness growth function for AD individuals compared to normal subjects. But it was concluded that the intensity discrimination function was not significantly different from the normal hearing individual, as the intensity increases, the difference limen decreased.

Zeng et al (2005), studied intensity discrimination as a function of intensity from threshold to maximum comfortable level and found that AN individual requires higher SL below 40 dB. At higher intensity levels, there was no significant difference obtained when compared to normal hearing individuals.

So, it can be concluded that, though there is variability in the different studies, it can be noted that the AD individual requires higher intensity difference than the normal hearing individual.

2.4.3. Temporal Processing

Temporal processing has been assessed in individual with AD using various test such as temporal integration, gap detection, temporal modulation transfer function, temporal masking and simultaneous masking.

2.4.3.1. Duration discrimination

Starr et al (1991) carried out study with duration discrimination in one child with AD and compared with five normal children. It was found that at lower duration tone of 50 msec, normal children required about 20 msec whereas for AD subject they required 20-30 msec. For 500 msec duration stimulus, normal children required a mean of 140msec, whereas AD subject required 200 to 300 msec more. In continuation to the study, Starr et al (1996) conducted psycho-acoustical discrimination, where the subject were asked to indicate the different stimulus. The results showed that compared to normal individual, the AD children had poorer ΔT.

2.4.3.2. Temporal Gap Detection

Starr et al (1996) considered a choice of three stimulus which were presented and the client was asked to indicate the stimulus with different duration compared to other two. The detection threshold was more in cases with AD who obtained 6 and 12 msec for left and right ear, whereas normal hearing children required 1-5 msec.

Zeng & Shannon (1999) found that the AD cases had poor gap detection compared to normal hearing individual. AD group had 2 to 25 time higher detection threshold compared to normal hearing control that is, as the intensity level increased the detection threshold decreased. In another study by Zeng et al in 2005, it was found that normal hearing individual required 50 msec interval at threshold level, but improved to 3 msec at higher sensation level of 40 to 50 dB.

2.4.3.3 Temporal Integration

First measure of temporal integration on AD cases was done by Starr et al (1991), they found that there was increase in threshold as the duration increased from 30 to 300 msec, for both normal and AD cases. Similar finding was observed by Zeng et al (1999, 2001) where it was found that temporal integration was similar to normal hearing individual, and decreasing the duration below 30 msec, there was abrupt increase in the threshold up to 20dB.

2.4.3.4: Temporal masking:

Studies have reported excessive masking for pure tones in simulatneous and non-simultaneous conditions in individuals with AD (Kraus, 2000; Zeng et al, 2001 & Zeng et al, 2005). Along with these, abnormal masking level difference is also consistently noted in them (Starr et al., 1991, 1996). Kraus, 2000 reported exaggerated masking thresholds in an individual with AD who had near normal pure tone thresholds. The physiological basis for excessive masking can be due to inner hair cell loss or loss of synchronous firing in the auditory nerve (Harrison, 1998 & Starr et al, 1996).

These psychophysical findings using non speech stimuli help to understand the speech perception ability as there are studies which have correlated the same to see the perceptual consequence of impaired frequency and temporal information (Kumar & Jayaram, 2006; Narne & Vanaja, 2008).

2.5 Speech perception:

Speech perception abilities are found to be variable in AD which can range from being very good to very poor in quiet condition (Kumar & Jayaram, 2006; Narne & Vanaja, 2008). But, often most of the individuals with AD exhibit a poor speech perception.

There is a disproportionate relationship between the tone identification (pure tone thresholds) and speech identification abilities in persons with auditory dys-synchrony (Zeng et al., 2004). This relationship is one of the clinical finding which is an important feature that is considered for the diagnosis of the same. On the other hand, there is a consistent result about relationship between speech perception ability and their temporal processing abilities (Zeng & Shannon., 1999; Rance, Fava, Chong, Barker, Corben & Delatycki, 2008). It is a consistent finding that the temporal processing is affected in individuals with AD and also, temporal cues are helpful in perception of speech cues like the voicing and the manner. Hence, a poor speech identification score in this case is not due to audibility of the speech signal as observed in a peripheral loss (Zeng & Shannon., 1999; Zeng et al., 2005). It is more of an accessibility issue that is the ability to utilize the temporal (time based) cues which is distorted at suprathreshold level (Kumar & Jayaram 2010).

2.4.1 Vowel and consonant perception:

The researches have reported that identification of speech sounds are affected across all classes of sounds (consonants and vowels). Jerger and Hanley (1985) studied phoneme identification in persons with retro-cochlear pathology and found that the vowel errors were more in them however, the consonant errors were comparable to that of cochlear hearing loss.

Rance, Gary, Barker & Janet (2008) observed a reduced score for perception of stops, nasals, semivowels and diphthongs in persons with auditory dys-synchrony. Identification of the nasals were impaired as it required low frequency spectral cues which is known to be more affected in these individuals. Similarly, the next class of sounds like semivowels and diphthongs were also affected as it required to track the rapid spectral changes within the sound. When the identification of stops were analysed, it was found that discrimination of sounds like /p and b/ and /t and d/ were significantly affected. It can be reasoned out as the inability of these individuals to perceive the subtle timing differences (w.r.t the Voice onset time, burst duration). However, discrimination of high frequency pairs like /s and f/ and /z and v/ were more accurate when compared to other sounds as high frequencies tend to be less affected in individuals with AD.

Narne and Vanaja (2008) studied the perception of consonants in the context of vowel /a/. They observed that the 8 participants of their study exhibited misinterpretation of place and voicing than manner cues. There was greater error of place of articulation for stop consonants. Furthermore in their study, it was observed that perception of stops and liquids were harder when compared to fricatives and affricates. The authors attribute these findings to the impaired perception of spectro-temporal transitions for stop consonants which is due to poor temporal processing. Similar findings were obtained by Ramirez and Mann (2005) who compared individuals with dyslexia and individuals with AD. They also reported that the trend of error in both the population were similar.

Kumar and Jayaram (2011) studied the perception of stops and analysed the error pattern. Out of the place and voicing cue, they observed more errors in place rather than voicing. Also they found that voiced stops were misinterpreted as voiceless more than the other way round. Authors observed a marked difficulty in discriminating between speech sounds of differing spectral onset. They reason out this impairment to the difficulty in temporal processing. The transitions comprise rapid spectral changes occurring at the onset of a speech stimulus. Results from this study, together with those of Kraus et al., 2000 provide evidence that individuals with auditory dys-synchrony have difficulties in processing temporal and spectral information presented at the onset of speech stimuli.

Thus, from the studies conducted on consonants and vowels, it can be summarised that vowels are affected more in auditory dys-synchrony individuals when compared to other peripheral pathologies. With respect to consonants, the stops suffer more when compared to continuants. In addition to that, the feature that is least affected is reported to be the manner cue. Whereas, the place and voicing cues are affected more.

2.4.2 Speech perception in noise

A reduction in performance in speech perception abilities in an unfavourable environment is a common feature observed in all population. Auditory dys-synchrony is no exception to this and the speech perception worsens even more in the presence of noise in these individuals as synchrony is critical for understanding speech in the presence of noise. (Starr et al, 1996; Kraus et al, 2000; Shallop, 2002). The underlying pathophysiology for the same is not clear yet. But, there are reports of excessive masking (10 – 20 dB higher than for normal controls) in detection of pure tones in the presence of noise has been reported (Kraus et al, 2000; Zeng et al, 2005). Hence, there is a possibility of the noise making it difficult to extract the envelope of speech as noise can reduce the modulation depth and add spurious modulations to the speech signal (Drullman, 1995). This phenomenon may help to understand excessive degradation in speech intelligibility in the presence of background noise especially for listeners with AD. However the amount of deterioration depends on the signal to noise ratio. The studies on speech perception in noise in AD has shown that the scores are poorer in noisy conditions when compared to quite condition or favourable environment (Zeng & Shannon, 1999; Rance et al., 2004; Narne & Vanaja, 2009).

Narne and Vanaja (2009) evaluated speech identification scores in quiet and in the presence of noise (signal to noise ratio of 10, 5 and 0 dB). A greater amount of reduction in identification scores in noise was observed for listeners with AD than those with normal hearing. Also, identification scores in noise were much lower for those who had poor identification scores (<50%), than those who had good identification scores in quiet (>50%). Similar results have been demonstrated for children by Rance et al., (2007) and for adults by Zeng and Liu (2006).

A study by Narne (2013) observed the speech in noise abilities of individuals with AD and compared the same to the temporal processing abilities. The signal to noise ratios that were tested in this study were 0, 5 and 10 dBs for the bi-syllabic Kannada words. The speech perception scores had a trend of reduction with reduction in the signal to noise ratios. The amount of reduction was reported to be significantly more in individuals with AD when compared to the reduction observed in normal hearing counterparts. This reduction in speech identification scores were attributed to the impaired temporal perception along with inability to encode temporal fine structure by the individuals with AD.

Owing to the pathophysiology of AD and the perceptual consequence that are exhibited by them, many management options were tried in this population. At present, the management options that are gaining importance for AD are cochlear implantation and spectro-temporal enhancement strategies which have taken the place of conventional hearing aids or ALDs.

2.5 Assistive hearing devices and implantable devices:

Conventional management strategies like amplification devices have been tried on individuals with AD. The results yield no significant improvement especially in speech intelligibility (Berlin, Hood, Morlet, Rose, & Brashears, 2003). Other management options like cochlear implants, FM systems, speech reading and signed language have been tried for its usefulness in individuals with auditory dys-synchrony (Kraus, 2001).

The use of sign language, ALDs were considered the management option earlier as the AD individuals had difficulty with conventional hearing aids (Widen, Ferraro & Trouba, 1995). The use of FM systems, speech reading or vibrotactile aids along with hearing aids are emphasized thus to overcome the drawback of hearing aid alone (Doyle, Sininger & Starr, 1998). Also, in patients where there are chance of recovery, usage of hearing aid can be harmful which might destroy the outer hair cell resulting in permanent sensorineural hearing loss. In such individuals, visual language exposure or sign language is recommended (Berlin, Hood, Morlet, Rose & Brashears, 2003). Also, there are studies which report that child learning spoken language along with signed language, the prognosis usually is less (Ketelaar, Rieffe, Wiefferink & Frijns 2012). Hence, a combination of visual and auditory mode for communication might not be a best option for management especially for children.

Keeping these drawbacks, the use of cochlear implantation was considered as another viable option. Sininger and Trautwein (2000) have described a case with normal electrical ABR findings following cochlear implantation surgery. Fabry (2000) reported on a patient with improved auditory function after receiving a cochlear implantation. Also, studies report that the outcome can be compared with the peers who has undergone cochlear implantation with pathology of cochlear origin (Budenz & Starr 2013). One of the studies by Sininger &

Oba (2000) compared the preoperative and postoperative ling sound perception in cases with auditory neuropathy and cases with hearing loss of cochlear origin. The results show that there was an improvement in the perception in individual with auditory dys-synchrony. Cochlear implant thus may be a useful option, however, the recent studies indicate that not all children with auditory dys-synchrony show benefit from it (Gibson, & Sanli, 2007; Range- Samuelson, Drake, & Wackym, 2008). These discrepancies are mainly because of the heterogeneous symptoms they exhibit from the etiology to the pathophysiology.

Thus, the current trend has now shifted to enhancement of envelop in terms of spectro-temporal parameters of speech to see if there will be an improvement in speech intelligibility. Some of the modification strategies that are tried as option for management of AD are clear speech, consonant or burst enhancement, transition duration enhancement and companding.

2.5.1: Signal enhancement techniques:

Clear speech is one of the enhancement strategies studied in individuals with learning disability, APD and cochlear implants. It has obtained the importance in managing AD individuals as the enhanced temporal properties can be beneficial in them as well (Zeng & Liu, 2006). Clear speech is different when compared to conversational speech in terms of reduced speaking rate, increased energy in the 1000–3000Hz range, enhanced temporal modulations, expanded voice pitch range and vowel space (Krause & Braida, 2002, 2004; Liu, Del Rio, Bradlow, & Zeng, 2004 & Ferguson & Kewley-Port, 2002). Zeng & Liu (2006) studied the speech perception in individuals with AD with clear speech. They considered 13 participants, out of which 7 had received cochlear implantation. Sentence perception was assessed in quiet and in the presence of noise (0, 5, 10 and 15dB SNR). They observed that participants with AD performed more poorly in speech recognition in noise than did the normal-hearing, cochlear-impaired, and cochlear implant controls. However, a significant clear speech advantage was observed, ranging from 9 to 23 percentage points in intelligibility for all listening conditions. Electric stimulation via a cochlear implant produced significantly higher intelligibility than acoustic stimulation in both quiet and in noise. Thus, the authors concluded that clear speech is a great strategy which enhances the temporal information which is helpful for population with AD. Although in this present study stimulation through electrical mode is emphasized, similar improvement with clear speech in auditory mode was also observed i.e. 16 percentage points improvement when compared to conversational speech.

Speech perception with lengthened transition duration was also tried on individuals with AD. A study by Kumar and Jayaram (2011) in 30 individuals with AD saw the effect of varying transition duration on perception of consonant-vowel syllables. The authors presented the CVs with enhanced transition duration and observed an improvement in perception both with respect to placement and voicing feature. However the mechanism underlying this improvement remains unclear.

Another strategy of envelop enhancement was studied by Narne and Vanaja (2008). They considered 8 participants with AD and assessed their consonant identification ability by enhancement of envelop. The magnitude of envelope enhancement was 15 dB and the bandwidths used were 3 to 10 Hz, 3 to 20 Hz, 3 to 30 Hz, and 3 to 60 Hz. In all these conditions, there was an improvement observed when compared to unprocessed condition. However, the improvement was greater for larger bandwidths (3 to 30 Hz) when compared to other 2 bandwidths. In the envelope-enhanced conditions, cues for manner and place of articulation were transmitted better than voicing. It was also observed that this difference was mainly due to improvement in identification of plosives and liquids in the broader bandwidth conditions. The authors reason that enhancing the envelope in broader-band width enhances the more salient cues such as formant transition and burst which in turn probably improved identification of consonants in their study.

The hallmark feature of auditory dys-synchrony is impaired temporal perception. Any management strategy that modified the temporal aspects of speech like burst duration, voice onset time or the transition time has shown to have some improvement in speech perception (Narne & Vanaja, 2009; Kumar & Jayaram 2011). Hence, any strategy that taps on the spectral or temporal changes in speech can bring about a change in perception of speech sounds. One such strategy is companding which refers to a technique for compressing and then expanding (or decompressing) of a signal.

Companding is a combination of the process "compressing" and "expanding." It is a well proposed technique which is developed based on relatively broadband compression followed by more frequency selective expansion. Turicchia and Sarpeshkar (2005) proposed this strategy for time domain spectral enhancement. This approach tries to mimic certain properties shared by the peripheral auditory system. Then, it is expanded at the receiving end using the same non-linear scale to restore it to its original form, but, the noise level would be reduced.

This strategy is used in cochlear implant subjects by various investigators who have found an improvement with respect to increased speech intelligibility (Oxenham, Simonson, Turicchia, & Sarpeshkar, 2007; Bhattacharya & Zeng, 2007). They observed that companding improved speech perception scores by 10 – 20% in steady-state noise. This technique enhances the spectral peaks in a stimulus, relative to nearby spectral valleys. It is observed that the spectral enhancement techniques mainly modify the regions of the signal that contain acoustic cues to consonant identity which helps individuals to perceive the speech better.

The effect of companding was studied in individuals with AD at different SNR (Shachi, 2010). They carried out speech perception testing (consonant identification, SNR 50 and Quick SIN) in 15 ears with AD. They obtained a significantly better performance with respect to speech identification with companding. The authors attribute the improvement to the increment in spectral and temporal contrast in the companded stimuli.

The same technology was used to study the speech perception in individuals with AD (Narne et al., 2014). They considered 10 participants with AD to check the perception of consonants and sentences in quiet and different SNRs (0, 5, 10 and 15 dB). There was a significant improvement observed in terms of increment in consonant identification scores in individuals with AD especially in quiet and 15dB SNR. However, there was no scores obtained in very low SNR (0dB) for both unprocessed and companded stimuli.

Another technology is burst enhancement strategy which incorporates enhancement of burst portion of the consonant modifying the consonant to vowel ratio. Guelke (1987) suggested that, increasing the amplitude of the burst of a stop consonant up to the level of the vowel in normal speech could result in increased intelligibility by perceiving information that would otherwise be inaudible to a person with hearing impairment. The increase in intensity following the gap would also, by itself, indicate the presence of a stop consonant. They found that the technique aided in improving the perception of speech.

A study by Hazan, Simpson and Huckvale in 1998 modified the gain of the vowels, frication and aspiration differentially and examined speaker and listener effects with native and non-native speakers using CVCs in only quiet conditions. Their results showed a significant improvement in processed conditions for all speakers and an increase in scores in the listeners. In other peripheral loss cases also, the enhancement has found to be beneficial (Simpson, Moore & Glasberg, 1990; Summerfield, Foster, Tyler & Bailey, 1985) Sumerfield et al, 1985 examined spectral enhancement by narrowing bandwidths and also broadening their bandwidths to reduce the contrasts at word level. When the bandwidths were broadened, the scores reduced in the group with cochlear hearing loss. Whereas when the formant bandwidth were reduced to half of the actual values, the perception improved. Following these results, the authors concluded saying that this technology can also be beneficial for individuals who exhibit temporal resolution deficit and excessive susceptibility to backward masking.

From these review, it is clear that any enhancement in the spectral or temporal domain can benefit the individuals with AD to perceive speech better. (Kumar & Jayaram, 2006; Narne & Vanaja, 2008).

In the earlier studies where companding is used, the stimulus considered are either VCV combinations or words. Hence there is a need to study the effect of companding on CV syllables to see if there would be an improvement if the consonants are in the initial position (which is considered to be more difficult to perceive, Sumerfield et al, 1985).

Studies where consonant enhancement for AD individuals are tried, either have enhanced with respect to bandwidth or duration of burst or transition. These are done owing to the perceptual consequence of the pathophysiology of AD where temporal perception is the severely affected. However, there is a need to see if a spectrally enhanced stimuli (amplitude of consonant enhanced) can bring about any change in the perception. Companding makes use of the concept of two tone suppression and alters both spectral and temporal aspects of speech whereas consonant enhancement aims to increase the spectral contrast. Thus, making the speech more accessible to individuals with AD.

The effect of noise in speech perception are also studied extensively in population with AD. The effect of noise is reported to be more detrimental in these individuals. Keeping these in mind, the study incorporated 4 different SNRs along with quiet condition to see if the techniques can aid in improvement of speech intelligibility in the difficult to listen environment as well.

Hence, a systematic comparison of both strategies under similar setting would give an appropriate idea to determine the best spectro-temporal strategy to incorporate to bring about better identification of speech both in quiet and noisy situations. This study was thus designed keeping the main objective to find the best possible modification of the spectral or temporal envelop of speech which would aid the speech intelligibility in individuals with AD.

Chapter 3

Method

This study was conducted with an aim to compare the effect of spectro-temporal enhancement of syllables using companding and consonant enhancement on speech perception in different listening environments including quiet and noisy situation. To achieve the aim of the study, two groups of participants were considered. Following procedure was administered to collect the data.

3.1: Participants:

Two groups of individuals, the control and the clinical group participated in the study. Both the groups had 10 native Kannada speakers in the age range of 14 to 40 years (mean age: 21.6years). This age range was taken as there is literature evidence that the psycho- acoustic abilities reach a plateau in normal hearing individuals by 12 years of age (Werner & Gray, 1998). The control group had participants with hearing sensitivity within normal limits and the clinical group included participants with auditory neuropathy.

3.1.1: Clinical group: Individuals with Auditory dys-synchrony:

This group consisted of 10 participants diagnosed as having auditory dys-synchrony at the Department of Audiology, All India Institute of Speech and Hearing. All of them reported to have acquired post lingual hearing loss. The hearing loss in terms of pure tone thresholds ranged from near normal hearing to moderate sensorineural hearing loss and the speech identification scores were disproportionate to the degree of hearing sensitivity. All the participants had no known speech and language deficits or any other associated neurological symptoms as reported. The Audiological details along with the demographic data are given in the Table 3.1

Table3.1: Demographic and Audiological details of subjects with auditory dys-synchrony

Subjects (Test ear)	Age/Gender	Pure tone average	SIS in quiet	Tympanometry	Acoustic reflexes	OAE	ABR
S1 (Left)	29y/M	33.75	32%	'A'	Absent	Present	Absent
S2 (Left)	21y/M	53.75	32%	'A'	Absent	Present	Absent
S3 (Right)	14y/F	34.75	68%	'A'	Absent	Present	Absent
S4 (Right)	15y/F	38.75	64%	'A'	Absent	Present	Absent
S5 (Right)	35y/M	27.5	60%	'A'	Absent	Present	Absent
S6 (Right)	18y/F	32.5	68%	'A'	Absent	Present	Absent
S7 (Left)	22y/M	45	52%	'A'	Absent	Present	Absent
S8 (Right)	18y/M	20	64%	'A'	Absent	Present	Absent
S9 (Left)	29y/F	32.5	52%	'A'	Absent	Present	Absent
S10 (Left)	15y/F	34.75	76%	'A'	Absent	Present	Absent

3.1.2: Control group: Individuals with hearing sensitivity within normal limits:

This group consisted of 10 adults who were age and gender matched with the clinical group. Their hearing sensitivity was within normal limits (four frequency average pure tone threshold, 500 Hz, 1000 Hz, 2000 Hz & 4000 Hz) and speech identification scores in noise was 60% or above at 0dB SNR. Transient evoked otoacoustic emissions were present in all the participants. 'A' type tympanogram with middle-ear acoustic reflexes (both ipsilateral & contralateral) and auditory brainstem responses were present for all the participants.

3.2: Instrumentation:

The following instruments were used in the study,

- Pure tone thresholds and speech testing was carried out using a calibrated dual channel diagnostic audiometer, GSI-61. TDH- 50P supra aural headphones and 71 bone vibrator were used to present the stimulus.

- Immittance testing was done using a calibrated GSI-tympstar (Grason-Stadler Incorporation, USA) clinical immittance meter.

- Transient evoked otoacoustic emissions was recorded using ILO 292 DPEcho port system (Otodynamics Inc., UK).

- Auditory brainstem responses were recorded using Intelligent Hearing Systems (IHS smart EP windows USB version 3.91) with AgCl electrodes and ER-3A insert earphones.

- MOTU MicroBook II with AHUJA AUD- 101XLR dynamic unidirectional microphone was used for recording the stimulus.

- MATLAB- 7 (Language of Technical computing, USA) was used to generate signal, mix the generated signal with noise and process the same for temporo-spectral modification.

- UCL enhance version 101.exe (2002) was used to enhance (gain in dB) the consonant portion of the syllable.

- Adobe Audition v5 was used to normalise the recorded CV syllables.

- The stimulus was presented through Dell Inspiron 14R laptop (Realtek sound card).
- The stimulus from the laptop was routed through a calibrated audiometer (MA-53). In order to control the intensity of the stimulus. Senheiser HD 200 headphones was connected to the audiometer to deliver the stimulus.

3.3: Stimulus generation:

A native Kannada speaker was used to record nineteen consonant-vowel combinations (CV syllables). Each syllable was recorded thrice digit. Non-sense syllables were considered as it reduced the redundancy of the stimuli which would affect the perception scores. The consonants (/p/, /b/, /d/, /g/, /tʃ/, /dz/, /r/, /v/, /n/, /m/, /l/, /l̲/, /j/, /k/, /d̲/, /s/, /ʃ/, /t/, /d̲/) were recorded in the vowel context /a/. All the recorded syllables were subjected to goodness test using a three point rating scale (3-good, 2-fair, 1-poor). This was done by asking five Kannada speakers with normal hearing sensitivity to listen and rate the quality of syllable. The syllables having maximum score were selected for the study. The selected signal was then mixed with six speaker speech babble developed by Jain, Konadath, Vimal, and Suresh (2014) to achieve signal to noise ratios of 0, +5, +10 and +15 dB. This was done using MATLAB- 7.8. Speech babble was chosen as the masker due to its resemblance to the natural environment, and also considering that people with hearing loss more often complain about understanding speech of a particular person in the presence of multiple speakers. The 19 syllables taken were temporally placed in the centre of the speech babble. To carry out testing in quiet condition, syllables without the babble were also retained. The mixed and the unmixed stimuli were subjected to companding and consonant enhancement processing. Companding and consonant enhancement were done using MATLAB 7.8

and UCL Enhance software respectively. The modifications were applied after adding the babble to the syllables to make it more natural as the speech reaching the ear of the listener would already be embedded in the noise present in the environment. Keeping this reason into consideration, the present experiment was designed to test the efficacy of enhancing the signals at this stage.

The speech stimuli mixed with 4 SNRs and quiet conditions were obtained finally. The syllables retained without processing were labelled as ‘unprocessed’ and whereas the stimuli to which companding was applied was labelled as ‘companded’ and the stimuli to which enhancement was applied was labelled as ‘enhanced’.

3.4: Procedure used for Companding:

Companding strategy was employed to spectrally enhance the syllables which was based on the algorithm provided by Bhattacharya & Zeng, (2007). A series of compression and expansion using different filters were involved in this process. This was based on the two-tone suppression which is one of the normal physiological processes observed at the level of cochlea. The incoming signal was first divided into 50 frequency channels by a bank of relatively broad band-pass filters. The signal within each channel was then subjected to amplitude compression. The amount of compression was dependent on the output of the envelope detector (ED) and the compression index (n1) which had a value of 0.3. The compressed signal was then passed through a relatively narrow band-pass filter before being expanded. The gain of the expansion block is dependent on the corresponding ED output and the ratio of (n 2- n 1)/ n 1. The n2 parameter of the algorithm is the expansion index and had a value of 1. Output from all the channels were then summed to obtain the processed signal. This combination of adding non-linearity and filtering in the algorithm enhances the spectral contrast by giving more importance to the spectral energies. The output after processing was normalised with RMS amplitude normalisation for -15dB in Adobe Audition software which was then equated to that of the original signals. Figure 3.1 and 3.2 explains the companding procedure used in the study.

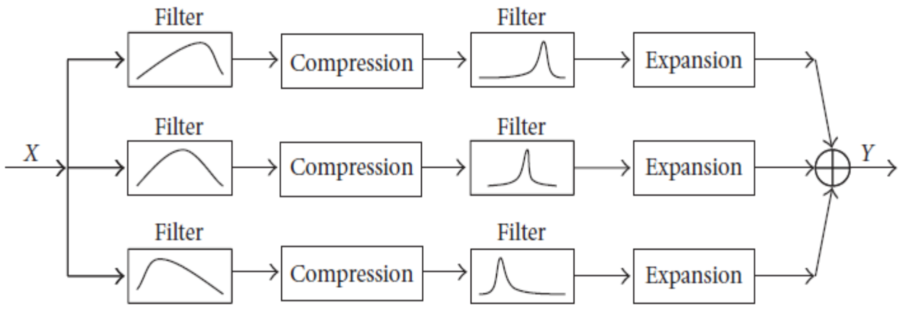


Figure 3.1: Block diagram of the companding architecture showing the stimulus being analysed by a bank of broad band prefilters. The output of each prefilter was subjected to compression, and the output was filtered again using sharper postfilters before it was subjected to expansion. The outputs from all the channels were then summed to obtain the processed signal. Extracted from “A Bio-Inspire Companding Strategy for Spectral Enhancement”, by L. Turicchia and R. Sarpeshkar, 2005, *IEEE Transactions on Speech and Audio Processing, 13,* p. 244.

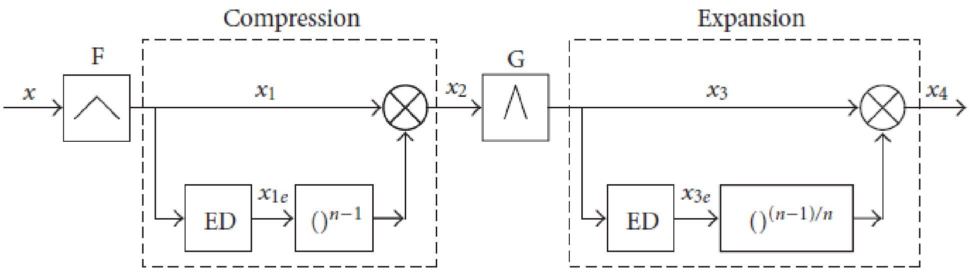


Figure 3.2: Detailed architecture of a single channel ED- envelope detector. Extracted from “A Bio-Inspire Companding Strategy for Spectral Enhancement”, by L. Turicchia and R. Sarpeshkar, 2005, *IEEE Transactions on Speech and Audio Processing*, *13*, p. 244.

3.5: Procedure used for Consonant enhancement:

Consonant enhancement technique was employed with the help of UCL enhance software to increase the amplitude of burst and transition portion of the syllables. The method followed was similar to the method previously used by Guelke (1987). The procedure involved an algorithm where the location of different parts of the syllable i.e., vowels, nasals, fricatives and gaps were automatically identified based on broad class phonetic recognition system. The algorithm then increased the amplitude of the selected portion of the syllable up to 6dB. The options selected for the enhancement of the stimuli are given in Table 3.2.

Table 3.2: Options chosen for enhancement for the syllables

Syllables	Options chosen (among Burst, Fricative, Nasal and transition)	Enhancement level
Stops (p, b, t, t̲, d, d̲, k, g)	Burst + Transition	6 dB
Fricatives and Affricates (tʃ, dz, s, ʃ)	Fricative + Transition	6 dB
Nasals (m, n)	Nasal + Transition	6 dB
Glides (j, r, l, l̲, v)	Transition	6 dB

RMS amplitude gain was given for all the syllables, along with the amplitude compression degree of 10. This was recommended by the software as it would help maintain an overall average of the non-silent portions of the signal which would not vary with additions of gaps due to variables like noise. This option was combined with amplitude compression to make sure any change in the intelligibility can be attributed to enhancement and not due to a general increase in signal to noise. Figure 3.3 gives an idea about how the different stimulus condition change the spectro-temporal features.

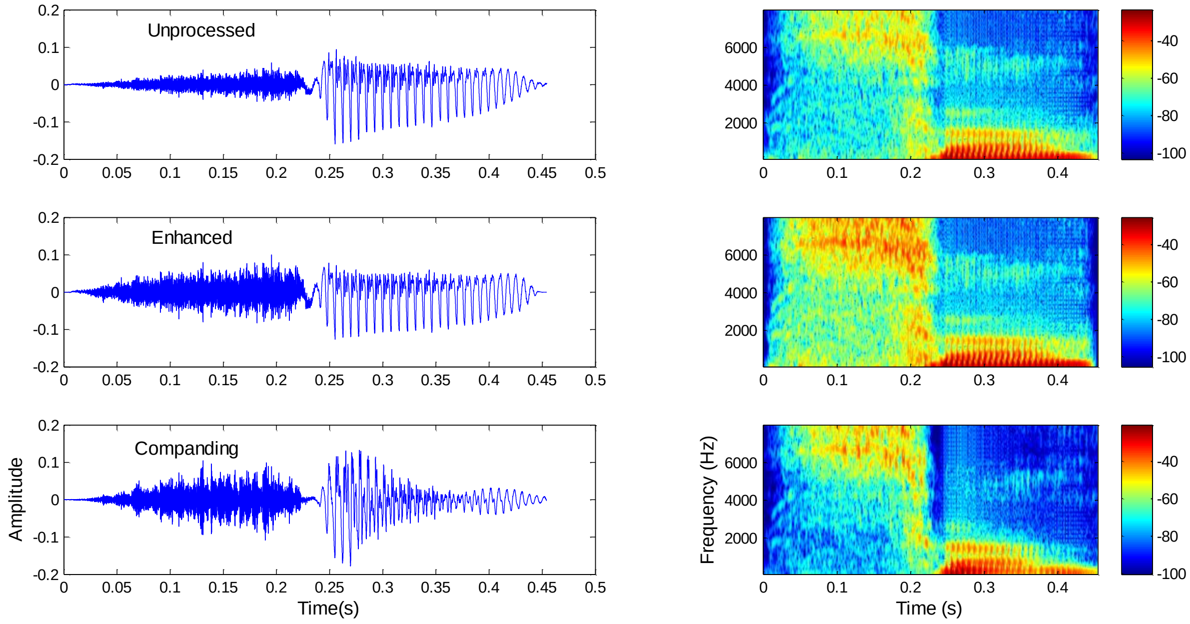


Figure3.3: Figure showing the spectrum and spectrogram of the syllable /sa/ in all the three conditions without noise.

3.6: Testing environment:

All the tests were carried out in an air conditioned, double room situation with ambient noise levels within permissible limits (ANSI S-3, 1991).

3.7: Procedure:

- A detailed case history: before the routine audiological assessment a case history was taken for all the participants in order to ensure that they do not report of any symptoms that would exclude them based on the subject selection criteria as described earlier.
- Pure-tone thresholds: modified version of Hughson and Westlake procedure was used (Carhart & Jerger, 1959) to obtain the thresholds. Hearing sensitivity was assured at octave frequencies between 250 Hz to 8000 Hz and 250 Hz to 4000 Hz for air conduction and bone conduction respectively.

- Speech audiometry: Speech recognition scores and speech identification scores were obtained using spondee word list given by Vandana (1998) and PB word list given by Yathiraj and Vijayalakshmi (2005) respectively. Using the same PB word list, speech in noise (SPIN) scores were also obtained. Uncomfortable level for running speech was also obtained.
- Immittance audiometry: using a probe tone frequency of 226 Hz immittance testing was performed by varying pressure from +200 dapa to -400 dapa. Ipsilateral and contralateral stapedial acoustic reflexes thresholds were measured for 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz pure tones with 226Hz probe tone frequency.
- Otoacoustic emissions: were obtained for 260 nonlinear click stimuli. OAEs were considered as present for the criteria in which the amplitude of OAEs were needed to be 6 dB SPL more than the noise floor in at least 3 consecutive octave frequency bands in both ears. The reproducibility criteria of greater than 50% was considered for the presence of OAEs (Wagner & Montgomery, 1975).
- Auditory Brainstem Responses: two recordings of the responses from all the participants were obtained using a repetition rate of 11.1/sec for 100usec click stimulus at 90dBnHL. The cut off frequency of 100 Hz was kept for high pass and 3000 Hz for low pass filtering. The polarity used was rarefaction.

3.8: Procedure used to obtain data:

To obtain the consonant identification scores, all the participants of both groups were made to undergo the following testing procedure.

- The most comfortable level for speech was obtained for each participant before the commencement of the actual testing. The further testing was carried out at that particular level (MCL).

- The participants were familiarised to the stimuli before the testing phase by writing the stimuli considered in their native language.
- The stimuli were presented through calibrated headphones attached to the laptop interface whose output was calibrated beforehand. MATLAB- 7.8 software was used for stimulus presentation. Every syllable was presented thrice under each stimulus condition. To start with, stimuli in quiet condition was presented and then moving to complex situation consisting of different SNRs.

The participants were asked to identify the syllable heard and that was selected as the response by the examiner on the response screen. The stimulus presentation was randomised by the software for every trial and thus making it blind-folded for the tester to avoid tester bias. The response screen is as shown in the figure 3.4

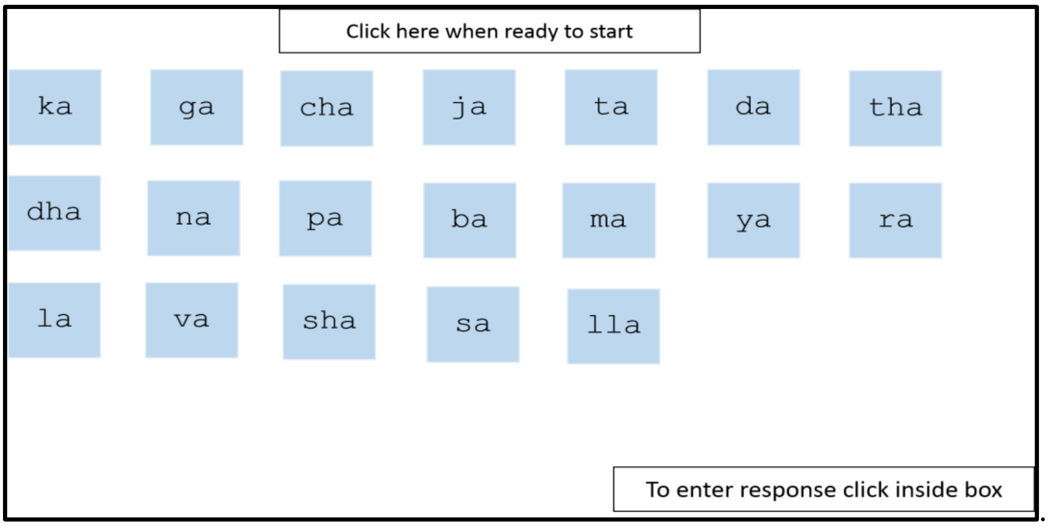


Figure 3.4: Response screen showing the arrangement of syllables considered for testing

- The next stimulus was presented following the response of the prior stimuli. A two minute interval was given after completion of every SNR condition and the testing was carried out in two sessions to avoid fatigue.

- A score of 1 was awarded for the correct response and 0 for incorrect response. No score was awarded for partial correct response to the stimulus. Using the response, a stimulus response matrix was constructed for each stimulus condition separately at each signal to noise ratio. In the matrix, the sum of scores of diagonal axis gave the total correct score for every condition.
- The obtained scores were analysed in every stimulus condition across all the SNRs considered for both the groups of individuals. On visual inspection of the consonant identification scores in clinical population, there was a lot of variability that was noted. Owing to this reason, the clinical group was further subcategorized based on their performance in syllable identification in quiet condition as good and poor performers. The criteria of 8 correct responses was set for this categorization. Out of the 10 individuals with AD, 5 individuals obtained scores greater or equal to 8 who formed the good performers group (ADG) and remaining 5 formed the poor performers group (ADP).
- The consonant identification scores were later compared across different SNRs in each condition separately and also across conditions in order to achieve the objectives of the study.

Chapter 4

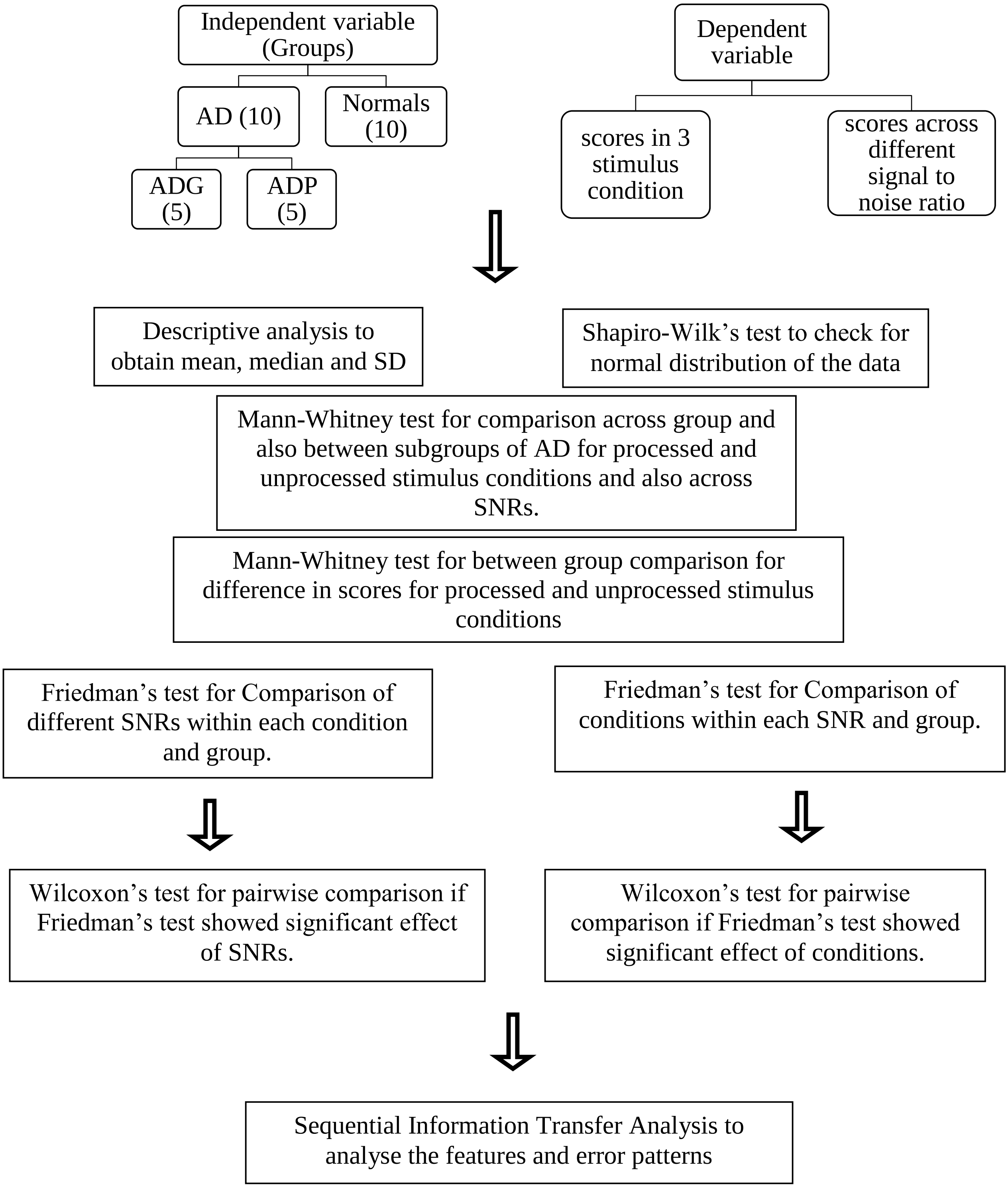
Results

The focus of the current study was to examine the performance of normal hearing individuals and individuals with auditory dys-synchrony for unprocessed, companded and consonant enhanced speech stimuli. To investigate the same, consonant identification scores (19 CV syllables) were obtained from all participants in two groups of individuals (10 individuals with normal hearing and 10 individuals with auditory dys-synchrony) in five different listening conditions.

The performance of individuals with auditory dys-synchrony was compared with individuals with normal hearing sensitivity across the three spectro-temporal conditions and 5 SNR conditions. Following the across group comparison, the effect of SNRs and stimuli conditions were individually analysed within group to arrive at the objectives of the study. The subgroups of AD (ADG and ADP) were also analyzed across the group and within the group to compare the effect of the stimulus condition across all the SNR. The analysis of these was carried out using Statistical Package for Social Sciences software (version 20).

The analysis was started by running Shapiro-Wilk's test of normality. The test revealed that the distribution of all data did not fulfil the normality criteria. Adding on, the descriptive statistics showed that the standard deviation is large in the clinical group. Due to these two reasons, non-parametric tests were chosen to analyse the data.

The following is the illustration of the statistical analysis carried out for the data obtained across different SNRs, stimulus conditions and groups and subgroups.


Independent variable (Groups)
AD (10)
Normals (10)
ADG (5)
ADP (5)
Dependent variable
scores in 3 stimulus condition
scores across different signal to noise ratio
Descriptive analysis to obtain mean, median and SD
Shapiro-Wilk's test to check for normal distribution of the data
Mann-Whitney test for comparison across group and also between subgroups of AD for processed and unprocessed stimulus conditions and also across SNRs.
Mann-Whitney test for between group comparison for difference in scores for processed and unprocessed stimulus conditions
Friedman's test for Comparison of different SNRs within each condition and group.
Friedman's test for Comparison of conditions within each SNR and group.
Wilcoxon's test for pairwise comparison if Friedman's test showed significant effect of SNRs.
Wilcoxon's test for pairwise comparison if Friedman's test showed significant effect of conditions.
Sequential Information Transfer Analysis to analyse the features and error patterns

The results of these statistical analysis are discussed under the following headings:

1. Comparison of consonant identification scores obtained across groups (AD v/s Normal)
2. Comparison between groups for benefit received by each stimulus condition across SNR
3. Comparison of consonant identification scores obtained across SNR within each condition and within the group
4. Comparison of consonant identification scores obtained across stimulus condition at each SNR and within group
5. Comparison between sub-groups of AD

4.1: Comparison of consonant identification scores obtained across groups (AD v/s Normal):

Data obtained from both the groups in all three stimulus conditions and 5 different SNRs was subjected to descriptive statistics to obtain the mean, median and standard deviation (SD). The same are tabulated in the Table 4.1.

Table 4.1 Mean, standard deviation and median for speech identification scores for 3 conditions at 5 different signal to noise ratio of both groups

Stimulus condition		**Unprocessed (across SNRs in dB)**					**Companding (across SNRs)**					**Consonant Enhancement (across SNRs)**				
Population		**Quiet**	**15**	**10**	**5**	**0**	**Quiet**	**15**	**10**	**5**	**0**	**Quiet**	**15**	**10**	**5**	**0**
AD	**Mean**	6.7	3.1	3	2.3	1.7	8.6	3.5	2.9	2.1	2	7.5	3	2.4	2.3	2.3
	SD	2.95	1.73	1.56	2.16	2.06	3.27	1.43	1.29	1.2	1.94	2.37	1.41	1.27	1.83	2.58
	Median	7.5	2.5	2.5	1	1	9.5	3	3	2	1.5	7	2.5	2	2	1.5
Normal	**Mean**	18	18.4	18	17.6	13.5	17.5	18.5	18.3	17.2	12.2	17.5	18	18	16.8	14
	SD	2.21	1.27	1.41	1.08	2.55	2.51	0.85	1.16	1.55	2.74	2.17	1.06	1.25	1.03	2.16
	Median	19	19	18.50	17	14	18.5	19	19	17.5	12.5	18	19	18	17	14

From visual inspection of the descriptive data from the Table 4.1, it can be said that the individuals with AD had a lower consonant identification scores when compared to normal hearing counterparts. However, in both the population there was a common trend of reduction in scores on an average of almost 50% correct score in quiet to 5.2% correct scores at 0 dB SNR in AD and average correct score of 100% to 73% in normal hearing group across all the stimulus conditions.

Mann-Whitney U test was carried out to compare the scores obtained across stimulus condition and SNRs between the two groups. The results revealed that the AD group differed significantly from the normal group at each stimulus condition and SNR. The results of this test are given in the Table 4.2 for all the conditions at different signal to noise ratio.

Table 4.2 |Z| values obtained for consonant identification scores between AD and normal group at each SNR and condition

Stimulus Condition	SNR (dB)				
	0	**5**	**10**	**15**	**Quiet**
Unprocessed	3.82	3.86	3.83	3.93	3.87
Companding	3.81	3.81	3.89	3.88	3.66
Enhancement	3.81	3.86	3.84	3.85	3.8

Note: p<0.001, 2-tailed

The Mann Whitney U test showed a significant difference between the groups for each stimulus condition and each SNR. However, it did not indicate the effect of stimulus condition on consonant identification scores at each SNR and also effect of SNRs on consonant identification scores at each condition. It is also required to know whether condition or SNR effect the consonant identification scores for each group. Thus, further statistical analysis was required which is given later.

4.2: Comparison between groups for benefit received by each stimulus condition across SNR:

To assess the benefit received by each processing condition, the difference between processed and unprocessed consonant identification scores obtained across all the SNRs were calculated that is, consonant identification scores obtained in companding/consonant enhancement – unprocessed condition. The mean difference data of the same is given in the Table 4.3.

Table 4.3 Mean, median and standard deviation of the difference between consonant identification scores between processed and unprocessed stimulus conditions across different SNRs for both groups

		Companding - Unprocessed					**Enhancement - Unprocessed**				
SNR		0 dB	5 dB	10 dB	15 dB	Quiet	0 dB	5 dB	10 dB	15 dB	Quiet
AD	Mean	0.2	-0.1	-1.3	0.7	1.9	0.5	-0.1	-0.5	0	0.8
	Median	0	0	-0.5	0.5	1.5	0.5	0.5	0	0	1
	SD	1.55	1.1	1.9	2.11	1.37	1.35	1.6	1.51	1.56	1.62
Normal	Mean	-1.3	-0.4	0.3	0.1	-0.5	0.5	-0.8	0.00	-0.1	-0.5
	Median	-2	0	0	0	0	0.5	-1	0	0	0
	SD	3.8	1.27	0.82	0.57	1.58	1.43	1.4	1.05	0.57	0.71

Note: C= companding, E= Enhancement, U= Unprocessed

From the table it can be noted that there is no specific trend obtained for both the groups. However, AD group showed a benefit at 15dB SNR and quiet for companding and a benefit at quiet for consonant enhancement. The negative values in the Table indicates that the scores were better in the unprocessed condition when compared to processed condition. Mann Whitney test for the difference in scores were carried out. The results of the same is given in the Table 4.4.

Table 4.4 |Z| values obtained for difference scores between the groups across SNRs

Difference between stimulus condition	SNR				
	0	**5**	**10**	**15**	**Quiet**
Companding-Unprocessed	1.5	0.93	2.45*	1.41	3.69*
Consonant enhancement-unprocessed	0.78	1.16	0.66	0.24	1.92^

Note: * indicates $p<0.05$, ^ indicates $p=0.05$

From the Table 4.4 it can be observed that there was a significant difference between the scores difference of companding and unprocessed in the quiet and 10 dB SNR between the groups. That is, the amount of improvement that AD group utilised was significantly greater than that of normal hearing peers. Also, there was a marginal significant difference for the difference of consonant enhancement and unprocessed condition in quiet condition where the AD group had a greater benefit than normal group.

4.3: Comparison of consonant identification scores obtained across SNR within each condition and within group:

The non-parametric Friedman's test was carried out under each stimulus condition where the different SNRs acted as the repeated measure in each group. The analysis were done separately for each stimulus condition and each group. The χ^2 values were obtained under each condition across different SNRs. The test results showed a significant effect of SNR on consonant identification scores for each stimulus conditions. This trend was same in both the groups of population. The χ^2 values for both groups at each stimulus condition are reported in the Table 4.5.

Table 4.5 ǀZǀ values obtained for difference scores between the groups across SNRs

Groups	**Stimulus Conditions**	χ^2 **(4) For AD**	***p*-value**
AD	**Unprocessed**	19	0.001
	Companding	25.72	0.00
	Enhancement	23.89	0.00
Normal	**Unprocessed**	18.71	0.001
	Companding	25.16	0.00
	Enhancement	22.29	0.00

From the Table, it can be seen that SNR affected the consonant identification scores significantly at all the conditions in both the group. Hence, a pairwise comparison was performed to see which pair of SNR is bringing about a significant difference at each condition. This was carried out using Wilcoxon's signed rank test. The same test was repeated in both clinical and control groups separately. The Wilcoxon's signed rank test results are tabulated in Table 4.6 for each condition.

Table 4.6 Pairwise comparison of consonant identification scores between SNRs at each stimulus condition and each group
Note: shaded area in the Table represents a significant difference for the particular pair of SNR under each stimulus condition and group

Group	Conditions	Quiet v/s15	10 v/s quiet	5v/s quiet	0 v/s quiet	15 v/s 10	15 v/s 5	15 v/s 0	10v/s5	10v/s0	5 v/s0
AD	**Unprocessed**	$p<0.05$	$p<0.05$	$p<0.05$	$p<0.05$	$p>0.05$	$p>0.05$	$p>0.05$	$p>0.05$	$p<0.05$	$p>0.05$
	Companding	$p<0.05$	$p<0.05$	$p<0.05$	$p<0.05$	$p>0.05$	$p<0.05$	$p>0.05$	$p>0.05$	$p>0.05$	$p>0.05$
	Enhanced	$p<0.05$	$p<0.05$	$p<0.05$	$p<0.05$	$p>0.05$	$p>0.05$	$p>0.05$	$p>0.05$	$p>0.05$	$p>0.05$
Normal	**Unprocessed**	$p>0.05$	$p>0.05$	$p>0.05$	$p<0.05$	$p<0.05$	$p>0.05$	$p<0.05$	$p>0.05$	$p<0.05$	$p<0.05$
	Companding	$p>0.05$	$p>0.05$	$p>0.05$	$p<0.05$	$p>0.05$	$p<0.05$	$p<0.05$	$p<0.05$	$p<0.05$	$p<0.05$
	Enhanced	$p>0.05$	$p>0.05$	$p>0.05$	$p<0.05$	$p>0.05$	$p<0.05$	$p<0.05$	$p<0.05$	$p<0.05$	$p<0.05$

From the Table 4.6 it can be observed that there is no significant difference across all pairs of SNRs. The consonant identification scores in quiet condition was higher and was significantly more than the other SNR conditions was observed in AD population. Furthermore, no significant difference was observed between any two SNR conditions in unprocessed condition and 15dB v/s 5dB in companded condition was observed in this population. Importantly scores obtained at 10dB SNR and 15dB SNR did not differ significantly in this group.

In the control group, at all the conditions the consonant identification scores obtained at quiet condition did not differ significantly from the scores obtained at 15dB SNR, 10dB SNR and 5dB SNR. It is significantly different only between 0dB SNR and quiet condition. Additionally, in unprocessed condition, 15dB and 10dB showed significant difference. Whereas in the processed condition (companding and enhanced), 10dB and 5dB and 15dB and 5dB SNR also showed a significant difference.

4.4: Comparison of consonant identification scores obtained across stimulus condition at each SNR and within groups:

To check for the effect of stimulus conditions on consonant identification scores within each SNR, Friedman's test was performed. The results of the same is given in the Table 4.7 where the three stimulus conditions are compared at each SNR for each group. The χ^2 values with the degrees of freedom and the p values are listed.

Table 4.7 χ^2 value along with degrees of freedom and significance obtained for across stimulus conditions at each SNR for each group

Groups	**SNR (dB)**	χ^2 **(2)**	***p* value**
AD	**Quiet**	8.72	Significant ($p<0.05$)
	15	0.27	Not significant ($p>0.05$)
	10	2.25	Not significant ($p>0.05$)
	5	0.5	Not significant ($p>0.05$)
	0	2.07	Not significant ($p>0.05$)
Normal	**Quiet**	5.16	Not significant ($p>0.05$)
	15	1.2	Not significant ($p>0.05$)
	10	1.75	Not significant ($p>0.05$)
	5	3.77	Not significant ($p>0.05$)
	0	4.42	Not significant ($p>0.05$)

It was observed that the conditions were not significantly different at each SNR except for quiet condition which was observed only in AD population. Hence, pairwise comparison was carried out using Wilcoxon’s signed rank test only for quiet condition for AD group. Since the control group showed no significant difference across conditions at any SNR, further statistical analysis was not administered.

The Wilcoxon’s signed rank test was performed to check for significant difference of the three stimulus conditions by running pairwise comparison for the scores obtained in quiet situation in AD group. The results are tabulated in Table 4.8.

Table 4.8 Pairwise comparison of consonant identification scores obtained across conditions in quiet situation for AD

Quiet situation	Unprocessed	Companding
Companding	Significant ($p<0.05$)	
Enhancement	Not significant ($p>0.05$)	Not significant ($p>0.05$)

It was observed that, there was a significant difference between companding and unprocessed condition. But, no other condition pairs (unprocessed v/s enhancement and enhancement v/s companding) showed any significant difference between them in the quiet situation. This differences in the conditions were obtained only in AD population.

4.5: Comparison of consonant identification scores between sub-groups of AD:

The data collected from 10 individuals having AD exhibited large variability. This presented a need for further categorization of the clinical group into two subgroups as good and poor performers. For this purpose, a score of 8 in unprocessed quiet condition was set as a criteria for categorization. This made the initial group of 10 divide into 2 subgroups of 5 individuals each. The data obtained in both the subgroups were subjected to descriptive statistics to obtain mean, median and standard deviation. Table 4.9 shows the descriptive statistical analysis of the 2 sub-groups (represented as ADG for good performers and ADP for poor performers henceforth) across different conditions and 5 SNRs.

Table 4.9 Mean, standard deviation and median obtained for consonant identification for 3 conditions at 5 different signal to noise ratio of both subgroups of AD

Stimulus condition		**Unprocessed (across SNRs in dB)**					**Companding (across SNRs)**					**Consonant Enhancement (across SNRs)**				
Population		**quiet**	**15**	**10**	**5**	**0**	**quiet**	**15**	**10**	**5**	**0**	**quiet**	**15**	**10**	**5**	**0**
ADG	**Mean**	9	3.80	3.40	2	2.4	10.80	2.80	2.20	1.80	2.40	9.40	4	2.80	2.80	2.80
	SD	1	2.05	3.4	2.83	2.8	1.64	0.84	1.48	1.3	2.8	1.52	1.23	1.3	2.5	3.5
	Median	9	4	3	1	1	10	3	2	1	1	10	4	2	2	1
ADP	**Mean**	4.4	2.4	2.6	2.6	1	6.4	4.2	3.6	2.4	1.6	5.6	2	2	1.8	1.8
	SD	2.3	1.14	0.89	1.51	0.71	3.05	1.64	0.55	1.14	0.55	1.14	0.71	1.23	0.84	1.48
	Median	4	2	2	3	1	7	5	4	2	2	6	2	2	2	2

From this Table 4.9, it can be said that the individuals of auditory dys-synchrony poor (ADP) group had a low score across at all SNRs whereas, the good performers had better scores in quiet condition but performed lower in the presence of noise. However, in both the population there was a common trend of reduction in the scores in the presence of noise as against quiet situation in all the stimulus conditions. Further analysis of the data were carried out across conditions and SNR for both the subgroups.

The results of the subgroups are discussed under the following headings:

1. Comparison of consonant identification scores obtained across subgroups (ADG v/s ADP)
2. Comparison of consonant identification scores across condition within different SNR for each subgroup

4.6: Comparison across subgroups (ADG v/s ADP)

To compare the subgroups ADP and ADG in the three stimulus conditions in 5 SNRs, Mann-Whitney U test was run. The |Z| values with the significance values are tabulated in the Table 4.10.

Table 4.10 |Z| values obtained between ADG and ADP at all SNR in different conditions

Stimulus Condition	**SNR (dB)**				
	0	**5**	**10**	**15**	**Quiet**
Unprocessed	0.67	1.01	0.54	1.2	2.64
Companding	0.11	0.99	1.75	1.4	2.22
Enhancement	0	0.44	0.57	2.17	2.53

Note: shaded area in the Table represents a significant difference ($p<0.05$, 2-tailed)

From the Table it can be noted that there was a better performance of ADG subgroup when compared to ADP subgroup, which was significantly different only in quiet condition. Also, in the enhancement condition there was a significant difference between the subgroups at 15 dB SNR. The Mann Whitney test showed a significant difference for quiet condition only. As the significant difference was seen only in quiet condition, data was not analysed across SNR at each condition for each subgroup.

4.7: Comparison of consonant identification scores at each stimulus condition in quiet situation in subgroups:

Obtaining a significant difference of stimuli conditions in the quiet listening situation necessitated the comparison of conditions to see which conditions are significantly different from each other in both the subgroups separately. For this reason, Friedman's test was performed. The results of the same are tabulated in the Table 4.11.

Table 4.11 χ^2 value along with degrees of freedom and significance level obtained across condition in quiet situation for each subgroups of AD

Subgroups	χ^2 **(2)**	***p* value**
ADG	5.78	0.05
ADP	3.44	0.17

From the above Table it can be observed that there was no significant difference that was obtained for the ADP subgroup but, there was a marginal significant difference observed in the ADG group. To see which stimulus condition is bringing about a significant change in the consonant identification, pairwise comparison was done using Wilcoxon's signed rank test.

The Wilcoxon signed rank test results in the ADG group are tabulated in the Table 4.12.

Table 4.12 Pairwise comparison of conditions in quiet situation for ADG

Quiet situation	Unprocessed	Companding
Companding	Significant ($p<0.05$)	
Enhancement	Not significant ($p>0.05$)	Not significant ($p>0.05$)

In quiet situation, the scores obtained in companding condition was significantly better than unprocessed condition. But, no other condition pairs (unprocessed v/s enhancement and enhancement v/s companding) showed any significant difference between them in the quiet situation. This differences in the conditions were obtained only in ADG population.

4.8: Consonant confusion matrices:

Nineteen syllables were considered with processed (companding and enhancement) and unprocessed stimulus conditions. The listeners were forced to guess within 19 syllables for every sound presentation. The final response was obtained in the form of matrix. In the matrix the first row indicates the responses and the consonants listed vertically in the first column indicates the stimulus presented. The number in each cell represents the frequency of the particular stimulus-response pair. The number in the cells along the principal diagonal axis would be the correct response. The analysis was started with adding all the response matrices of AD population for the respective stimulus condition in quiet situation. The different SNRs were not considered as the correct scores obtained were less than 50% of total correct scores in most of the participants. An example of the added matrix is provided in Table 4.13.

Table 4.13 Example of a stimulus response matrix showing the results obtained for companding condition for 10 participants with correct response highlighted in the diagonal axis

	b	**tʃ**	**d**	**d̲**	**g**	**k**	**l**	**l̲**	**m**	**n**	**p**	**r**	**s**	**ʃ**	**t**	**t̲**	**j**	**dz**	**v**
b	**5**			1	1						1				1			1	
tʃ		**6**					1											3	
d	1		**4**	1	2										1	1			
d̲	1		3	**3**	1										2				
g			3	1		3			1			1			1				
k			1		1	**6**							1		1				
l					1		**4**		3	1									1
l̲	1						1		1	4					2		1		
m	1						1		**4**	3	1								
n							1		4	**4**							1		
p	2					1	1				**4**				2				
r									2		1	**6**				1			
s													**4**	5				1	
ʃ		1												**9**					
t			1	1							2				**5**	1			
t̲			2	1	1										1	**5**			
j							1		3								**6**		
dz		3															1	**6**	
v							1		1	1	2								**5**

Sequential information transfer analysis (SINFA) (Wang & Bilger, 1973) was performed using the ‘Feature Information Xfer (FIX) ’software (developed by University College of London, Department of Linguistics) on the added confusion matrices to assess the amount of information transfer from stimulus to response for place, manner and voicing. SINFA considers the amount of information transmitted in terms of electronic units of ‘bits’. To analyse the 19 consonants, a feature matrix was formed using the voicing, place and manner features as a basis for classification. The same is shown in the Table 4.14.

Table 4.14 Feature matrix of the 19 syllables considered

	b	**d**	**g**	**dz**	**k**	**l̲**	**l**	**m**	**n**	**p**	**r**	**s**	**t̲**	**v**	**j**	**tʃ**	**t̲**	**d̲**	**ʃ**
Voicing	+	+	+	+	-	+	+	+	+	-	+	-	-	+	+	-	-	+	-
Place	b	a	v	p	v	p	a	b	a	b	a	a	a	l	p	p	d	d	p
Manner	p	p	p	a	p	l	l	n	n	p	l	f	p	g	g	a	p	p	f

Note: Voicing: +=voiced, -=voiceless
Place: b=bilabial, a=alveolar, v=velar, p=palatal, l=labial, d=dental
Manner: p=plosives, a=affricates, l=laterals, n=nasals, f=fricatives, g=glides

Zero information transmitted indicates no transmission of a particular feature and one indicates complete transmission. In this experiment, the maximum information that can be transmitted for 19 syllables was 4.24.

The total information transmitted is observed to be more in processed conditions (companding and consonant enhancement) when compared to unprocessed condition. Among the two enhancements applied, the information transmitted is greater for the companding when compared to consonant enhancement condition. The information transmitted for different phonetic features and the total information transmitted are represented in the figure 4.2.

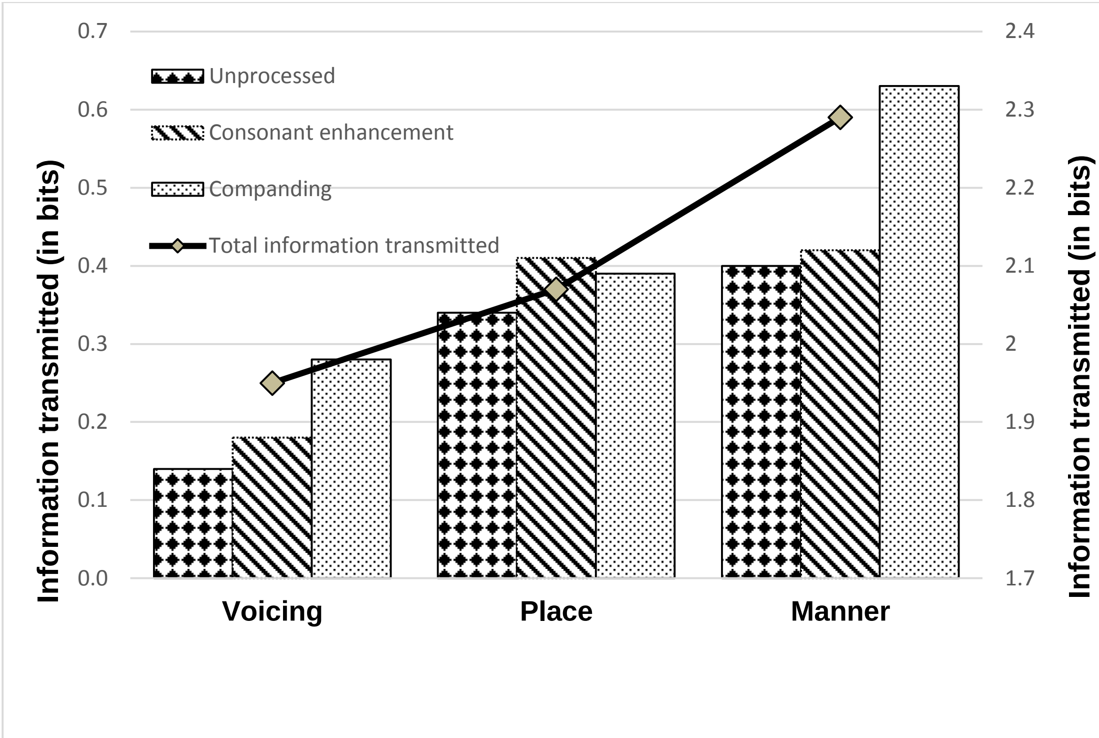


Figure 4.1 Relative information transmitted across different stimulus condition

From the figure 4.1 it is noted that manner cue is transmitted the maximum in all the stimulus condition, whereas voicing cue is transmitted the least across the three stimulus condition, which is followed by place. Among the stimulus conditions, companding showed maximum improvement for voicing and manner. Whereas, place showed marginal improvement in the consonant enhancement condition.

Results obtained in the study can be summarised as follows:

Groups

ADG

ADP

Significant difference between groups for amount of improvement with companding in quiet.

No Significant difference between subgroups for all conditions and SNRs except in quiet situation

Conditions

Normal

AD

No Significant difference between conditions for all SNRs in normal group

No Significant difference between conditions for all SNRs except in quiet situation in AD group

SNR

Normal

AD

Significant difference of low SNR (0dB) with other SNR conditions and quiet situation

Significant difference of quiet condition with other SNRs.

SINFA

Place

Manner

Voicing

Manner cue is transmitted the maximum in all the stimulus condition, followed by place. Voicing cue is transmitted the least across the three stimulus condition.

Chapter 5

Discussion

The objective of this study was to investigate the effect of spectro-temporal enhancement of speech on speech perception in individuals with AD and normal hearing ability. Consonant identification task was administered in quiet, and in the presence of noise (with different SNRs) to see the effect. The scores from all the participants were compared across different stimulus conditions, SNRs and also between groups and within group. The results are discussed below.

5.1: Effect of signal enhancement techniques and SNRs on speech perception between clinical and control group:

Syllable identification was significantly better in normal hearing individuals when compared to individuals with AD. This finding was consistent across all the SNRs and the stimulus conditions considered.

Studies on speech perception in normal hearing individuals show that they utilise both envelop and fine structure cues to understand speech, especially in adverse listening conditions (Dorman, Loizou & Tu, 1998). Whereas, owing to the pathophysiology, individuals with AD are not able to extract the cues available in the speech. Adding noise in the background would exaggerate the speech perception problem due to reduction in modulation depth and addition of spurious waveform (Drullman, 1995; Zeng & Liu, 2006). This reduced accessibility of cues probably is the reason for the reduced scores in AD when compared to normal hearing individuals in the present study. Also, excessive masking can be another reason for the poor performance by AD when compared to normal hearing individuals in the presence of noise (Zeng et al., 2005).

5.2: Between group comparisons of amount of improvement obtained from different stimulus techniques:

The findings obtained showed that there was significant benefit from companding over unprocessed condition in individuals with AD as against normal hearing individuals who demonstrated no improvement in quiet situation and also at 10dB SNR. The consonant enhancement showed a marginal improvement when compared to unprocessed condition in quiet situation.

The amount of benefit from companding can be attributed to the spectral and temporal enhancement that is provided by the technique which is proven to aid speech perception in AD (Shachi, 2012; Narne et al, 2014). This is especially helpful for AD as they fail to utilize the cues owing to poor temporal perception and the enhancement helped to improve the spectral contrast and probably made the cues more accessible. There was a marginal benefit noted from consonant enhancement in individuals with AD as well. This can be attributed to the increment given to consonant which had improved the spectral contrast and hence increasing the speech perception. However, it did not reach the significance level. This can be because the increment given might not have necessarily increased the spectral contrast sufficiently or the amplification might be insufficient which still could get masked by the following vowels (backward masking effect).

Though there was a significant improvement obtained in AD with processed stimuli when compared to unprocessed stimuli, the benefit obtained for companding or enhanced stimulation was not consistent across SNRs and it did not show any specific trend.

Individuals with AD are very sensitive to the presence of background noise. This can be due to excessive masking which makes it difficult to utilise the enhanced cues, making the consonant identification scores almost similar to that of unprocessed stimuli. Thus, the benefit of processed signal was restricted to quiet condition with no significant difference at other SNRs.

However, benefits were not noted in normals as they could utilise the envelop cues and fine structure cues available in speech in quiet and noisy conditions. Also, in normal hearing individuals there was ceiling effect in unprocessed condition, the amount of improvement could not be appreciated well in them.

5.3: Effect of SNR on consonant identification scores obtained for each stimulus condition within the groups:

5.3.1: Control group:

The consonant identification scores obtained were maximum in quiet situation and was found to deteriorate with decreasing SNR. The number of syllables identified were least at 0dB SNR. The same trend was noted for all the stimulus conditions.

This is a well-established observation that speech identification ability reduces in the presence of noise and makes it an adverse listening condition to communicate (Dorman et al., 1998). However, in normals, the difficulty is not present until the SNR is low. This is credited to the normal auditory physiology which is able to differentiate the wanted signals (speech) form the unwanted signals (background noise). Hence, at higher SNRs, normals can compensate by extracting the fine structure cues along with envelop cues to understand the speech (Wang & Bilger, 1973).

The consonant identification scores showed a significant difference of 0dB SNR with other SNR conditions and quiet situation for all stimulus conditions.

The accessibility to envelop and fine structure cues is not possible at a very low SNR. It is because, the noise reduces the modulation of speech envelop and also distort the temporal fine structure of speech (Houtgast & Steeneken, 1985; Drullman, 1995). Thus, reducing the consonant identification scores in the presence of noise at higher levels (low SNR).

5.3.2: Clinical group:

In the present study, it was noted that the consonant identification scores were low and variable in AD individuals. Maximum scores were obtained in quiet condition and a drastic reduction in scores was noticed even at high SNR itself.

This poor performance can be attributed to inability to process the temporal information due to asynchronous neural firing. This variability in performance in speech perception is a common finding across studies (Kumar & Jayaram, 2006; Narne & Vanaja, 2008). *In the presence of noise, at 15dB SNR, the scores decreased greater than 50% when compared to quiet situation. Further reduction in SNR did not show a significant reduction in consonant identification scores*. There was a significantly poor score obtained at 15dB SNR itself almost reaching the floor effect. Thus, further reduction in SNR did not change the scores significantly. This suggests that individuals with AD fail to perceive the stimulus information even if the environment is slightly not favourable. The physiology behind this extreme sensitivity to noise in AD is unclear. But, it can be attributed to excessive masking effects that is observed in them for both pure tones and speech (Zeng et al, 2005).

5.4: Effect of stimulus condition on consonant identification scores obtained across SNRs within the groups

5.4.1: Control group:

There was no significant difference obtained between the processed stimuli and the unprocessed stimuli and also between two processed stimuli across SNRs for normal hearing individuals.

The normal hearing individuals identified almost all consonants in quiet condition for unprocessed stimuli (ceiling effect). They do not face any difficulty in perceiving the envelop or the fine structure cues in quiet or higher SNR conditions. Hence, the role of enhancement or companding contributing to the intelligibility could not be appreciated especially in the quiet situation or at high SNRs.

In the lower SNRs, the consonant identification scores were higher for unprocessed stimuli than the processed stimuli but this was not statistically significant. These findings can be explained on the basis that the normal hearing individuals depend majorly on the spectral characteristics and the format transition to perceive the consonants (Liberman, Delattre and Cooper, 1952; Blumstein & Stevens, 1979). This might be because, the normal hearing auditory system can derive sufficient information from amplitude modulation of the speech in quiet situation. But, depends on frequency modulation in the presence of noise as the noise affects the intensity cues first preserving the frequency cues (Zeng et al., 2004). Whereas, application of the techniques changed either the spectro-temporal characteristics or the transition and hence probably acted as a distortion to naturally available cues for normal hearing individuals. This might have led to reduced scores in low SNRs in the processed conditions. However, these differences were not statistically significant and thus it can be said that the techniques did not cause distortions to such an extent that it compromised the intelligibility. These findings are however not in consensus with previous study by Shachi (2012) who found a significantly better consonant identification in VCV condition in the presence of speech shaped noise at 0 dB SNR with companding when compared to unprocessed stimulus in individuals with normal hearing. They also obtained SNR 50 at a significantly lower SNR (better) with companded stimulus for normal hearing individuals. Along with Shachi (2012), other studies by Baer et al, (1993); Bunnel, (1990) also support the finding that enhancement provided better speech identification especially in low SNR conditions. These discrepancies in the findings might be due to the stimuli considered, training with the processed stimuli and background noise used across studies.

5.4.2: Clinical group:

Companding was proven to be significantly beneficial than consonant enhancement and unprocessed condition only in quiet condition in the current study. However, there was no difference noted between consonant enhancement and unprocessed condition in quiet situation also.

The advantage of companding can be owed to the time domain spectral enhancement which enhances spectral and temporal contrasts of speech making cues more obvious and in turn aiding speech perception. This findings are consistent with other studies where effect of companding was studied on VCV syllables and word identification in AD population with companding (Shachi, 2012; Narne et al, 2014). Thus, the improvement in speech perception observed in the present finding along with previous studies can be acknowledged to the enhancement in the spectral and temporal domain which is brought about by companding.

Between the two modifications, companding condition (average = 9.5) did show a better improvement when compared to the consonant enhancement (average= 7), however was not statistically significant.

The process behind consonant enhancement was that the burst and the transition portion of the syllable were given an additional gain of 6 dB to improve the consonant vowel ratio and aid the speech identification ability. However, in the present study, no improvement was observed following this modification. It can be reasoned that the increment given to the consonant could have masked the transition of the consonant to vowel or there is also a possibility that the enhancement was not sufficient to bring about a change in the consonant to vowel ratio. In addition to that, the amplification given probably might not help to overcome the effect of excessive masking effect on consonants by vowels. Also, the scores in enhancement being less than companding can be probably that the spectral contrast that is brought about is not as efficient as the contrast that is obtained due to companding.

At 15dB SNR, there was a difference in the performance between companding (average = 4), enhancement (average = 2.5) and unprocessed condition (average = 2.5), but was not significant statistically.

This lack of significance suggests that the individuals with AD could not benefit from the enhancement provided in the presence of noise. However, the same effect was not observed across other lower SNRs that is, the performance did not vary significantly with different spectro-temporal enhancement and the unprocessed condition. This is attributed to the inability of the individuals with AD to perceive the temporal fine structure information of the speech which helps in speech perception especially in the presence of noise which is a documented result (Zeng & Liu, 2006; Kumar & Jayaram, 2008).

5.5: Effect of stimulus condition and SNR on speech perception in subgroups of AD

Following a large variability in the performance of individuals with AD, the group was further divided as good and poor performers based on the speech perception ability in quiet situation. The effect of enhancement conditions in the subgroups were analysed across different SNRs. *The application of the stimulus modification did not bring about a significant difference between the two sub-groups across different SNRs. But, a better improvement for auditory dys-synchrony good performers (ADG) group was observed in quiet situation with companding when compared to auditory dys-synchrony poor performers (ADP) group*. This can be reasoned out as the individuals with AD who had a better speech perception scores in quiet situation were able to utilise the enhancement provided better than ADP group. The individuals in ADG group probably had a less severe form of dys-synchrony when compared to ADP group. Hence, their temporal processing ability is not severely affected which is aiding them to utilise the enhancement better. This findings is in consensus with the study by Narne and Vanaja (2008) who also observed an improvement following envelop enhancement in the group of AD. But, AD with severe dys-synchrony exhibited greater variability in the improvement when compared to mild to moderate degree of dys-synchrony who had less variability in the improvement.

5.6: Consonant perception and analysis of error pattern:

The information transmitted showed that manner cue is transmitted the information more efficiently when compared to place and voicing across all the stimulus conditions.

5.6.1: Perception of Manner:

Information transmission analysis revealed that manner was transmitted more efficiently in the unprocessed condition, when compared to the other features. The major cue for manner perception is believed to be the envelop modulation in the speech signal. The errors observed were also within category of substitutions. The perception of continuants (fricatives, nasals) were less affected as they are cued by slow fluctuations of the envelop. The perception of slow fluctuations of the sound are relatively preserved in individuals with AD (Zeng et al 1999; Narne & Vanaja, 2008). The sounds with faster fluctuations like the stops were found to be most affected as the perception of faster fluctuations of the envelop are compromised in AD (Zeng et al 1999; Kumar & Jayaram, 2006 and Narne & Vanaja, 2008)

With processed stimuli, there was a greater increase in information transfer that was observed for manner.

With companding, the spectrum of the speech sound is enhanced across all the frequencies. It also would result in increase in contrast of the sound. This would make the modulations or the fluctuations of the envelop more robust leading to improvement in perception of manner. This probably would have led to the greater increase in the transmission of the manner feature in this experiment.

With consonant enhancement, the consonant burst portion were enhanced. This would also improve the spectral contrast. However, this resulted only in marginal improvement in transmission of manner cue. This can be attributed to the amplification given (6dB) which probably was not sufficient to bring a change in the fluctuations so much to improve manner perception.

5.6.2: Perception of Place:

Perception of place feature mainly depends on the burst and formant transition (<50ms). Individuals with AD have difficulty in processing the short duration cues and hence there is less information transmission of the place feature. *The number of errors were also high with respect to place of articulation especially of stop consonants*. This also is a consistent finding across studies (Kraus et al, 2000; Kumar & Jayaram, 2011).

With processed stimuli, there was a marginal increase in information transfer that was observed for place.

Individuals with AD require more depth of modulation to process faster envelopes in the speech signal. This would have hampered the perception of major acoustic cues to identify place of articulation (formant transition, burst amplitude) (Lahiri, Gewirth, & Blumstein, 1984; Ohde & Stevens, 1983). In this study, the consonant enhancement technique mainly amplified the burst and the transition portions of the consonant. The studies where enhancement on speech perception was considered are in consensus with the result of the current experiment. Guelke (1987) also found improved discrimination ability of place of articulation especially of stop consonants following enhancement of the burst portion of the syllables in sensorineural hearing loss individuals. Hence, enhancement of the transition and burst portion could probably be the reason for greater transmission of place feature.

5.6.3: Perception of Voicing:

Voicing feature information was transmitted least when compared to other phonetic features in all the stimulus conditions.

Perception of voicing cues mainly depends on voice onset time (VOT) and first formant onset and transition (Summerfield & Haggard, 1977). This can be attributed to altered temporal processing which affects the perception of short duration cues like VOT. In addition to that, voicing is predominantly a low frequency cue and perception of low frequency is noted to be poor due to dys-synchronous firing in AD (Rance, McKay, & Grayden, 2004). These factors might be the reason for voicing to be transmitted least in the present study.

However, with companding, there was an improvement in the transmission of the voicing feature. This can be attributed to the spectral enhancement that is applied across all the frequencies in companding. The enhancement given may have probably made the

VOT and the F1 onset and the transition more robust which would have increased the perception of voicing in AD.

When consonant enhancement was applied, there was a marginal improvement observed with respect to transmission of voicing. This can probably be due to the enhancement of the transition of the consonant stimuli (Kumar & Jayaram, 2011 & Narne & Vanaja, 2008).

Thus, there was an improvement in consonant identification scores in processed condition when compared to unprocessed stimulus condition. The spectro-temporal enhancement also improves the spectral contrast of the speech stimulus (Turicchia & Sarpeshkar, 2005). This makes the cues more prominent for individuals with AD to utilise. Hence, the total information transmitted was greater in processed than the unprocessed condition. Out of the two processed conditions, companding was found to be more efficient in transmitting the information when compared to consonant enhancement.

Chapter 6

Summery and Conclusion

It is said that "The right timing is in all things the most important factor" (Hesiod, 750-650 BC). Any deviation in perception of timing information can lead to misrepresentation of the image, be it visual or auditory. Auditory dys-synchrony is an example of such disorder of the auditory system where the perception of the temporal features is affected leading to distorted auditory image (Zeng et el., 1999; Zeng & Liu, 2006). Since its first description by Starr (1991), various researchers have studied the underlying pathophysiology and the perceptual consequence. The striking feature of AD is found to be impaired speech perception in the presence of relatively spared audibility.

Management of such disorder continue to pose many problems as there is no standard treatment protocol to overcome all the difficulties faced by individuals with AD. In addition to that, the improvement of speech perception in natural environment with noise remains as one of the biggest concerns in today's rehabilitation professionals. In this line of research, there were other modifications like enhancement, companding that were suggested in order to rehabilitate more efficiently (Kraus, 2000; Zeng & Liu, 2006; Kumar & Jayaram, 2006 & Narne & Vanaja, 2008). Hence, this study was taken up with the aim to evaluate the effect of companded and consonant enhanced speech on speech perception in quiet and in noise among normal hearing individuals and individuals with auditory dys-synchrony.

To fulfil this aim, two groups of participants were considered where group 1 consisted of 10 individuals with normal hearing sensitivity and group 2 consisted of 10 individuals with auditory dys-synchrony. 19 consonant vowel combinations were considered to obtain the identification scores in three stimulus conditions (unprocessed, companding and consonant enhancement) and in five different SNRs they are, quiet, 15, 10, 5 and 0 dB SNR. The scores from all the participants were compared across different stimulus conditions, SNRs and also between groups and within group. Statistical analysis of these data revealed:

1. There was a significant group effect observed at each SNR and each condition.
2. With the increase in SNR, there was an improvement in consonant identification scores in both the groups for processed and unprocessed conditions.
3. In quiet condition, maximum scores were obtained in both the groups. The normal hearing group reached a ceiling effect for both processed and unprocessed condition.
4. Individuals with AD obtained a significant improvement with companding when compared to unprocessed condition. There was an improvement in performance in companding when compared to consonant enhancement however it was not statistically significant.
5. The amount of benefit observed in individuals with AD with processed stimulus was significantly greater as against normal hearing individuals who did not show any improvement. This finding was noted only in quiet situation.
6. In the subgroups of AD, the amount of improvement with signal enhancement was found to be more in good performers when compared to poor performers.

7. The consonant errors obtained were subjected to SINFA to quantify the error patterns. In individuals with AD, majorly voicing and place errors were noted when compared to manner in the unprocessed condition.
8. The pattern of errors were the same across stimulus conditions. However, the amount of transmission showed an improvement with respect to manner and voicing transmission in companding and marginal improvement of place transmission in consonant enhancement condition when compared to unprocessed stimulus.

The use of enhancement techniques presented with betterment in consonant identification scores. This can be attributed to the spectro-temporal enhancement which brought about a change in the spectral contrasts of speech. This enhancement in the contrast helped the individuals with AD to extract the acoustical cues available and improved the speech perception.

6.1 Conclusion:

There are various enhancement strategies which has revealed improvement in speech perception but, it has not yet implemented clinically. This study is another attempt to show the behavioural improvement with spectro-temporal enhancement which can benefit the individuals with AD. Out of the different techniques considered, companding showed a significantly better improvement in consonant identification. Consonant enhancement showed a marginal improvement when compared to unprocessed stimulus condition.

It is also an indirect approach of showing that if hearing aids can implement the companding or consonant enhancement strategy, the failure of hearing aids for management of AD can be overcome. But, all the benefit noted for stimulus condition are significant and less variable in the group where the temporal processing is relatively spared. Thus, conclusion from the study should be cautiously made before generalisation.

6.2 Implication:

1. This study has highlighted the understanding of the speech perception ability of individuals with AD.
2. It helps to understand the consequence of temporal processing on speech perception (pathophysiological basis for speech understanding).
3. This study gives an insight about comparison of spectral enhancement and burst enhancement. On the basis of the outcome of this study, signal processing strategies can be designed for better speech perception.

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.

Baer, T., Moore, B.C.J., & 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., Morlet, T., Rose, K., & Brashears, S. (2003). Auditory Neuropathy/Dys-synchrony: Diagnosis and Management. *Mental Retardation and Developmental Disabilities Research Reviews, 9*, 225-231.

Berlin, C., Hood, L., & Rose, K. (2001). On renaming auditory neuropathy as auditory dys-synchrony.*Audiology Today, 13*(6), 15-17.

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

Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America, 66*(4), 1001-1017.

Budenz, C. L., Telian, S. A., Arnedt, C., Starr, K., Arts, H. A., El-Kashlan, H. K., & Zwolan, T. A. (2013). Outcomes of cochlear implantation in children with isolated auditory neuropathy versus cochlear hearing loss. *Otology & Neurotology*, *34*(3), 477-483.

Bunnel, H. T. (1990). On enhancement of spectral contrast in speech for hearing impaired listeners. *Journal of Acoustic society of America, 88,* 2546-2556.

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

Dolphin, W. F. (2004). Auditory neuropathy and configured hearing loss: the case for two-stage screening. *Hearing Review*, *11*(2), 28-33.

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

Doyle, K. J., Sininger, Y., & Starr, A. (1998). Auditory neuropathy in childhood. *The laryngoscope, 108*(9), 1374-1377.

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

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America, 112, 259–271.

Gibson. W. P., & Sanli, H. (2007). Auditory neuropathy: an update. *Ear and Hearing, 28* (2 supplement), 102S-106S.

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

Hazan, V., Simpson, A. and Huckvale, M. (1998) Enhancement techniques to improve the intelligibility of consonants in noise: Speaker and listener effects. *Proceedings of International Conference of Speech and Language Processing,* Sydney, Australia, December 1998.

Jain, C., Konadath, S., Vimal, B., Suresh, V., (2014). Influence of native and non-native multitalker babble on speech recognition in noise. *Audiology Research,* 4:89.

Jijo, P. M., & Yathiraj, A. (2012). Audiological characteristics and duration of the disorder in individuals with auditory neuropathy spectrum disorder (ANSD)—a retrospective study. *J Indian Speech Hear Assoc*, *26*(1), 17-26.

Ketelaar, L., Rieffe, C., Wiefferink, C. H., & Frijns, J. H. (2012). Does hearing lead to understanding? Theory of mind in toddlers and preschoolers with cochlear implants. *Journal of pediatric psychology,37*(9), 1041-1050.

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, *115*(1), 362-378.

Kraus, N. (2001). Auditory neuropathy: an historical and current perspective. Auditory Neuropathy: *A New Perspective on Hearing Disorders. San Diego, CA: Singular Thomson Learning*, 1-15.

Kraus, N., Bradlow, M. A., Cunningham, C. L., King, C. D., Koch, D. B., Nicol, T. G., Mcgee, T. J., Stein, L. K., & Wright, B. A. (2000). Consequence of neural asynchrony: A case of auditory neuropathy. *Journal of the Association for Research in Otolaryngology, 01*, 33-45.

Kumar, U. A., & Jayaram, M. (2011). Speech perception in individuals with auditory dys-synchrony.*The Journal of Laryngology & Otology*, 125(03), 236-245.

Kumar, U. A. (2011). *Perception of Some Temporal Parameters of Speech in Individuals with Auditory Dys-synchrony*. Unpublished doctoral dissertation, University of Mysore, Mysore

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

Lahiri, A., Gewirth, L., Blumstein, S. E. (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross language study, *Journal of the Acoustical Society of America*, 76, 391–404.

Liberman A.M., Delattre P.C., & Cooper F.S. (1952). The role of selected stimulus variables in the perception of the unvoiced stop consonants. *American Journal of Psychology, 65*, 497-516.

Liu, S., Del Rio, E., Bradlow, A. R., & Zeng, F. G. (2004). Clear speech perception in acoustic and electric hearing. *Journal of the Acoustical Society of America*, 116(4, Pt. 1), 2374–2383.

Montgomery AA, Prosek RA, Walden BE, Cord MT., (1987). The effects of increasing consonant/vowel intensity ratio on speech loudness. *Journal of Rehabilitation Research and Development*. *24*(4), 221-228.

Moore, B.C.J., Garnier, S., Carn, G., Lorenzi, C., (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. Proceedings of the National Academy of Sciences of the United States of America.

Moore, B. C. J. (2003). Speech processing for the hearing-impaired: Successes, failures, and implications for speech mechanisms. *Speech Communication*, *41,* 81–91.

Moore, B. C., & Oxenham, A. J. (1998). Psychoacoustic consequences of compression in the peripheral auditory system. *Psychological review,105*(1), 108.

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

Narne, V. K. (2013). Temporal Processing and Speech Perception in Noise by Listeners with Auditory Neuropathy. *PLoS ONE, 8*(2). doi:10.1371/journal.pone.0055995

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

Ohde, R. N., Stevens, K. N. (1983). Effect of burst amplitude on perception of stop consonant place of articulation. *Journal of the Acoustical Society of America*, 74, 706–714.

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

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking Clearly for the Hard of Hearing IIIAn Attempt to Determine the Contribution of Speaking Rate to Differences in Intelligibility between Clear and Conversational Speech. *Journal of Speech, Language, and Hearing Research, 32*(3), 600-603.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking Clearly for the Hard of Hearing IIAcoustic Characteristics of Clear and Conversational Speech. *Journal of Speech, Language, and Hearing Research, 29*(4), 434-446.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking Clearly for the Hard of Hearing IIntelligibility Differences between Clear and Conversational Speech. *Journal of Speech, Language, and Hearing Research, 28*(1), 96-103.

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

Rance, G., Barker, E.J., Mok, M., Dowell, R., Ricton, 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., McKay, C., & Grayden, D (2004). Perceptual characterization of children with auditory neuropathy. *Ear and Hearing, 25*, 34-46.

Range-Samuelson, C. L., Drake, S., & Wackym, A. P. (2008). Quantitative analysis of electrically evoked auditory brainstem responses in implanted children with auditory neuropathy/dys-synchrony. *Otology and Neurology, 29*, 174-178.

Shachi (2012). *Perception of Spectrally Enhanced Speech through Companding in Individuals with Auditory Neuropathy*, unpublished masters dissertation, University of Mysore, Mysore.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304.

Simpson, A. M., Moore, B. C. J., & Glasberg, B. R. (1990).Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners. *Acta Oto-Laryngology*, *469*, 101-107.

Sininger, Y. S., & Trautwein, P. (2002). Electrical stimulation of the auditory nerve via cochlear implants in patients with auditory neuropathy. *Annals of Otology Rhinology And Laryngology*, *111*(5; PART 2), 29-31.

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

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

Stone, M. A., & Moore B. C. J. (1992).Spectral feature enhancement for people with sensorineural hearing impairment: effects on speech intelligibility and quality. *Journal of Rehabilitation Research and Development*, *29* (2), 39-56.

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

Summerfield, Q., Haggard, M. (1977). On dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America*, *62*, 435–448.

Trautwein, P. G., Sininger, Y. S., & Nelson, R. (2000). Cochlear implantation of auditory neuropathy. Journal-American Academy of Audiology, 11(6), 309-315.

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

Varga, R., Kelley, P. M., Keats, B. J., Starr, A., Leal, S. M., Cohn, E., & Kimberling, W. J. (2003). Non-syndromic recessive auditory neuropathy is the result of mutations in the otoferlin (OTOF) gene.*Journal of medical genetics*, *40*(1), 45-50.

Vandana (1998) *Speech Identification Test for Kannada Speaking Children,* unpublished masters dissertation, University of Mysore, Mysore.

Wagner, W., Heppelmann, G., Vonthein, R., & Zenner, H. P. (2008). Test-retest repeatability of distortion product otoacoustic emissions. *Ear and hearing*, *29*(3), 378-391.

Werner, L. A., & Gray, L. (1998). Behavioral studies of hearing development. In *Development of the auditory system* (pp. 12-79). Springer New York.

Yang, J., Luo, F., and Nehorai, A. (2003).Spectral contrast enhancement: Algorithms and comparisons. *Speech Communication, 39*, 33–46.

Yathiraj, A. & Vijayalakshmi, C. S. (2005). *Phonemically Balanced Word List in Kannada*, unpublished departmental project. Developed in Department of audiology, AIISH, Mysore

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

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology*, *93*(6), 3050-3063.

Zeng, F. G., Oba, S., Garde, S., Sininger, Y., & Starr, A. (2001). Psychoacoustics and speech perception in auditory neuropathy. *Auditory Neuropathy*, 141-164.

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