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

Madhuri Sharma
13AUD011

ALL INDIA INSTITUTE OF SPEECH AND HEARING
★MYSORE★

This Dissertation is submitted as part fullfillment
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 Cochlear Hearing Loss”** is a bonafide work in part fulfillment for the degree of Master of Science (Audiology) of the student with Registration No. 13AUD011 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 Cochlear Hearing Loss"** 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 Cochlear Hearing Loss is the result of my own study under the guidance of , 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**: 13AUD011

May, 2014

Acknowledgements

‘I can no other answer make but thanks, and thanks, and ever thanks!’ - William Shakesphere

I thank the superpower that always looked upon our shoulders, giving us strength to overcome obstacles and blessing us with all the good times.

I extend my profound gratitude to my parents and family who gave me the freedom to pursue my interests and always supported me in all my endeavours. You have inspired me in various ways over time and are the reason for any bit of success I have achieved till date. Mayur, you are a powerhouse and I’m sure you’ll channelize all your energy and achieve great heights! Good luck to you! ☺

Animesh sir, I’m awed by how amazing a teacher you are! Your experience and knowledge speak for themselves. You have guided us not only through this dissertation but have also taught us the a b c’s of various aspects of research and at the same time, made the whole process fun. I’m lucky to have been your student and I fall short of words to thank you for all the encouragement, support and guidance you have provided all through the 6 years I have spent here.

I express my endless gratitude to Vijay sir. Sir, I have immense respect for your knowledge, the dedication and passion you show towards Audiology and yet, be the humble person you are. Never did we hear ‘come back later’ when we approached you for help, even if it meant staying back after 5 30. Thank you for everything sir!

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 shall fail in my duty if I do not thank every single staff here who have taught me and imparted knowledge in me about everything related to speech and hearing. All of you have inspired me in your own ways, for the dynamic personalities you possess! I specially thank my lecturers in Audiology for being the awesome people they are and strengthening my interest in the field. Sharath sir, Antony sir, Ismail sir, Rama ma'am and Baba sir, thank you for making us better clinicians and teaching us how to apply our knowledge clinically. I extend my gratitude to Nisha akka, Deepashree akka, Nike sir, Srikar sir, for your valuable inputs and staying back for us afterhours to keep the department open.

My sincere thanks to Vasanthalakshmi ma'am and Santosh sir for helping me make sense out of the numbers. Thank you ma'am for making statistics simple and comprehensible.

Bhuvana and PG, I'm so glad to have had you as my co-pros! Thank you for making this otherwise serious process, a lot of fun! :D From being good classmates to becoming hippo, Julian and panda, the journey has been all smiles, be it arguing over silly conversations or hanging out together. It has helped me remain sane even during times of stress. Thank you! :D

Knowledge, wisdom, and a bunch of amazing people always willing to help, all under one roof. That's what I call our library here. My heartfelt thanks to this department for providing us with all the resources required for completing this dissertation. Thank you electronics, especially Ravi sir, for helping with any technical issue that occurred.

I never realised how time flew in masters. Thanks to our amazing ‘mafia family’ for adding the much needed zest and entertainment in the otherwise (not so entertaining :P) academic life. Classes were ALWAYS fun, especially, the time before and after the lecturer was present in class. ;) I also wish to thank my whole masters batch for being such wonderful classmates and to have given me enough memories to cherish for a lifetime. ☺

Whether I needed a shoulder to cry on or some company to have a crazy time, I knew who to turn to. Anju and Amruth, if not for you guys, I wouldn’t be grinning so widely thinking of ALL the times we spent, right from the time I stepped into this college. Zeena and Riniya, my life here, especially in hostel was what it was, because of you guys. A few ‘normal’ times and a plethora of great times! Usha, you are blessed with talent and a great friend like me. Haha! Phebe, thank you for being the ‘cherry on the top’ in my masters life. You complete it. Tithi, you are a very helpful person. Thank you for being so nice to me and trying to solve all my problems with your ultimate ‘ajwain therapy’! ;) All of you guys not only helped me put together this dissertation, but also myself, whenever necessary. What would I have done without you guys? ☺

Thank you Paddy, Preeta, Chucha, Kiran, Bibek, Husna, Akshay, Akhil, Thareeque, Nirmala and Cheta. Class or clinics, you guys made it memorable.

Rohini aunty and Kishore uncle, a mere thank you is insufficient for all that you have for us. You are my home away from home, family away from family! ☺

Thank you Achu, for standing by me through thick and thin. You have it in you to make your dreams come true. Thank you for all the help with my dissertation and much more in life. Good luck to you! ☺

Pavan, I can’t thank you enough for believing in me and for your ever encouraging words. Thank you for being the big brother I always wanted! Prathyusha, when I doubted myself, you told me I had it in me to achieve what I wanted to. A thank you is too formal for you but anyway, thank you! :P Niki, Mad, and Manasa, I know I can always count on you. Good times or not, you were a part of it! ☺

Adi and Sharanya, you guys are two of the nicest people I know. Thank you for the much needed breaks from time to time! ☺

I am grateful to many of my seniors and juniors who have directly or indirectly helped my finish this dissertation with a smile on my face, still intact. If I name one, I’ll have to name them all and this section just ain’t enough for that! I wish you all a bright future and all the luck and happiness! ☺

My most sincere gratitude to the participants of this study. Thank you for your selfless cooperation that made this dissertation happen.

Abstract

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

Table of Contents

Results

List of Tables

List of Figures

Chapter 1

Introduction

Communication plays a key role in one’s daily living. It helps in expressing one’s ideas or needs and is central to any individual’s life. The most common form of communication involves information in the form of speech which needs to be heard, decoded into its acoustic information and then perceived. Hearing is a sequence of events wherein the ear transforms the sound waves into electrical signals which acts like a biological microphone and sends these nerve impulses (electrical signal) to the brain where they are processed and interpreted as sound. It starts with the sound waves impinging on the pinna. They pass through the outer and middle ear and get enhanced in the process, before arriving at the cochlea. The basilar membrane, a part of cochlea is responsible for decoding the acoustic stimulus. It acts as having a series of band pass filters, called the critical bands which extract frequency, amplitude and duration information from the signal (Moore, Glasberg & Baer, 1997). It also bears the property of an automatic gain control (AGC) which manipulates the loudness to be comfortable (Rhode, 1971). This signal is then converted into nerve impulses and passed further to higher nuclei and finally to the auditory cortex for us to perceive and comprehend what has been said.

When the stimulus to be processed is speech, the auditory system must be capable of extracting certain cues that are essential for perception of any speech sound. They are, the faster oscillations called temporal fine structure and the relatively slowly varying envelope. It has been hypothesised that coding of this fine structure across the basilar membrane is by place coding, and that of the envelope is through phase locking by neural fibres (Rose, Brugge, Anderson & Hind, 1967; Joris & Yin, 1992).

It is therefore evident that, in order to be able to hear and understand speech completely, a normal functioning of the cochlea and auditory pathway is essential. Owing to the same, a normal hearing individual can perceive speech easily. However, the perception of speech becomes difficult in the presence of background noise, in a reverberant condition or when more than one person is talking. The reason behind this is that the available spectral contrast and other temporal cues are reduced due to the presence of a competing signal (Brungart, Chang, Simpson, & Wang, 2009; Brungart et al., 2001; Freyman, Balakrishnan, & Helfer, 2004). However, it is only relatively hampered in these conditions for a normally hearing individual due to the physiological compensations made by a normally functioning auditory system. The advantage is that, they can still adaptively listen to a specific sound in a concoction of multiple talkers by focusing on the required stimulus. This phenomenon is widely known as the "cocktail party" effect (Cherry, 1953; Yost, 1997).

Any alteration in the structure of the auditory system and its physiology will result in hearing loss which could be of conductive, sensorineural or mixed in nature. In a conductive type of hearing loss, the speech perception can be restored by compensating for the loss of audibility since conductive mechanism is understood to be non-analytic in nature. However, in a sensorineural or mixed type of pathology, the perceptual consequences are different and cannot be compensated the way it can be done in a conductive type of pathology as they also involves damage to the cochlear and/ or auditory nerve. Sensorineural hearing loss, which results from disturbances in the cochlea, is the most common type of pathology seen. The most often cited complaint from them is failure to comprehend speech, especially in noisy or reverberant conditions, or when more than one person speaks. The magnitude of problem is likely to increase with severity and change in configuration of hearing loss (Hornsby, Johnson & Picou, 2011).

Cochlear hearing loss (CHL) is most often a consequence of loss or damage of outer hair cells (OHCs). OHCs are known to sharpen the auditory filters which in turn help in finer frequency discrimination. Abnormal structure or functioning of these cells would result in widening of auditory filters, as a result of which, many kinds of perceptual consequences can arise. They include, impaired frequency resolution, temporal resolution, and reduced sensitivity to low level sounds (Glasberg & Moore, 1986; Tyler, 1982; Thibodeau & Van Tasell, 1987; Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006).

While decoding speech, the auditory system represents its spectral shape as excitation pattern which resembles a slightly smoothed version of the input spectrum. Since individuals with cochlear pathology exhibit broader tuning curves, they produce a highly smoothened representation of the spectrum which makes it difficult to perceive fine acoustic information of speech. The reduced frequency resolution in cochlear pathology results in impaired discrimination of formant frequencies of vowels and consonants. On the other hand, reduced temporal resolution fails to quantify the subtle duration differences and hence results in poor discrimination of speech sounds (Schorn &

Zwicker, 1990; Lorenzi et al., 2006). This problem worsens in the presence of background noise. Hearing impaired individuals require a higher signal to noise ratio (SNR) to understand speech when compared to normal hearing listeners (Glasberg & Moore, 1989). SNR value usually ranges from 2.5 dB for mild to up to 7 dB for moderate to severe degree of Sensorineural hearing loss. They also fail to make use of spectral dips in fluctuating background noise unlike normal hearing listeners, who are able to perform better in steady state noise (Festen & Plomp, 1990).

Further, the configuration of hearing loss also has an effect on speech perception. It has been well established that different speech sounds comprise of different frequencies. For example, nasals (/m/ & /n/) are more low frequency dominant while, fricatives and affricates are of higher frequencies. Therefore, a person with a sloping configuration would tend to miss cues to perceive high frequency speech sounds and vice versa in case of a rising configuration.

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

To account for the consequences of cochlear pathology, the currently available hearing aids have features like, channels with band specific amplification, frequency transposition, etc. Although these features improve speech perception in individuals with cochlear hearing loss (Kompis & Dillier, 1994; Robinson, Baer & Moore, 2007), they fail to preserve the normal cochlear function as they do not restore frequency resolution. In fact, multiple channels might in turn introduce distortion in the signal as it passes through several filters. Their inability to recombine a signal across several channels may also distort the acoustic signal reaching the cochlea (Moore & Glasberg, 1993). To specifically overcome perceptual problems in noise, advanced hearing aid technology also provides different noise reduction strategies. However, these technologies prove effective only if noise and signal are spatially separated. If they arrive from the same direction, they get attenuated equally and hence do not result in noise reduction or an improvement in speech perception.

Thus, the current trends in hearing aids may provide adequate gain in quiet situations but are unable to give sufficient improvements in speech perception in the presence of background noise or in reverberant conditions. This remains as one of the biggest concerns in today's rehabilitation professionals. Therefore, to address these problems, several experiments with acoustical modifications in speech signal were carried out to improve the speech perception ability of individuals with cochlear hearing loss (CHL) and auditory neuropathy.

Enhancing the signal by digital signal processing of speech which manipulated the short term spectrum of words embedded in speech shaped noise showed a small but statistically significant improvement in intelligibility of the same for moderate to severe hearing-impaired listeners (Simpson, Moore & Galsberg, 1990). Similarly, signal enhancement with syllables and sentences also showed benefits in speech perception in quiet and in noise (Bunnell, 1990; Baer, Moore & Gatehouse, 1993; Lyzenga, Festen, & Houtgast, 2002). When, spectral enhancement was compared with and without phonemic compression, results showed an improvement in spectral enhancement alone condition for recognition of phonemes in cochlear loss with sloping configuration (Franck, Van Kreveld-Bos, Dreschler, & Verschure, 1999).

One of the other popular techniques is clear speech. This technique showed substantial benefit in both normal hearing listeners and hearing impaired listeners of various degrees and configuration in both quiet and noise (Picheny, Durlach & Braida, 1985; Bradlow & Bent, 2002). However, Ferguson and Port (2002) indicated no benefit in vowel discrimination in elderly hearing impaired with mild to moderate sloping hearing loss although benefit with sentences was consistent with previous studies.

In population like auditory neuropathy, signal enhancement techniques like envelope enhancement and lengthened transition duration proved beneficial in consonant identification. This performance was attributed to an improvement in the spectral and temporal contrast in speech (Narne & Vanaja, 2008; Kumar & Jayaram, 2011).

Despite the above discussed solutions, the currently existing technology has not been completely successful in restoring normal speech perception (Stone & Moore, 2003), especially on the frequency and temporal resolution domain. Hence, there is a need to analyse the shortcomings of the existing strategies and carry out research to modify them and evaluate its usefulness.

Need for the study:

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

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

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

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

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

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

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

Objectives of the study

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

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

Chapter 2

Literature Review

Speech, the key element in verbal communication, can be understood as a string of complex signal with intricately varying intensity and frequency over time. It is made up of different combinations of vowels and consonants which are the nucleus of speech. Vowels are voiced sounds made with an open oral cavity while consonants are either voiced or unvoiced and are made with a constriction at various levels by the articulators to produce different speech sounds like, labial, dental, alveolar, etc. It is believed that consonants cue primarily about the lexicon, whereas vowels cue about the syntax and add energy to the content (Toro, Nespor, Mehler & Bonatti, 2008). Hence, it is clear that, vowels and consonants differ in terms of their production, function and their acoustic features as well.

When we hear speech, it undergoes a series of processes before it is finally encoded and perceived as a meaningful utterance at the cortical level. This signal, after being acoustically enhanced by the outer and middle ear, is primarily processed at the level of the cochlea or basilar membrane before it is transmitted in the form of neural impulses through the rest of the auditory pathway. In order to perceive a speech signal, the auditory system makes use of certain acoustic features extracted from speech that serve as cues. They are the rapid temporal fine structure cues and the slowly varying envelope superimposed on it. LaRiviere, Winitz, and Herriman (1975) reported that the fine structure cues correspond mostly to place features like the consonant release burst and frication noise in obstruents, while Dorman, Studdert- Kennedy and Raphael (1977) stated that onset frequency location of the formants and the resulting formant transitions in sonorants and obstruents correspond to place features as well. On the other hand, the envelop helps extract manner cues like the duration of utterance and the rise time at the onset of the consonant (e.g., /ʃ/-ʤ/-/d/). Rosen (1992) and Stevens, Blumstein, Glicksman, Burton and Kurowski (1992) reported that vowel identity (e.g., /u/- /U/) and consonant voicing are also derived from the envelope. Put in simpler terms, each of these components is responsible for providing segmental cues. For example, envelope mainly cues for manner, tempo, rhythm and syllabicity while periodicity and fine structure majorly account for voicing, stress, intonation and place, voice quality. Hence, collectively, they represent the overall speech signal. A person with normally functioning auditory system, can therefore detect, decode, process and comprehend all the acoustic features to understand speech.

In one's daily routine, a person comes across more than just a quiet or a non-reverberant situation during communication through speech. For example, conversing at market, party, railway station, etc. At such places, the background noise is generally uncontrollable. In such situations, degradation of the available cues for speech perception is seen depending on the type, property and level of the background noise. Hazarati, Sadjadi, Loizou and Hansen (2013) hypothesised that reverberation, can cause two types of spectro- temporal smearing- self masking, caused by early reflections and overlap masking, caused by late reflections. The early reflections smear energy within each phoneme while the late reflections cause temporal smearing of high energy phonemes that mask their succeeding sounds. However, Nabelek, Letowski and Tucker (1989) stated that noise masks the weak consonants to a greater degree than the high intensity vowels, but unlike reverberation, this masking does not depend on the energy of the preceding segments.

When a background noise is introduced, the spectrum of the noise masks those of the speech, making only those spectral peaks available that are above the level of noise. Eisenberg, Dirks and Bell (1995) compared syllable identification scores of normal hearing individuals with that of mild to moderate sensorineural hearing loss individuals in various types of noises. They found that, normal hearing listeners are able to extract these cues and understand speech better than their hearing impaired counterparts. They performed better in the presence of amplitude modulated noise than a steady state noise. Similar results pertaining to normal hearing listeners having an advantage in speech perception in noise were reported by other authors as well (Pekkarinen, Salmivalli & Suonpaa, 1990; Nabalek, 1988).

In case of any alteration to the normal physiology of the auditory system, there can be impairment in terms of conductive, sensory, neural or mixed hearing loss. While conductive pathology reduces the overall audibility of the signal reaching the inner ear, sensory or neural hearing loss will lead to deficits in processing and transmission of the signal to the auditory cortex. Hornsby, Johnson and Picou (2011) examined the effect of type and degree of sensorineural hearing loss (SNHL) by obtaining speech perception scores for Connected Speech Test (CST) in noise in individuals with sloping and rising configuration with various degrees of losses. The test material was assessed in multiple low and high pass, band pass and wideband filter conditions. Speech was also amplified to increase audibility based on audiogram of participants. Based on the linear regression made, the authors attributed degree of hearing loss to be the strongest predictor for speech performance. The configuration of hearing loss affected the availability of cues in the region where they had elevated thresholds.

2.1 Cochlear hearing loss

This is a term given when there is an alteration in the function or structure of the cochlea leading to a pathophysiology that is different from normal hearing individuals. The damage is often caused to the outer hair cells (OHCs), inner hair cells (IHCs), stereocilias which get distorted, stria vascularis, or, the entire hair cell might die (Bohne & Harding, 2000; Mac Mohan & Patuzzi, 2002). Khanna and Leonard (1982) suggested that loss of OHCs result in partial or full loss of active mechanism and other non-linear functions which are the main functions of cochlea. This leads to widening of auditory filters, disturbing the first step of signal analysis- resolving into different frequency bands. Following this, the input output function, which is responsible for differential gain across frequency bands, loses its non-linearity resulting in loss of audibility to low level sounds and recruitment at high level sounds which increase with increase in amount of damage. Due to broadening of tuning curves, a range of perceptual problems are seen like, reduced frequency selectivity and temporal resolution (Kiang, Moxon & Levine, 1970; Glasberg & Moore, 1986). Hence, their ability to extract the envelope, periodicity and temporal fine structure (TFS) cues are compromised due to which there is a loss of segmental and suprasegmental information reaching the cortex (Rosen, 1992).

Psychophysical Tuning Curves (PTCs) are the most commonly used tool to measure frequency selectivity. A number of studies have compared PTCs in normal hearing individuals and those with cochlear pathology and have gone to establish that PTCs in cochlear hearing loss are broader (Florentine, Buss, Scharf & Zwicker, 1980; Stelmachowicz, Jesteadt, Gorga & Mott, 1985; Kluk & Moore, 2005). Carner and Nelson in 1983 have also showed different types of PTCs using simultaneous masking. They are, flat, erratic, broad and inverted patterns and they indicate variability amongst the group of cochlear hearing loss itself.

To understand the role of TFS in the perception of speech in hearing impaired, Lorenzi, Gilbert, Carn, Garnier and Moore (2006) generated a speech stimuli consisting of only TFS and no envelop cues. Consonant identification task in this stimulus by young and elderly normal hearing listeners and individuals with moderate flat CHL showed that individuals with moderate flat CHL could perform almost at par with normal with unprocessed and envelope cues but they scored very poorly with only TFS cues. Ghitza (2001) and, Hopkins and Moore (2007) also indicated the inability of individuals with CHL to extract TFS cues from a speech signal. Evidence from Lorenzi, Debruille, Garnier, Fleuriot and Moore (2009) showed deficits in the ability of individuals with high frequency hearing loss to use TFS cues that occurred within their normal hearing thresholds (i.e., less than 2 kHz), as compared to that of normal hearing individuals. All the above mentioned literature points to the fact that their impaired capacity to excerpt TFS cues has a profound impact on their speech perception. According to the authors, this was probably more severe in the presence of noise, as dip listening was disrupted

(Schooneveldt & Moore, 1987; Moore & Glasberg, 1987). In quiet situations, only envelop cues were sometimes sufficient to understand speech in individuals with cochlear hearing loss (Shannon, Zeng, Kamath, Wygonski, and Ekelid, 1995; Qin & Oxenham, 2003; Stone & Moore, 2003).

Another problem faced by individuals with CHL is, reduced phase locking abilities. This may be attributed to the alteration in travelling wave properties owing to the presence of cochlear pathology. This in turn, affects the TFS processing by higher centres (Leob, White & Merzenich, 1983).

In individuals with cochlear hearing loss, the effect of masking is strong. To evaluate the same, Dubno and Schaefer (1995) measured thresholds in the presence of notched noise and narrow band noises in individuals with mild to moderate type of hearing loss and normal hearing adults. Spectrally shaped broadband noise (SSBB) was used to make the thresholds of clinical and control subject pairs equal. They noted an increase in masked threshold in the clinical group as compared to the control group in frequencies above the masker pass band. These results were consistent with their previous study in 1991. Trees and Turner (1986) also observed the effect of masking was high especially in those with high frequency hearing loss. They suggested that a person with hearing loss at high frequencies may be unable to hear some high frequency sounds, even at high intensities, in the presence of a low frequency masker. This very well mimics the situation of perceiving speech in the presence of noise as the composition of noise is generally believed to be dominant in the low frequency region. Therefore, due to the above mentioned perceptual consequences of cochlear hearing loss, speech perception is also affected.

2.2 Speech perception in individuals with cochlear hearing loss

Nabelek and Pickett (1974) reported that the ability to perceive speech in individuals with cochlear hearing loss is largely dependent on the listening environment as it affects their control over the cues available. It was therefore, highest at favourable signal to noise ratios (SNRs) and decreases as a function of reduction in SNR.

The altered perception of various features of speech due to the presence of cochlear hearing loss has been well established in literature. Boothroyd (1984) showed that these individuals follow supra segmental cues better than segmental. Within segmental, vowels are better perceived than consonants. Similarly, consonant voicing and continuance are easier to estimate than consonant place and, word initial consonants are comprehended better when compared to word final.

2.2.1 Perception of vowels

Vowels are speech sounds associated with a steady state articulatory configuration and a steady state acoustic pattern. They are usually voiced and thus are high energy speech sounds. Tasell, Fabry and Thibodeau (1987) showed that vowel perception is more affected in people with cochlear pathology in the presence of a masker, than in an individual with normal hearing.

Boothroyd in 1984 examined vowel perception in children with profound hearing loss. These children with profound hearing loss were able to distinguish between vowels that differed in their Formant 2 (F2) frequencies, such as /u/ and /a/. However, they were able to do it only if they had measurable high frequency hearing. In children with very poor high frequency thresholds, they could distinguish only between vowels varying in Formant 1 (F1) such as, /a/ and /i/. The author explained that children with very poor high frequency thresholds depended upon F1 as a cue and perceived vowels mainly on the basis of relative intensity and duration. It also suggested that although vowel perception could be affected in individuals with cochlear hearing loss, they are less affected as compared to consonants.

A study by Richie, Kewley- Port and Coughlin in 2003 examined the discrimination and identification of vowels in two conditions, namely, frequency specific gain and flat frequency response conditions. The population chosen was individuals with mild to moderate sensorineural hearing loss. Results proved that, even in loud listening condition, there was no difference between identification and discrimination tasks. It indicates that the perception of vowels is not hampered by upward spread of masking in these individuals. Though, they noticed a higher likelihood of confusion between /e/ - /i/ and /e/- /ɛ/ because of similar spectral characteristics. Liu and Kewley- Port (2004) further proved that although upward spread of masking of F2 by F1 is not evident in up to moderate hearing loss group, it is more pronounced in listeners with severe SNHL and hence results in poorer perception of vowels.

In a study done on cases with moderate degree of hearing loss with sloping configuration, Coughlin, Kewley-Port and Humes (1998) found that F2 discrimination was significantly poor in hearing impaired listeners at lower presentation levels.

However, given at higher presentation levels discrimination ability improved. The authors attributed this to their reduced audibility at high frequencies due to which F2 was not audible at low levels but was compensated at high levels. Although F1 and F2 were audible, these individuals had impaired formant identification which manifested in confusion between vowels having similar spectral characteristics.

To sum up, affected vowel identification in hearing impaired listeners is observed in severe degree and above. Due to unavailability of high frequencies in most cases, they rely on F1 cues the most which sometimes leads to confusions amongst vowels with similar spectral characteristics.

2.2.2 Perception of consonants

Consonants are sounds whose acoustic characteristics are defined by subtle cues which can be grouped into place of articulation, manner of articulation and voicing. Owing to their widespread frequency composition, they could be affected differently according to the severity and configuration of hearing loss (Walden & Montgomery, 1975).

Johnson, Whaley and Dorman (1984) studied the location of phonetic boundaries of voice onset time (VOT) as a cue for place of articulation (POA) by varying it in a continuum in three-formant CV syllables. They found that, the results for VOT were on par with normal hearing persons for those with mild and moderate SNHL whereas, severe and profound hearing losses had abnormal VOT boundaries. The authors partially attributed this to the high level of presentation in high degree of hearing loss which could have distorted the signal. Revoile, Pickett and Holden (1982) found similar results with place of articulation cues like burst and formant transition. Although they were present in hearing losses up to moderate degree, they were weakly defined. Their utility as cues depended on the audibility which varied to an extent from one person to another. However, as noted by the authors, the voicing errors within the hearing impaired group remained consistent across any degree of loss. It depended on the availability of preceding vowel duration and murmur as cues for perception of stops.

Manner of articulation cues, as reported by Subtenly (1983), were found to be poorest for fricatives followed by stops. Nasals and glides were found to be perceived best when evaluated in 160 individuals with CHL of different degrees using a consonant identification task. Perception of fricatives was affected in any degree of hearing loss including mild degree. They found that even though steady state spectral cues in the frication portion were available sometimes for HI group of subjects, transition cues that usually assist a normal hearing subject to distinguish place information especially at low levels, were not always available due to audibility factor. Hence, they made errors in recognizing voiceless fricatives.

Tsui and Ciocca (2000) reported that, when the degree of hearing loss was greater than moderate- moderately severe, place of articulation for stops did not serve as a cue due to reduced audibility of low intensity burst and formants above F3. Reduced frequency selectivity and upward spread of masking from high intensity F1 on other formants resulted in the reduced ability to process rapid frequency changes contained in formant transitions and release burst in individuals with CHL.

To evaluate the effect of configuration of hearing loss on utilization of short duration spectral cues, Dubno, Dirks and Schaefer (1987) evaluated plosive recognition in a CV combination by reducing the duration of the synthetic CV in a continuum from 300 to 10 ms. The population considered was 10 normal listeners and 15 individuals with SNHL of varying configuration. They found that, a flat and gradually sloping configuration had similar scores for recognition of /b/, /d/, and /g/ although the overall scores were poorer than normal. Authors hypothesised that they used the onset spectrum as cues for place of articulation. Identification of voicing showed significant improvement between 10-30 ms increase in preceding vowel duration. Improvement was observed with /a/ and /u/ but not /i/. On the other hand, a steeply sloping loss resulted in high scores for /g/ and lower or equivalent scores for /b/ and /d/. The authors attributed this to the pronounced spectral peak in onset spectra for /g/ which was more resistant to changes than rising/falling spectra for /b/ or /d/. They use later occurring cues like formant transition as the shape of the onset spectra was distorted. Voicing results were similar to that of flat or gradual sloping group.

To study the relation between speech perception and region of cochlear damage, Preminger and Wiley in 1976 measured psychophysical tuning curves (PTCs) for 500 Hz and 4000 Hz in subjects with cochlear hearing loss of rising, falling and flat configurations having two subjects in each group. The stimulus used to study consonant identification was divided into sounds of low frequency (/w/,/b/,/m/,/l/), high frequency (/t/,/d/,/k/,/z/,/s/) and diffuse (/v/,/f/). Subjects with high frequency hearing loss had broadened PTCs at 4 kHz but normal at 500 Hz and hence, they performed well for low frequency consonants but not high frequency ones. Out of the two subjects with flat hearing loss, one of them had relatively better PTC at 500 Hz and this subject had better perception scores than the other. Rising configuration with poor PTCs at 500 Hz however, showed equivocal results. The authors attributed the reason to possibility of variation in the integrity of the cochlear regions associated with this configuration. Another reason, according to the authors, was greater impairment of temporal mechanism (encodes for low frequency) than place mechanism (encoded for high frequency). Hence, they concluded that, frequency selectivity alone might not be able to clearly explain the errors in speech perception.

2.2.3 Speech perception in quiet

The development of research in this area began with researchers believing that loss of audibility was the main cause for reduced speech perception. Authors like Zurek and Delhorne (1987) and Lee and Humes (1993) argued that their pure tone threshold being higher than normal led to portions of speech spectrum lying below their audible range. They mainly attributed perception difficulties to these speech sounds not being heard. Simultaneously, a larger group of researchers (for example, Baer & Moore, 1993) expanded on this thought and added loudness recruitment. Other researchers like Plomp (1978) and Glasberg and Moore (1989) added reduced frequency selectivity, and temporal resolution as contributing factors as well.

2.2.4 Speech perception in noise

Pekkarinen, Salmivalli, and Suonpaa (1990) studied word discrimination scores in quiet and in the presence of pink noise in individuals with normal hearing, conductive pathology and CHL. They reported that although**,** speech perception for one with cochlear hearing loss was better in a quiet situation, it was still not always at par with normal hearing individuals. They suggested that it partly depended on the severity of hearing loss. While those with mild to moderate hearing loss performed reasonably well in quiet situations, people with severe to profound loss faced a lot of difficulties even in quiet conditions. This problem in perception of speech further increased to a large extent in the presence of background noise.

It has been well established in literature that individuals with CHL find it challenging to understand speech when background noise or reverberation was introduced. Many researchers like Eisenberg et al. (1995) and Pekkarinen et al. (1990) have suggested that an interaction between reduced audibility, poor frequency selectivity, temporal resolution, and recruitment is the reason for deterioration in speech perception in noise in individuals with CHL and that the independent contribution of these effects vary with degree and configuration of hearing loss.

a. Reduced audibility

Humes and Roberts (1990) compared the closed set responses for CV syllables in quiet, background noise and reverberant conditions for populations of young nor hearing, elderly hearing impaired and young normal hearing with simulated hearing loss to match the group of hearing impaired. By performing correlational analyses and by using an acoustic index with adjustments for threshold elevations, they concluded threshold elevation as the primary determinant of speech perception in elderly individuals with hearing impairment in all listening conditions. The authors attributed this result to missing out of information at those frequencies where their threshold of audibility was poor whereas, in the presence of background noise, even if those cues were audible, they were masked by the noise making it unavailable for perception.

Bronkhorst and Plomp (1989) mentioned another way where loss of audibility could result in poor perception of speech in the presence of noise. In a condition with spatially separated background noise, head shadow effect provided a natural advantage with improved SNR in one ear, especially at high frequencies. Their study revealed that hearing loss at high frequencies would deteriorate the advantage that could have been obtained by this phenomenon.

Another study by Moore and Glasberg (1993) examined the effect of loss of audibility as an independent factor by simulating three groups of cochlear hearing loss-moderate flat, severe flat, and moderate to severe sloping. A linear amplification with National Acoustic Laboratory (NAL) prescription in these populations showed good intelligibility in quiet conditions but deteriorated when a competing talker was introduced. It also indicated that linear amplification could improve speech intelligibility in flat configuration but was less effective for sloping type of simulation. It was noted that, as the gain was increased to compensate for loss of audibility, the signal became unpleasant before they could obtain speech scores comparable to normal. The authors attributed the combined effect of reduced audibility and loudness recruitment for this finding. Similar results were obtained by Festen and Plomp (1990) who attributed the poor performance of listeners with CHL in both modulating and steady state noise to their inability to make use of spectral dips, which, normal listeners are capable of utilizing.

This was because of reduced audibility in cochlear pathology. These participants performed poorer than estimated by articulatory index (AI). Hence, the authors concluded that, factors apart from reduced audibility also contribute to reduction in speech perception in noise.

b. *Frequency selectivity*

To understand the effect of widened auditory filters in individuals with CHL, Baer and Moore (1993) quantified speech perception using sentences in both quiet and in the presence of speech shaped noise. The population they considered were that of normal hearing listeners with simulated frequency selectivity of varying degrees of hearing loss as described by Moore, Glasberg and Simpson (1992). They found that in quiet condition, the speech intelligibility was minimally affected by spectral smearing even when the simulated auditory filter was six times broader than that in normal. Though, when noise was introduced, speech intelligibility was adversely affected, especially for low SNRs and large degrees of smearing. To explore further, the same authors in 1994 examined the perception of smeared speech in a single competing talker and steady background noise. Results indicated that speech is more affected in the former condition than the later. The authors attributed this finding to a reduced frequency resolution ability which manifested as an inability to separate speech frequencies from that of background noise.

c. *Loudness recruitment*

As discussed earlier, Moore and Glasberg (1993) studied the isolated effect of elevated thresholds and loudness recruitment. They measured speech perception in quiet and in noise by simulating higher thresholds with recruitment in normal hearing individuals. The results of quiet situation were better when the signal level increased. In a single talker background, the subjects required an increase of 13dB in the target signal to yield scores on par with normal. In speech shaped noise however, 6dB increase in SNR was required for the same (Moore, 2007). The difference in results because of the difference in background noise was explained by the authors in terms of the cues available. Individuals with a normal auditory system were able to make use of the spectral and temporal cues in a single talker background. Hence, owing to dip listening, their perception was better in fluctuating noise than in steady noise. As dip listening required a wide dynamic range, listeners with CHL were unable to take advantage of the fluctuation in noise. Small dynamic range caused intense parts of speech to be too loud and weaker parts to be almost inaudible. Therefore, they concluded that loudness recruitment too could be an influential factor in speech perception.

To summarize, the major contributing factors for speech intelligibility in listeners with CHL can be regarded as a combination of reduced frequency, loudness recruitment and loss of audibility. The effect of this combination varies across degrees of hearing loss. Any rehabilitation approach aiming at improving speech perception in these individuals must address the above mentioned issues.

2.3 Strategies to improve speech perception in noise in individuals with cochlear hearing loss

Skinner (1980) suggested that amplification in order to restore audibility can sometimes lead to an increase in speech understanding. However, if the hearing loss is very high, excess amplification would not be beneficial, or could even lead to deterioration of speech understanding. Therefore, to improvise on the previously existing technology, a number of advancements were evaluated in hearing aids. The benchmark advancements in the history of amplification devices include directional microphones, compression, frequency transposition and signal enhancement.

Kompis and Dillier (1994) evaluated a directional signal processing approach with and without a combination of directional microphones and adaptive beamform. Results indicated that a combination method proved more beneficial in noise, when compared to any of the two approaches independently, for intelligibility of words in both normal-hearing and hearing-impaired subjects.

Lippmann, Braida and Durlach (1981) suggested the use of compression in hearing aids after obtaining a benefit in individuals with sensorineural hearing loss with CVCs and sentences as stimuli. Robinson, Baer and Moore (2007) evaluated the benefit of frequency transposition in hearing aids by improvising and presenting the signal in a simulated output of hearing aid. The authors transposed the frequency components well within the dynamic range (DR) to a region just within the DR without applying compression. This was done in subjects with high frequency sensorineural hearing loss. To assess the performance, VCVs were used as stimulus. Only after a period of training, they found a significant improvement for identification of word final ‘s’ and ‘z’ in two out of seven subjects. The authors attributed this improvement to the availability of cues that were otherwise lost in regions of poor audibility. However, they suggest that training is essential to show any benefit from this technique.

However, Plomp in 1988 brought to notice two features of sensorineural hearing loss, attenuation and distortion. Although hearing aids were able to partially overcome the audibility factor, the distortion caused by reduced frequency selectivity still remained an issue. Hence, several advancements were made in technology which aimed at spectrally modifying the signal itself which can be implemented in hearing aids.

Baer, Moore and Gatehouse (1993) suggested that by pre-processing the signal by enhancing spectral contrasts, the problem of reduced frequency selectivity and temporal selectivity can partially be overcome. The enhanced signal increases the prominence of the peaks which can help in producing an excitation pattern that closely resembles the excitation pattern evoked in a normal auditory system by an unprocessed signal. The disturbance of temporal patterns due to the presence of noise through broadened filters can be overcome as the processed signal enhances those portions of the spectrum where the signal-to-noise ratio is highest (the peaks) and suppresses those where it is lowest (the valleys). They found that the enhanced signal resulted in better intelligibility and quality of signal no significant improvement in speech perception. The authors reported that speech identification scores for enhanced signal with moderate degree of compression improved in the control group with practice for a few days. Hence, they suggested the same for clinical group as well.

Simpson, Moore and Glasberg (1989) described a method of digital signal processing of speech which would help increase the spectral contrasts in the spectrum. They manipulated the short term spectrum of speech in noise by calculating an auditory excitation pattern from the magnitude spectrum of overlapping short segments of the speech signal which was then filtered according to the normal auditory filter bandwidths. Magnitude values from this enhanced pattern were combined with the unchanged phase spectrum from the original signal to produce the enhanced speech which was combined with speech shaped noise at SNRs between -3 and 6 dB and presented. The population considered was moderate cochlear hearing losses who were asked to identify sentences in speech shaped noise without any amplification devices. Results showed small but statistically significant improvement of 6% to 7% for processed speech. They attributed this improvement to the increase in spectral contrast that was brought about by this technique and suggested further studies to be carried out to implement the technique in hearing aids.

Picheny, Durlach and Braida (1985) evaluated the benefit of clear speech. This technique showed substantial benefit in both normal hearing listeners and hearing impaired listeners of various degrees and configuration in both quiet and noise using sentences. However, Ferguson and Port (2002) indicated no benefit in vowel discrimination in elderly hearing impaired with mild to moderate sloping hearing loss although benefit with sentences was consistent with previous studies.

Despite the discussed approaches and solutions the results so far, in terms of improved speech intelligibility, have not been particularly promising (Moore, 2003), especially on the frequency and temporal resolution domain.

2.4 Effect of companding

Based on the idea of relatively broadband compression followed by more frequency selective expansion which is seen in auditory system, Turicchia and

Sarpeshkar (2005) proposed a strategy for time domain spectral enhancement. This technique combines two-tone suppression and dynamic gain control in order to increase the spectral contrast. Following this, Oxenham, Simonson, Turicchia, and Sarpeshkar (2007) examined the advantage of companding for cochlear implant (CI) listeners by simulating CI processing in normal hearing individuals. They observed that companding improved sentence perception scores by 10 – 20% in steady-state noise. The improvement observed was attributed to enhanced spectral and temporal cues in the speech signal. These results were supported by Bhattacharya and Zeng (2007) who compared the benefit of companding in normal individuals and CI users using phoneme and sentence recognition tests. Their tests revealed that normal hearing individuals showed an improvement in noise condition for vowel perception but not consonants and sentences. CI users however, demonstrated a benefit in in both phoneme and sentence recognition in noise. The authors indicate that, apart from an improvement in spectral contrasts, this technique also enhanced the temporal cues which sufficiently contributed to the increased scores in speech perception.

Narne, Barman, Deepthi and Shachi in 2014 evaluated the benefit of companding in individuals with Auditory Neuropathy Spectrum Disorder (ANSD) using consonant and sentence identification task in noise and noticed an improvement in both sentence and consonant identification tasks in quiet and higher SNRs. They too attributed the benefit to enhanced temporal and spectral representation of signal which helped individuals with ANSD to code the cues better.

A similar study was done in individuals with cochlear hearing loss by Deepthi (2012). She checked for benefit from companding in syllable identification in 0, 10 and 15 dB SNR and sentence identification from -10 to +20 dB SNR. SNR50 was also found for sentences using speech shaped noise. The outcome measures indicated a significant improvement in consonant recognition at 0, +10 and +15 dB SNRs in processed conditions compared to unprocessed, indicating a benefit from this strategy. A greater benefit from processed signal was seen at 0 dB SNR than at +15 dB ANR. The author explained the improvement with respect to the availability of the enhanced spectral and temporal cues in individuals with reduced frequency selectivity.

2.5 Effect of consonant enhancement

Fletcher (1953) reported that vowels usually contained more acoustic energy than consonants. To make both components of speech available, Guelke (1987) suggested that increasing the burst amplitude of a stop consonant up to the level of the vowel seen in normal speech. According to the author, this could result in increased intelligibility. This is because information that would have otherwise been inaudible to a hearing impaired person is made available by increasing the burst amplitude. 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 their perception of speech. This finding was also supported by Baer, Moore and Gatehouse (1993) who showed an improved subjective quality and intelligibility rating for consonant enhanced stimuli.

Bunnell (1990) studied signal enhancement by spectrally amplifying the consonants in CV syllables in population of moderate to severe SNHL. The author observed a moderate improvement in stop consonant recognition with this technique and attributed the results to an increased spectral contrast in the stimulus.

Franck, Van Kreveld-Bos, Dreschler and Verschure (1999) studied the separate and combined effects of spectral enhancement with phonemic compression with CVCs and words as stimuli in individuals with sloping SNHL. They showed a benefit in vowel perception with spectral enhancement only. However, they also indicated that a combination of enhancement and compression deteriorated speech perception.

A study by Hazan, Simpson and Huckvale in 1998 examined speaker and listener effects with native and non-native speakers using CVCs in only quiet conditions. They differentially varied the gain for vowel, consonant burst, frication and aspiration. Their results showed a significant improvement in processed conditions for all speakers and an increase in scores in the listeners.

The above discussed studies with respect to companding and consonant enhancement clearly indicate that these strategies improve the overall signal envelope by enhancing the spectral and temporal aspects of the signal. As, many authors like Moore et al., (1992 1993), Moore (2007) established the cause of poor speech identification of speech in persons with CHL as being reduction in spectral contrast and temporal modulations due to addition of noise, and these techniques aim at compensating for the lost spectral and temporal cues, there is a need to study more number of strategies in individuals with CHL.

Most of the previous researches with companding and signal enhancement have been done in cochlear hearing loss with high redundancy stimuli like words and sentences in very few SNRs. The most often used background competing signal was noise which does not provide an accurate estimate of the noise encountered in real life situation. Also, the literature does not report of any study that has evaluated both the techniques in same population. Hence, the present study was taken up to overcome these limitations and thus, evaluate the benefit of companding and consonant enhancement strategies in individuals with normal and CHL in various listening conditions.

Chapter 3

Method

The current study was taken up to see the effects of processed speech through companding and consonant enhancement on speech perception at five different stimulus environments. To evaluate the same, two groups of participants with the clinical group having two subgroups were considered. The following procedure was administered to study the objectives mentioned.

3.1 Participants:

The participants selected were divided into two groups, the clinical and the control group. Those with cochlear hearing loss belonged to the clinical group while those with a normal auditory system were included under control group.

3.1.1 Clinical group (Individuals with cochlear hearing loss)

- Fourteen Adults (16 ears) were selected as participants of this group. All the participants of this group were diagnosed as having post lingual acquired cochlear hearing loss at the Department of Audiology, All India Institute of Speech and Hearing. All of them were native speakers of Kannada.
- The age range of the participants was from 23 to 55 years, with a mean age of 39.87. This age range was taken as it has been reported that psycho- acoustical abilities reach a plateau in normal hearing individuals by 12 years of age (Warner, & Gray, 1998).

- The degree of hearing loss of the participants ranged from mild to moderately severe sensorineural hearing loss with either flat or gradually sloping configuration.
- They had speech identification scores that were in proportion to their pure tone thresholds.
- All participants in the group had ‘A’ type tympanogram with ipsilateral and contralateral reflexes present elevated or absent depending on the degree of hearing loss.
- Auditory brainstem response patterns were as expected with the severity of hearing loss.
- Transient evoked otoacoustic emissions were absent indicating outer hair cell dysfunction in conjunction with tympanogram results.
- None of them had any history or presence of any other neurological or middle ear related problems as reported.

The clinical group was further subdivided in to two groups based on their configuration. They are as follows,

- ✓ Subgroup I (individuals with flat hearing loss)
- ✓ Subgroup II (individuals with sloping hearing loss)

3.1.1.1 Group I (individuals with flat hearing loss)

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

3.1.1.2 Group II (individuals with sloping hearing loss)

The participants of this subgroup included four adults (5 ears) in the age range of 28 to 55 years and with a mean age of 49 years. The degree of hearing loss ranged from mild to moderately severe sensorineural hearing loss with gradually sloping configuration (Stephen & Rintelmann, 1978). The demographic and audiological details of all the participants of cochlear hearing loss group are shown in table 3.1.

Table 3.1: Demographic and audiological details of subjects with cochlear hearing loss

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

3.1.2 Control group (Individuals with normal hearing sensitivity)

The participants of this group consisted of 14 individuals (14 ears) who were selected based on the following criteria,

- All the subjects had hearing sensitivity less than or equal to (four frequency average pure tone threshold, 500 Hz, 1000 Hz, 2000 Hz & 4000 Hz) 15 dB HL, with no history or complaint of difficulty in understanding speech in noise.
- ‘A’ type tympanogram with ipsi and contralateral reflexes were present in all the participants.
- All of their SPIN scores were 60% and above at 0 dB SNR.
- Normal auditory brainstem responses at 90 dBnHL were obtained in all the participants.
- Presence of otoacoustic emissions in both ears was observed in all the subjects.
- Through a structured interview, it was ensured that, none of them had any history or presence of any neurological, otological or any other associated problems.
- All the participants of this group were native speakers of Kannada.

3.2 Instrumentation:

The following instruments were used in the study,

- A two channel diagnostic audiometer, GSI-61 (Grason-Stadler Incorporation, USA) was calibrated with Telephonics TDH-50P supra aural headphones and Radio ear B-71 bone vibrator which was calibrated as per ANSI S-3.6, (2004), was used for threshold estimation and speech audiometry.

- In order to do tympanometry and estimate acoustic reflexes, a calibrated GSI-tympstar (Grason-Stadler Incorporation, USA) clinical immittance meter was used which was calibrated as per ANSI (1987).

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

- To record brainstem auditory responses, Intelligent Hearing Systems (IHS smart EP windows USB version 3.91) was used with AgCl electrodes and ER-3A insert earphones.

- For generation of stimulus, mixing the generated signal with noise and to process the same for temporo-spectral modification using companding, MATLAB- 7 (Language of Technical computing, USA) was used. It was also used for presenting the stimulus and recording the responses.

- The consonant portion of the syllable was enhanced (gain in dB) using UCL enhance version 101.exe (2002).

- All the CV syllables used in this study were recorded and normalized using Adobe Audition v5.

- Motu MicroBook II instrument connected with AHUJA AUD- 101XLR dynamic unidirectional microphone was used to record the stimulus. A Dell Inspiron 14R laptop (Realtek sound card) was used to present the stimulus.

- In order to control the intensity of presentation, the stimulus from the laptop was routed to a calibrated audiometer (MA 53). The output was delivered through Sennheiser HDA 200 headphones that were connected to the audiometer.

3.3 Stimulus generation

A set of 19 consonant- vowel (CV) non sense syllables in the context of the vowel /a/ (tʃ, ʤ, r, n, m, v, j, l, ḷ, s, ʃ, t̲, p, b, ṭ, d̲, ḍ, k, g) were digitally recorded. Non-sense syllables were considered in order to reduce the redundancy of the stimuli which would affect the perception scores. Each of these syllables was recorded thrice from an adult male native Kannada speaker. The recording was done in a sound treated room using a data acquisition system with 44.1 kHz sampling frequency and a 16 bit analogue to digital converter.

Following this, the recorded samples were given to five individuals with normal hearing sensitivity for goodness test. They were asked to perceptually rate the quality of the three samples of each syllable. A three point rating scale was used for the same with 1corresponding to poor intelligibility, 2 to fair intelligibility and 3 to good intelligibility. For each syllable, only the sample with the highest rating was chosen against the remaining two. This way, a set of recorded unprocessed 19 CV syllables with maximum relative intelligibility scores were considered for the study.

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

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

The resulting processed stimuli tokens using companding and consonant enhancement were labelled as ‘companded’ and ‘enhanced respectively, while the unmodified tokens were labeled as ‘unprocessed’. The signal was subjected to companding using MATLAB 7.8 and consonant enhancement using UCL Enhance software.

3.4 Procedure used for companding

The algorithm followed to carry out the spectro- temporal enhancement was based on that given by Turicchia and Sarpeshkar (2005). It was done using MATLAB- 7.8 software. The basis of this strategy is two tone suppression that otherwise happens in a normally functioning cochlea. At the psychoacoustic level, it is manifested as simultaneous masking which results in the reduction of one signal due to the presence of another in its frequency vicinity, thereby increasing the spectral contrast. Hence, a multichannel syllabic compression and spectral contrast enhancement through two-tone suppression forms the core idea of this strategy.

The incoming signal was first passed through 50 independent frequency channels by a bank of relatively broad band filters. The authors recommended the number of independent frequency channels to be 50 based on the dependence of degree of spectral enhancement on the number of channels. They found that, a lesser number would only enhance the stronger first formant as compared to the following weaker formants.

Every channel of the companding architecture had a relatively broad prefilter, a compression block; a relatively narrow band post filter and an expansion block. Following the initial filter, it was subjected to envelope detection (ED), whose output along with compression index (n1) having a value of 0.3 determined the amount of amplitude compression the signal underwent at this second stage. The EDs consisted of a full wave rectifier with a first order low pass filter. This compressed signal was then expanded after being passed through a relatively narrow band-pass filter. The gain of the expansion block depended on the corresponding ED output and the ratio of (n 2- n 1)/ n 1. The n2 parameter of the algorithm is the expansion index and had a value of 1. The outputs from all the channels were then non-linearly summed to obtain the processed signal. A block diagram of the above procedure is shown in Figure 3.1 and 3.2.

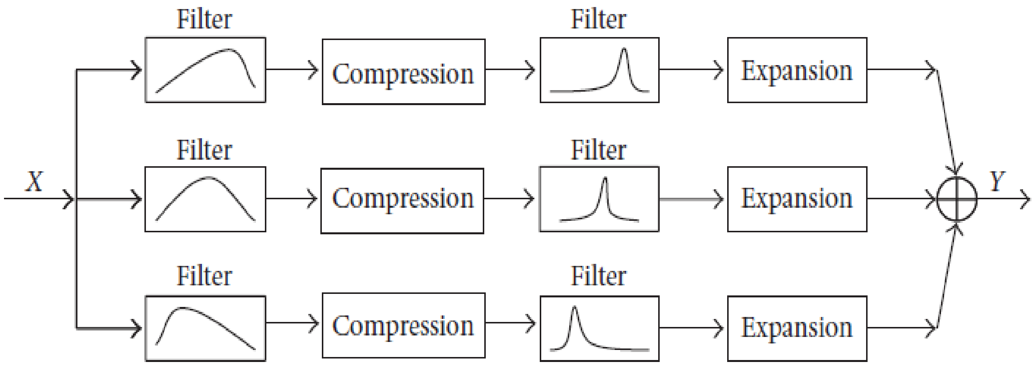


Figure 3.1: Block diagram of the companding architecture showing the stimulus being analysed by a bank of broad band prefilters. The output of each prefilter was then subjected to compression, and the output was filtered again using sharper postfilters before it was subjected to expansion. The outputs 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.

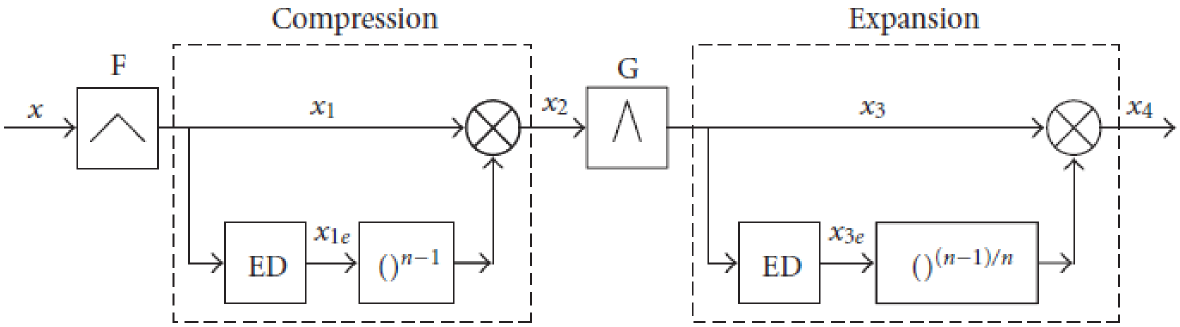


Figure 3.2: Detailed architecture of a single channel ED- envelope detector. Extracted from "A Bio-Inspire Companding Strategy for Spectral Enhancement", by L. Turicchia and R. Sarpeshkar, 2005, *IEEE Transactions on Speech and Audio Processing, 13*, p. 244.

3.5 Procedure used for consonant enhancement

This enhancement strategy was based on the rationale that, by increasing the spectral contrast, the otherwise weak consonants are made more perceivable in the presence of high energy neighbouring vowels. Taking the method adapted by Guelke in 1987 as reference, the procedure was carried out similarly. UCL enhance software was used to process the incoming signal with consonant enhancement technique. The procedure consisted of an algorithm that automatically identified the location of vowels, nasals, fricatives and gaps based on broad class phonetic recognition system. The algorithm then increased the amplitude of the selected portion of the syllable up to the specified level of the vowel in normal speech. The following options were selected for the enhancement of the stimuli:

Table 3.2: Table showing details of options chosen to enhance the consonant part of various consonants in the context of /a/

Syllable	Options chosen (among Burst, Fricative, Nasal and transition)	Enhancement level
Stops (p, b, ṭ, t̠, ḍ, d̠, k, g)	Burst + Transition	6 dB
Fricatives and Affricates (tʃ, ʤ, s, ʃ)	Fricative + Transition	6 dB
Nasals (m, n)	Nasal + Transition	6 dB
Glides (j, r, l, v, ḷ)	Transition	6 dB

For all the syllables, RMS amplitude gain was selected with an amplitude compression degree of 10 as recommended by the software. The RMS amplitude was used as it maintains an overall average of the non-silent portions of the signal which would not vary with additions of gaps due to variables like noise. This option was combined with amplitude compression to make sure that the increase in intelligibility is due to enhancement and not due to a general increase in signal to noise. Figure 3.3 gives a bird's eye view of the sound /s/ for unprocessed, enhanced and companded signal condition, in the context of vowel /a/.

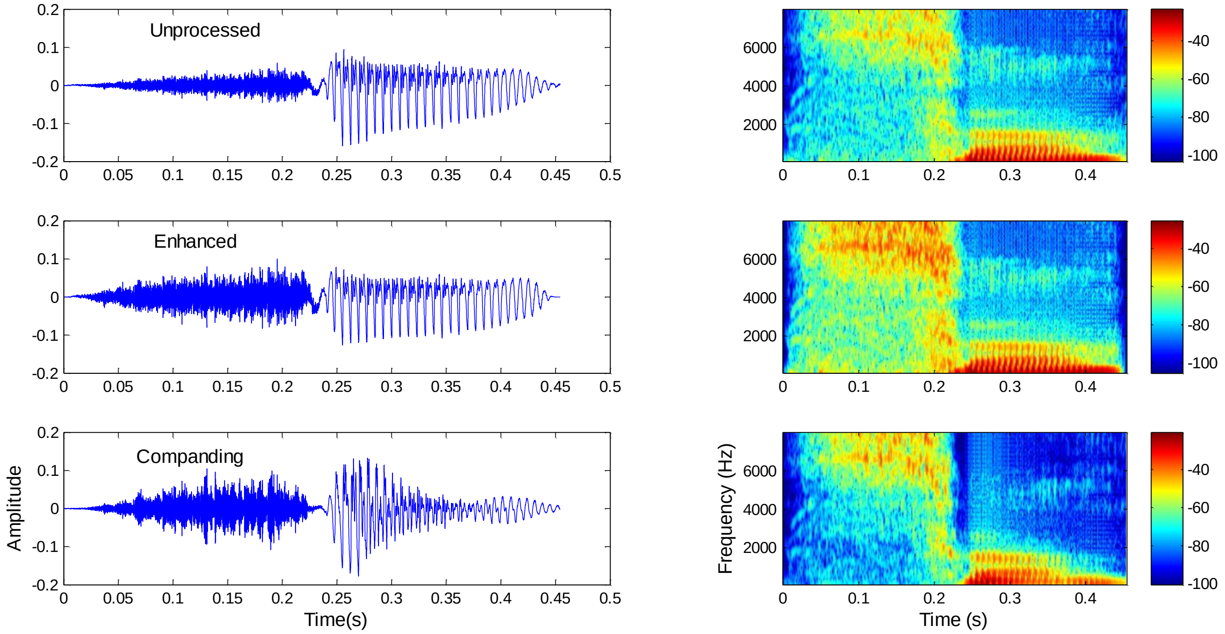


Figure 3.3: Figure showing the spectrum and spectrogram of the syllable /sa/ for all the three conditions without noise.

The signal was spectro-temporally enhanced using companding and consonant enhancement at various SNRs. Following this, they were normalized along with the unprocessed stimuli in order to avoid any intensity differences amongst them serving as an unrequired variable. This was done using RMS amplitude normalisation at -15dB in Adobe Audition software v5.

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). The following procedure was used for subject selection.

- A detailed case history was taken for all the participants before the routine audiological assessment was carried out in order to ensure that they do not report of symptoms that would exclude them from the study, based on the subject selection criteria as described before.
- The modified version of Hughson and Westlake procedure (Carhart & Jerger, 1959) was used to obtain air conduction and bone conduction thresholds at octave frequencies between 250 Hz to 8000 Hz and 250 Hz to 4000 Hz respectively.
- As a part of speech audiometry, speech recognition scores were obtained with spondee word list that was given by Vandana (1998). Speech perception in noise (SPIN) scores were obtained for PB word list that was given by Yathiraj & Vijayalakshmi (2005). Uncomfortable level (UCL) for running speech was also obtained.
- Immittance audiometry was carried out by varying the pressure from +200 dapa to -400 dapa for a probe tone frequency of 226 Hz. Ipsilateral and contralateral stapedial acoustic reflexes thresholds were measured for 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz pure tones.
- Otoacoustic emissions were obtained for 260 nonlinear click stimuli. SNR of more than 6 dB SPL in at least 3 consecutive octave frequencies in both ears with reproducibility greater than 50%, was considered as presence of OAEs (Wagner, Heppelmann, Vonthein & Zenner, 2008).
- Auditory Brainstem Response (ABR) was recorded using standard ABR protocol (Hall, 2006) with 11.1/ sec repetition rate and 100 μs click stimulus at 90 dBnHL

in all the participants. The filter settings that was kept had a high pass cut off of 100 Hz and low pass cut off of 3000 Hz. At least two recordings with stimulus in rarefaction polarity was done for each ear to ensure reproducibility.

3.7 Procedure used to obtain data

In order to obtain responses for a consonant identification task, the response screen on the laptop was shown to the participant before beginning the trial phase to familiarize him/ her with the possible 19 syllable options. Hence, a closed set task was administered.

All the 19 CV syllables (unprocessed and processed at different SNRs) were fed into MATLAB 7.8 software and programmed. They were programmed in such a way that all 19 CV syllables were presented in a randomized sequence for a particular SNR in a particular signal condition. This procedure was done for presentation of syllables at all SNRs in each of the signal conditions. Each CV was presented thrice with an inter-stimulus interval of 50ms. Following this, the participant was asked to respond for that token by verbally repeating the CV syllable that was perceived. The tester then clicked on the same syllable token on the laptop screen. As the stimulus was being presented to the participant through headphones, the tester was not aware of the stimulus that was being presented. Hence, it was a blind folded procedure. Figure 3.4 is an example of the procedure in which stimulus was presented for two out of nineteen syllables.

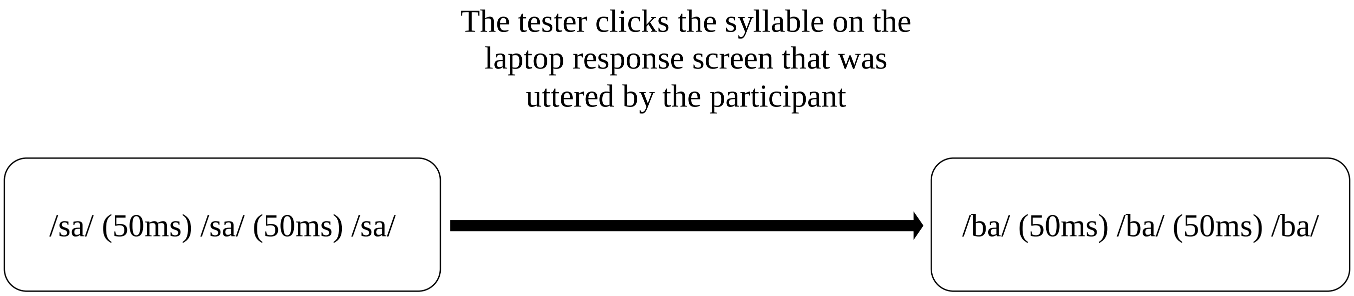


Figure 3.4: An example of the procedure used for presentation of two out of nineteen stimuli

3.8 Instructions

The participants were instructed to listen and repeat each CV syllable perceived by them. They were informed that the stimulus could be embedded in varying levels of noise and could occur in quiet condition as well. In situations where the participant was not able to correctly recognize the syllable, he or she was asked to guess the syllable that could have most probably occurred. Therefore it was a forced choice response.

3.9 Trial phase

This stage started with establishing their most comfortable level by using a file of stimulus token that was selected randomly. The file selected did not matter as all the stimuli were normalized prior to this. Following this, a trial run for unprocessed quiet and unprocessed SNR 10 conditions was given to familiarize the participant and to stabilize the consistency of their responses.

3.10 Testing phase:

The testing phase was carried out similarly for both the groups. The test material consisted of 15 files corresponding to five SNR conditions under each of the three stimulus conditions. These files were presented in a semi random manner by initializing this phase with either quiet or SNR15 condition of any stimulus condition (unprocessed or processed). This was followed by a mixture of the remaining conditions. This ensured that the order effect was avoided. The sequence of syllables within any SNR file was automatically randomized by the software to avoid any practice effect. A two minute interval between each SNR presentation in both processed and unprocessed condition was given. Also, the testing was done in two sittings, consisting of 7 and 8 presentations respectively. This was done to avoid fatigue.

The stimuli were presented through calibrated headphones attached to a calibrated output system of a laptop interface via MATLAB 7.8 software. The response screen consisted of the 19 syllables in English script, as choices. On hearing the stimulus, the participant was asked to identify and repeat the stimulus perceived. The experimenter then selected the corresponding option on the screen by clicking on the same. Hence, the inter stimulus interval varied depending on the participant’s response. The next syllable was present as soon as the experimenter clicked on an option on the screen. The response screen was as shown below,

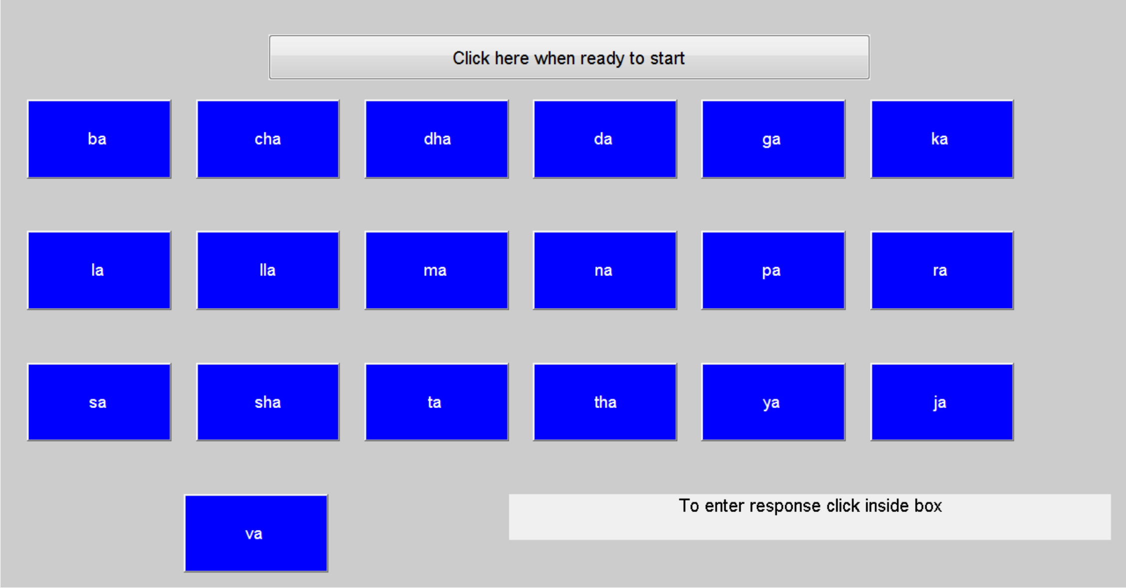


Figure 3.5: The response screen that was showed to the participants

After each trial, a response matrix was saved by the software indicating the response specific to each syllable as stimulus. The sum of correct score for each stimulus condition was obtained separately at each signal to noise ratio. These scores were later tabulated in order to be compared and analysed according to the groups, subgroups, conditions and SNRs.

Chapter 4

Results

The current study was taken up with an aim of investigating the effect of spectro-temporal enhancement of speech stimuli on speech perception, by comparing three signal conditions, namely, unprocessed, companded and consonant enhanced stimuli in individuals with cochlear pathology and normal hearing. This was done in a total of five SNR conditions- quiet, 15 dB SNR, 10 dB SNR, 5 dB SNR and 0 dB SNR for each of the signal conditions used. Nineteen consonants in the context of vowel /a/ were used to see the effects of signal conditions. The data obtained was from 16 individuals with cochlear hearing loss and 14 normal hearing individuals. The clinical group was again subdivided into subjects with flat (N=11) and sloping hearing loss (N=5) based on their audiograms. Consonant identification scores were tabulated and analysed using Statistical Package for Social Sciences (SPSS, version 16.0). The following is a summary of the statistical analysis that was performed to investigate the objectives of the present study.

4.1 Descriptive analysis was done to obtain the mean, median and standard deviation values for each of the following groups:

- Individuals with normal hearing
- Individuals with cochlear hearing loss
    - ➢ Cochlear hearing loss with flat configuration
    - ➢ Cochlear hearing loss with sloping configuration

Following the descriptive analysis, it was necessary to check if the data followed normal distribution or not in order to select the appropriate statistical analyses to be used.

For this reason, the data was subjected to Shapiro-Wilk's test. The results showed that, all data did not follow normal distribution. Also, the sample size was less and descriptive analysis revealed that the standard deviation was large in a few conditions in the clinical group. Hence, the data was analysed using non parametric tests. As the mean values would not be a true representative of the data in this case, median values were considered. The following is a summary of the non-parametric statistical analyses that was carried out.

4.2 Comparison of consonant identification scores between groups

To compare and see if there existed a significant difference between the performances of the two groups, individuals with normal hearing and those with CHL, Mann-Whitney U test was performed. Similarly, in order to compare the results of the two subgroups of cochlear hearing loss i.e. cochlear hearing loss with flat configuration and cochlear hearing loss with sloping configuration, Mann-Whitney U test was administered.

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

The consonant identification obtained at five SNR conditions (0dB, 5dB, 10dB, 15dB and quiet) were compared for each of the three signal conditions (unprocessed, enhanced and companded) separately. This was achieved by running Friedman's test over this data. If and when the comparison showed a significant difference, Wilcoxon signed rank test was administered to see which pair amongst them showed a significant difference. It was carried out for each of the following groups,

- Individuals with normal hearing
- Individuals with cochlear hearing loss

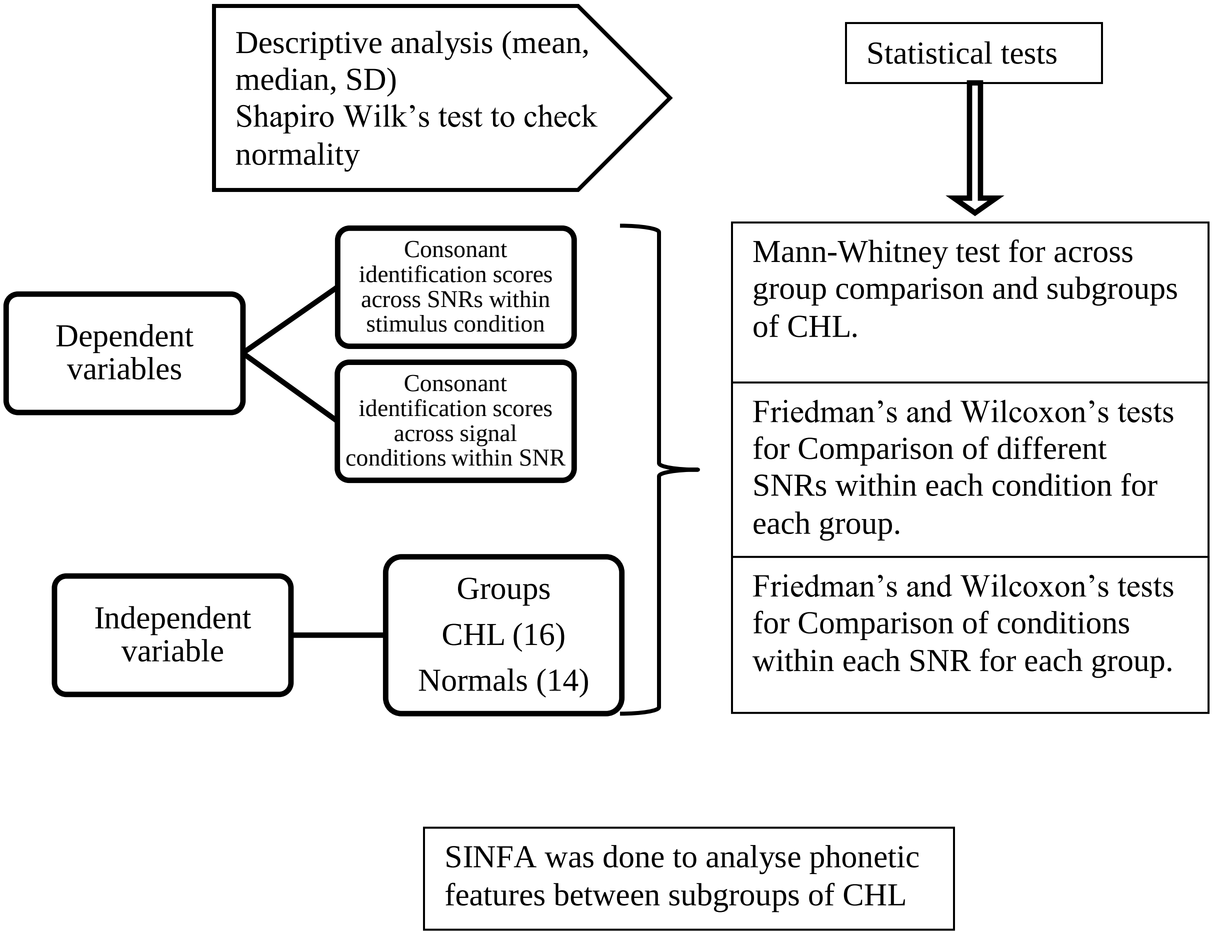
4.4 Comparison of consonant identification scores obtained for different signal conditions within each SNR and within each group

The scores obtained at the three signal conditions were compared at each of the five SNRs using Friedman’s test. Further, Wilcoxon signed rank test was carried out for those comparisons that showed a significant difference to see which pair had a significant difference and which did not. The following groups underwent this analysis.

- Individuals with normal hearing
- Individuals with cochlear hearing loss

4.5 Sequential Information Transmission Analysis (SINFA)

This was carried out for the added response matrices of the subgroups of CHL to determine the amount of information transmitted for phonetic features like, voicing, place of articulation (POA) and manner of articulation (MOA), independently and compare the same. The following is an illustration of the analysis that was carried out.



4.1 Descriptive analysis

The number of correct syllables identified was obtained for every individual out of a maximum of nineteen syllables that were presented. These consonant identification scores were considered for the calculation of mean, median and standard deviation. This was done across signal conditions and SNRs for all the groups considered. The same has been represented in Table 4.1. The median values have been represented in Figure 4.1.

Table 4.1: Mean, standard deviation (SD) and median values of consonant identification scores across different SNRs and different stimulus conditions for control and the clinical groups

Stimulus condition		Unprocessed (across SNRs in dB)					Consonant Enhancement (across SNRs)					Companding (across SNRs)				
Population		0	5	10	15	Quiet	0	5	10	15	Quiet	0	5	10	15	Quiet
Normal	Mean	11.86	16.5	17.79	17.64	18	11.21	15.43	17.07	18.43	17.86	11.93	16.07	17.57	17.71	17.71
	SD	4.27	2.24	1.67	1.21	2.07	4.45	2.31	2.16	0.85	2.24	3.4	1.77	1.08	0.91	1.97
	Median	13	17	18.5	18	19	12	16	17.5	19	19	13	16.5	18	18	18.5
CHL	Mean	7.06	11.37	14.06	15.06	15.38	7.5	10.81	13.75	14.94	15.06	5.62	10.5	12.88	15.06	15.88
	SD	4.15	2.98	3.53	3.66	3.79	3.67	4.24	3.21	2.99	3.27	3.2	4.57	4.04	3.35	2.65
	Median	6.5	12	14	16	17	7	9.5	13	14.5	16	5.5	11.5	14.5	16	16.5

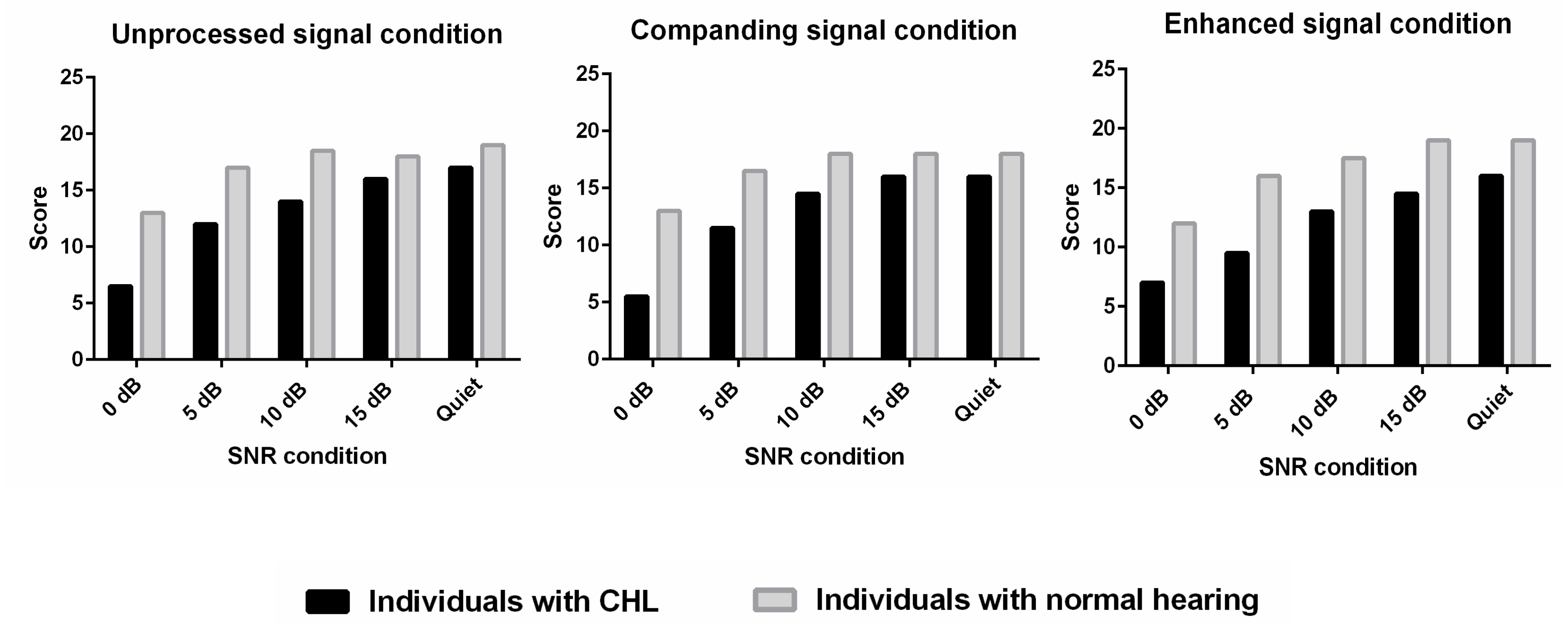


Figure 4.1 Median values for consonant identification scores across SNRs in each signal condition for clinical and control group

From the above table and figure, it can be inferred that the scores follow a trend with respect to SNR, irrespective of the condition or group. That is, as the amount of background noise increased, the consonant identification scores decreased. The only exception is the mean value of 15 dB SNR being more than quiet condition in consonant enhanced condition in normal hearing listeners, although the median values in both conditions are the same.

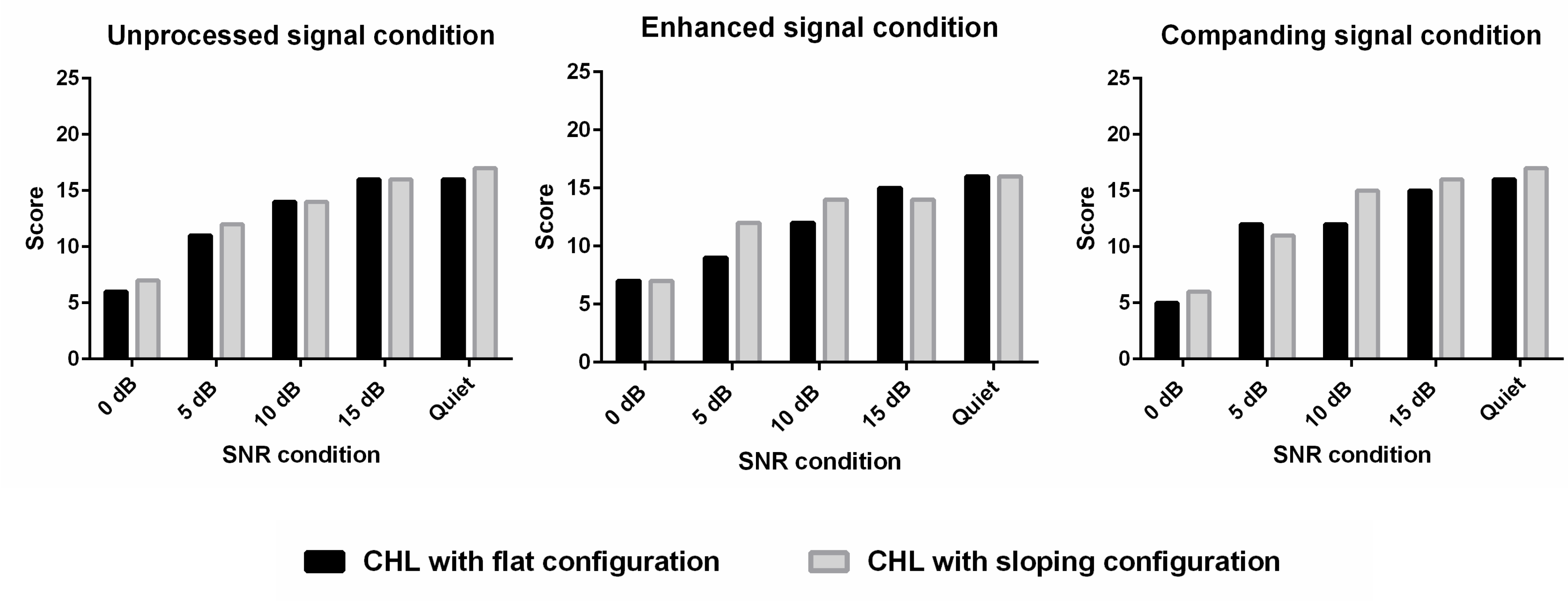
Although both normal hearing individuals and those with cochlear hearing loss had maximum scores in quiet and minimum at 0 dB SNR condition, the corresponding values were higher in that of the normal hearing group. On an average, the clinical group had a range of 33% to 87% from 0 dB SNR to quiet condition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

The standard deviation of both groups suggests that the performance was more variable with lesser SNR than when SNR was high. The individual variation existed more in the clinical group than in the control group.

Further, the clinical groups were divided in to two sub groups, namely, cochlear hearing loss with flat, and with sloping configuration. The mean, median and SD values have been displayed for these two subgroups as well in Table 4.2 and Figure 4.2.

Table 4.2: Mean, standard deviation (SD) and median values of consonant identification scores across different SNRs, at each stimulus conditions for the two subgroups of the clinical group

Stimulus condition		Unprocessed (across SNRs in dB)					Signal Enhancement (across SNRs)					Companding (across SNRs)				
Population		0	5	10	15	Quiet	0	5	10	15	Quiet	0	5	10	15	Quiet
Flat	Mean	6.91	11.09	14	14.55	14.82	7.64	10.91	13.64	14.82	15.09	5.27	10.18	12.45	14.55	15.64
	SD	4.52	3.53	4.14	4.29	4.33	4.2	4.46	3.58	3.4	3.83	3.28	5.09	4.43	3.93	2.94
	Median	6	11	14	16	16	7	9	12	15	16	5	12	12	15	16
Sloping	Mean	7.4	12	14.2	16.2	16.6	7.2	10.6	14	15.2	15	6.4	11.2	13.8	16.2	16.4
	SD	3.64	1.22	1.92	1.3	2.07	2.49	4.21	2.55	2.16	1.87	3.2	3.56	3.27	1.09	2.07
	Median	7	12	14	16	17	7	12	14	14	16	6	11	15	16	17


Unprocessed signal condition

Enhanced signal condition

Companding signal condition

Figure 4.2 Median values for speech perception scores across SNRs in each signal condition for the two sub groups of CHL

Similar to the data of normal hearing individuals and cochlear hearing loss group as a whole, the trend in the sub groups of cochlear hearing loss was also similar, with maximum scores obtained in quiet or SNR 15 condition. Deterioration in perception was noticed as the SNR decreased. It is evident that, with addition of noise, the consonant identification scores in both the subgroups decreased, making evident, the detrimental effect of noise on speech perception. This is well represented in Table 4.2 and Figure 4.2.

On an average, the group with flat configuration had a range of 32% to 84% from 0 dB SNR to quiet condition, while the sloping configuration group had a range of 35% to 88% indicating similar performances. Upon visual inspection it can be noted that, the trend in standard deviation of both the sub groups is similar to the previous comparison made between clinical and control group. The performance became more variable as the SNR decreased. The individual variations however, were slightly more in the group with flat configuration compared to the group with sloping configuration.

The difference in improvement between unprocessed and processed was calculated by subtracting the processed from unprocessed scores. Hence, a negative value would indicate the direction of improvement as processed signal being better. The mean, median and standard deviation of the same are shown in Table 4.3. The highlighted region shows a mean improvement in the processed condition.

Table 4.3: Mean, standard deviation (SD) and median values of the differences in consonant identification scores between processed and unprocessed signal conditions, across different SNRs obtained for both the groups. The shaded area shows a mean improvement in the processed signal

Stimulus condition		Unprocessed- Signal enhancement (across SNRs in dB)					Unprocessed- Companding (across SNRs)				
Population		0	5	10	15	Quiet	0	5	10	15	Quiet
Normal	Mean	-0.21	0.57	0.14	0.00	0.21	0.43	1.07	0.64	-0.71	0.07
	SD	1.57	1.82	1.09	0.87	0.69	3.22	1.54	2.40	0.72	0.47
	Median	0.00	0.00	0.50	0.00	0.00	0.00	1.00	0.00	-1.00	0.00
CHL	Mean	-0.44	0.56	0.31	0.13	0.31	1.44	0.87	1.19	0.00	-0.50
	SD	2.89	2.75	1.62	1.85	1.88	2.96	3.03	2.34	1.09	1.96
	Median	-0.50	0.50	1.00	0.00	0.00	1.50	1.00	2.00	0.00	0.00

From the above table, it can be noted that, there was a mean improvement in scores in the processed condition in only signal enhanced condition at 0 dB SNR and companding at 15 dB SNR for normal hearing individuals. In individuals with CHL the improvement noticed was at 0 dB SNR for signal enhancement and quiet for companding condition. The rest of the scores suggest that there was a higher or equal improvement in unprocessed condition.

4.2 Comparison of consonant identification scores between groups

In order to compare the consonant identification data of clinical and control groups and to check for any significant differences, Mann-Whitney U test was administered. This was done at each SNR and each stimulus condition. The results indicated a significant difference ($p<0.05$) in the consonant identification scores in all five SNRs in each of the three conditions between the two groups. The results of this test are given in Table 4.4 for all the conditions at different signal to noise ratios.

Table 4.4: The |Z| values and significance level (p value) obtained from Mann-Whitney U test for consonant identification task between CHL and normal group at all SNR and different stimulus conditions

Signal Condition	SNR (dB)									
	0		5		10		15		Quiet	
	|Z|	P	|Z|	P	|Z|	P	|Z|	p	|Z|	P
Unprocessed	2.86	0.004	3.87	0.000	3.15	0.002	2.22	0.026	2.68	0.007
Enhanced	2.13	0.033	2.95	0.003	3.01	0.000	3.73	0.004	2.92	0.000
Companding	3.82	0.000	3.77	0.000	3.76	0.000	2.90	0.004	2.46	0.014

Note: p<0.001, 2-tailed

Mann-Whitney U test was also done to see if the consonant identification performance of the two subgroups of cochlear hearing loss differed significantly at each SNR and stimulus condition. Results revealed an insignificant difference across all the SNRs in each of the signal conditions. Therefore, further statistics for the sub-groups were not carried out separately. The data was combined and was referred to as the clinical group for further statistical analyses.

As Mann-Whitney U test revealed a significant difference in consonant identification scores across individuals with normal hearing and with CHL, it was necessary to know if the groups showed a significant difference in consonant identification task across SNRs or across conditions. Hence, Friedman test was carried out. The following sections consist of results of the same.

In order to see if the difference caused by processing the signal made a significant difference in CHL when compared to individuals with normal hearing, Mann-Whitney test was then carried out on the differences obtained between processed and unprocessed scores (unprocessed-processed). The results indicated no significant difference in the improvement in individuals with normal hearing as compared to individuals with CHL across SNRs and conditions.

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

All the participants were given a task to identify consonants in three signal conditions, and five SNRs. To compare the resultant scores across SNR in each stimulus condition, Friedman test was administered for clinical and control groups separately. They are given below,

4.3.1 Individuals with normal hearing

To check if there was a significant difference across the consonant identification scores at different SNRs within each of unprocessed, consonant enhanced and companded signal conditions, Friedman test was administered. The test showed a significant difference (p<0.01) across SNRs within all the signal conditions. The results of the same are displayed in Table 4.5.

Table 4.5: Results of Friedman test with χ^2(df) and significance levels across all SNRs at each stimulus condition for individuals with normal hearing

Signal condition	χ^2 **(4)**	***p* value**
Unprocessed	29.21	0.00
Consonant enhanced	39.39	0.00
Companding	36.67	0.00

Note: p<0.001, 2-tailed

As Friedman test showed an overall significant difference in speech scores across SNRs within each condition, Wilcoxon signed rank test was chosen to further evaluate which of the ten SNR pairs had a significant difference in the consonant identification scores. The details of Wilcoxon signed rank test are shown in Table 4.6.

Table 4.6: Results of Wilcoxon signed ranked test showing significant differences for SNR pairs with each signal condition obtained in normal hearing individuals

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

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

In the above table the SNR pairs that did not have any significant difference have been shaded. It is clear that these are the pairs with higher values of SNR. As the SNR value increased, the difference amongst its adjacent SNRs decreased.

In the unprocessed condition, a significant difference ($p<0.05$) was evident in all SNR pairs except for 5 vs quiet, 10 vs 15, 10 vs quiet and 15 vs quiet, which did not show a significant difference. In the consonant enhanced condition, all the SNR pairs differed significantly ($p<0.05$), except, quiet vs 10 dB SNR and quiet vs 15 dB

SNR. In companding signal condition, the test indicated significant difference across all SNR pairs except 10 vs 15, 10 vs quiet, and 15 vs quiet conditions. In all the shaded SNR pairs, the higher SNR had better scores compared to the lower SNR in the pair.

To sum up, out of the three conditions, unprocessed signal had the least amount of significant differences amongst higher SNR pairs, followed by companding and then enhanced condition, which showed significant differences in all SNR pairs except, 10 vs. quiet and 15 vs. quiet conditions.

4.3.2 Individuals with cochlear hearing loss

Friedman test was administered separately for each signal condition to check if there existed a significant difference in the consonant identification scores across the five SNR conditions within each signal condition, separately. The test showed significant differences (p<0.01) for all three signal conditions which is represented in Table 4.7.

Table 4.7: Results of Friedman test with χ^2(df) and significance values across SNRs within each signal condition for individuals with CHL

Signal condition	$\chi^2_{(4)}$	***p* value**
Unprocessed	51.09	0.00
Consonant enhanced	55.29	0.00
Companding	53.54	0.00

As Friedman test showed a significant difference in scores across signal conditions, Wilcoxon's signed rank test was further administered to check which of the SNR pairs showed significant differences in consonant identification scores. In all the three signal conditions, similar results were obtained for Wilcoxon signed rank test. A significant difference ($p<0.05$) was found for all SNR conditions except quiet vs. 15 dB SNR. The results of the same are displayed in Table 4.8.

Table 4.8: Results of Wilcoxon signed ranked test showing significant differences for SNR pairs within each signal condition for individuals with CHL

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

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

Each of the three signal conditions- unprocessed, enhanced and companding were compared within each of the five SNR conditions- 0, 5, 10, 15dB SNR and quiet. This was done using Friedman test in each group separately. The results are explained below.

4.4.1 Individuals with normal hearing

Friedman test for comparison of consonant identification scores across stimulus conditions at various SNRs revealed that a significant difference existed across signal conditions for only 15dB SNR. There was no significant difference seen across signal conditions for any other SNR. The same has been shown in Table 4.9.

Table 4.9: Results of Friedman test with χ^2(df) and significance level across signal conditions at all SNRs for individuals with normal hearing

SNR (dB)	χ^2**(2)**	***p*-value**
Quiet	3.80	0.150
15	11.619	0.003
10	2.47	0.290
5	3.17	0.205
0	1.73	0.420

As Friedman test revealed significant differences across signal conditions at only 15 dB SNR, Wilcoxon signed rank test was administered for scores obtained at 15 dB SNR to see which two signal conditions at this SNR showed a significant difference ($p<0.05$). It was seen that the speech scores differed significantly ($p<0.05$) across unprocessed- enhanced and companded- enhanced signal condition pairs with consonant enhanced condition resulting in significantly higher scores in both pairs. The results have been displayed in Table 4.10.

Table 4.10: Results of Wilcoxon signed ranked test for stimulus condition pairs in individuals with normal hearing at 15 dB SNR

15 dB SNR	Unprocessed	Companding
Unprocessed		
Companding	Not significant (p<0.05)	
Enhancement	Significant (p<0.05)	Significant (p<0.05)

4.4.2 Individuals with cochlear hearing loss

While comparing the consonant identification scores at each SNR across conditions, Friedman test showed a significant difference (p<0.05) only at 0 dB SNR. The same has been shown in Table 4.11.

Table 4.11: Results of Friedman test with χ^2(df) and significance level across signal conditions at all SNRs for individuals with CHL

SNR (dB)	χ^2**(2)**	***p* value**
Quiet	3.800	0.150
15	0.122	0.941
10	4.66	0.097
5	2.33	0.311
0	6.87	0.032

As Friedman test showed a significant difference across signal conditions at 0 dB SNR, Wilcoxon signed rank test was administered for the same. The test indicated a significant difference (p<0.05) only between companded and enhanced signal conditions with consonant enhanced condition having significantly higher scores than companding. Table 4.12 represents the same.

Table 4.12: Pairwise comparison of SNRs across conditions in 0 dB SNR for individuals with CHL

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

4.5 Sequential Information Transfer Analysis (SINFA)

Sequential information transfer analysis (SINFA) (Wang & Bilger, 1973) is a method of determining the amount of information transmitted for each of the defined phonetic features, independently. This analysis was carried out to assess the same in subgroups of CHL- CHL with sloping configuration and CHL with flat configuration. Following is the gist of the procedure that was carried out,

- Patient responses for each of the condition and SNR were obtained in the form of a confusion matrix.
- The individual matrices were added using MATLAB 7.8 for each SNR across conditions.
- The matrices were subjected to SINFA using the software Feature Information Xfer (FIX) (developed by University College of London, Department of Linguistics).

4.5.1 Stimulus response matrices

For each condition and SNR, the responses from the participants were obtained in the form of a stimulus-response matrix. Each participant had fifteen matrices, i.e. 5 SNRs × 3 conditions. The first row on the top of the matrix indicated the responses while the first column from the left indicated the stimulus presented. The number in each cell represented the frequency of the particular stimulus-response pair. The number in the cells along the principle diagonal axis was the correct response. The individual responses for five SNRs in each stimulus condition were added across participants. For example, stimulus-responses matrices of five subjects with sloping configuration were added for 0 dB SNR in unprocessed condition. Similarly, the stimulus-response matrices of subjects were added for all SNRs in each of the stimulus conditions. This was done for individuals with flat and sloping loss separately. The following is an example of an added matrix (Table 4.13).

Table 4.13: Example of a stimulus response matrix showing the results obtained for consonant enhanced condition at 0 dB SNR for 5 CHL participants with sloping configuration. The correct response have been highlighted in the diagonal axis.

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

4.5.2 Results of SINFA

The confusion matrices were subjected to SNIFA for assessing information transmitted for place, manner and voicing across conditions and SNR. The feature matrix was constructed with the phonetic features of each of the nineteen syllable considered. The same is represented in table 4.14.

Table 4.14: Feature matrix of the 19 syllables considered

	b	ḍ	g	ʤ	k	l̲	l	m	N	p	R	s	t̲	v	j	tʃ	d̲	ʃ	t̲
Voicing	+	+	+	+	-	+	+	+	+	-	+	-	-	+	+	-	+	-	-

Place	b	a	v	p	v	p	a	b	A	b	A	a	a	l	p	P	d	p	d
Manner	p	p	p	a	p	l	l	n	N	p	L	f	p	g	g	A	p	f	p

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

The information transmitted is calculated in bits for each of the features- voicing, POA and MOA along with total information transmitted. The total information transmitted in this experiment, ranged from 0 - 4.24. However, for each of the individual components like voicing, POA and MOA, the information transmitted ranged from a minimum of 0 and maximum of 1. The following Figure 4.3 represents the information transmitted for all the parameters in the two subgroups considered.

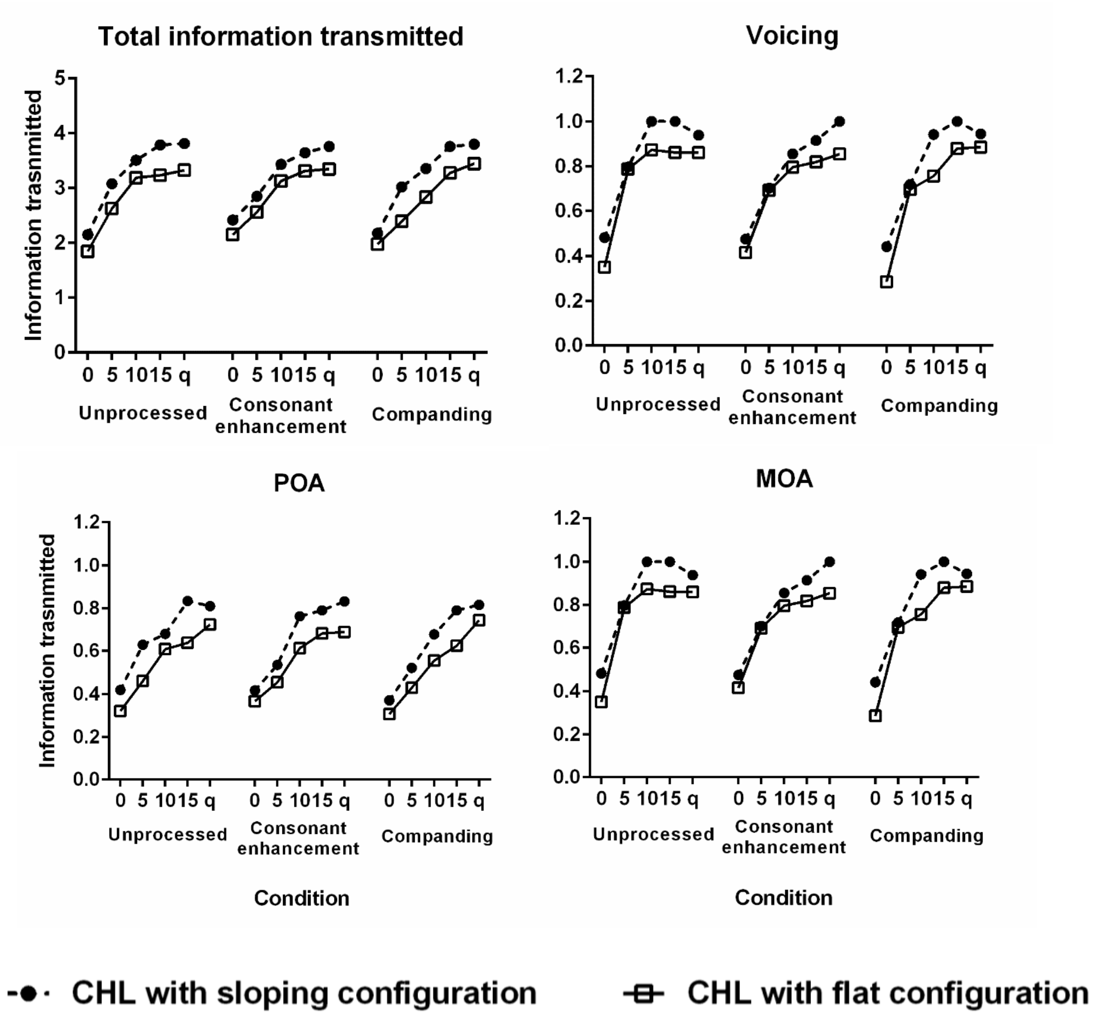


Figure 4.3: Information transmitted in bits for voicing, MOA, POA and total information transmitted across SNRs and signal conditions for the two subgroups of CHL

It can be inferred from Figure 4.3 that, the maximum information transmitted was for MOA cue followed by voicing and POA which were equally transmitted. The information transmitted across voicing and POA was similar within both individuals with flat and sloping configuration. Individuals with flat hearing loss, however, showed a higher information transmission for MOA cues with consonant enhancement condition only at 0 dB SNR. At the rest of the SNRs, the information transmitted for MOA cues was similar within the subgroups of CHL.

The results of the study can be summarised as follows,

1. A trend of reduction in speech scores with increase in noise in both groups was observed. Individuals with CHL had lesser scores when compared to normal listeners in all conditions. The detrimental effect of noise was greater in the clinical group.
2. There was a significant difference in consonant identification scores between clinical and control group but not between the subgroups of CHL.
3. No significant differences between the improvements by processed signal across clinical and control group were obtained.
4. There was a significant difference seen across lower SNRs within all conditions in normal hearing individuals. It was present in all SNRs except 15 vs. quiet in individuals with CHL.
5. Significant effect across conditions within SNR with better consonant identification scores in consonant enhanced condition was seen at 15 dB SNR and 0 dB SNR in normal and CHL listeners respectively.

6. Results of SINFA indicated that the best transmitted parameter in both the subgroups of CHL (CHL with flat and sloping configuration) was MOA followed by voicing and POA which were equally transmitted.
7. There was no benefit seen from the processed signal in the information transmitted for voicing and POA. Individuals with flat hearing loss showed benefit for MOA cues with consonant enhancement condition at 0 dB SNR.

Chapter 5

Discussion

The present study was taken up to evaluate the effect of spectro-temporal signal enhancement using companding and consonant enhancement on speech perception and comparing it with unprocessed signal condition in population with normal hearing and with CHL across five SNRs. The consonant identification scores were obtained for a set of 19 consonants presented in the context of vowel /a/. The obtained data were tabulated and analysed using Statistical Package for Social Sciences (version 16.0). The results of the study are discussed below.

5.1 Effect of noise on consonant identification scores

Results of descriptive analyses of the clinical, control group and subgroups of cochlear hearing loss showed a trend for consonant identification scores with respect to SNR, irrespective of the condition or group. As the level of background noise increased, the consonant identification scores decreased.

This detrimental effect of addition of noise on speech perception has been well established in literature (Nabelek et al., 1989; Dorman, Loizou & Tu, 1998). It is believed that the addition of a background noise reduces the distance between the peaks and troughs, thereby reducing the available spectral cues in order to identify speech. Hence, speech scores are poorer in the presence of noise (Baer & moore, 1993).

The only exception to the trend of speech scores decreasing with decrease in SNR was the mean values at 15 dB SNR and quiet for consonant enhanced condition. Here, the mean at 15 dB SNR was more than the mean at quiet condition. The reason for this slight difference could have been due to chance factor which can be justified with a slightly increased standard deviation at quiet condition. It could also be observed due to the consideration of a small subject sample which could have resulted in more variability. Hence, median is a better representation of this data, which is equal at 15 dB SNR and quiet conditions for enhanced stimulus in normal hearing listeners.

In spite of both the groups following the trend, the corresponding scores at each SNR were lesser for clinical group than the control group. This result could be attributed to the classical features of cochlear pathology like, reduced audibility, reduced frequency selectivity and temporal resolution which has been well supported in literature as well (Eisenberg et al., 1995; Pekkarinen et al., 1990).

However, reduced audibility could not have been the primary reason for this decrease in speech scores as poor thresholds were compensated for by presenting the signal at the participant's most comfortable level (MCL). The dependence on loss of audibility as the main reason for decreased perception has also been discarded in literature (Dubno & Dirks, 1982; Dubno & Schaefer, 1992). However, loss of frequency selectivity, loudness recruitment and temporal resolution could be the contributing factors. These parameters have been well correlated with degree of hearing loss and speech scores as well (Moore & Glasberg, 1993; Baer & Moore, 1993; Nejime & Moore, 1997). Their effects are more pronounced in noisy situations. When the background noise reduced the spectral contrasts, individuals with CHL failed to resolve the remnant frequency cues into its frequency components due to widened auditory filters. This resulted in a highly smoothened representation of the input signal, which, the clinical group was unable to decode (Moore, Glasberg & Simpson, 1992).

Another reason could be the dependence of individuals with CHL on envelope cues rather than fine structure cues (Lorenzi, Gilbert, Carn, Garnier & Moore, 2006). An addition of noise would reduce the modulation depth of speech signal and further distort the envelope. Therefore, due to the interaction of reduced frequency selectivity, temporal resolution and loudness recruitment, as mentioned above, their speech perception scores in both quiet and noise was less than that of their normal hearing counterparts although the loss of audibility was compensated for.

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

6.2 Comparison of consonant identification scores between groups and subgroups considered

There was a significant difference in the consonant identification scores between the clinical and control group, but not between the subgroups of CHL. As discussed earlier, due to problems like reduced audibility, reduced frequency selectivity, temporal resolution and loudness recruitment, the performance of group with CHL became significantly poorer than that of normal listeners in both quiet and noise conditions (Eisenberg et al., 1995; Pekkarinen et al., 1990). Similarly, as discussed above, due to the consideration of participants with only up to gradually sloping hearing loss, the two subgroups did not differ significantly.

5.3 Comparison of improvement obtained in consonant identification with processed stimulus across groups for all SNRs and stimulus conditions

There was no significant difference in the improvement seen between clinical and control group for the benefit obtained by processed speech stimuli (unprocessed-processed condition). This means that processing of signal, although gave benefit within a group across some SNRs and conditions, did not show improvement in consonant identification in normal hearing individuals and individuals with CHL. The improvement obtained in both groups for processed signal, is similar. The results of benefit with consonant enhancement within group comparison have been equivocal in the past. Few researchers like Bunnell (1990) and Hazan et al. (1998) found improvement. However, various other authors like Baer et al. (1993), Stone and Moore (1992), did not obtain a significant different in intelligibility of speech. Bunnell (1990) found moderate level of improvement for /b/ and /g/ and inconsistent results for /d/. The reason attributed for the improvement as opposed to others who did not find an improvement was that Bunnell used consonant enhancement technique that enhanced only the mid frequencies to avoid masking of higher formants by F1 that otherwise happened in previous studies. Also, Bunnell’s study was conducted in only quiet condition. Hence, the results of the current study need not be in accordance with Bunnell’s.

Hazan et al. (1998) found a significant difference in the processed stimulus (VCVs) at 0 dB SNR for normal hearing listeners while examining speaker and listener effects with population of different languages. In their study, noise with same long term average speech spectrum as the stimulus was used, which was added to the stimulus after enhancement. This procedure is different from the current study with respect to processing the stimulus (CV, in this case) after the addition of noise (using 6 talker babble as background noise). By adding noise after enhancing the signal, the possible degradation of spectral and temporal contrast due to noise is not taken into consideration. Hence, the difference in results could be attributed to the procedural variations.

Baer et al. (1993) did not find an improvement in sentence perception with moderate enhancement although the improvement increased, though not to significant levels, with practice effect. The present study used stimulus that are very less redundant. Hence, if stimulus that is as redundant as sentences required familiarity with the strategy to show benefit, it is justified that CVs did not show a marked improvement in individuals with CHL. Hence, due to the above mentioned reasons, the improvement seen in CHL individuals themselves is small and seldom significant.

A study by Deepthi (2012) showed a significant difference in individuals with CHL using companding at 0, +10 and +15 dB SNR. The study used speech shaped noise as background noise. The current study used a six talker babble, which is known to be a better masker than speech shaped noise (Sperry, Wiley & Chial, 1997). Hence, the effects of companding, although present, might have been reduced to extent in the current study.

Further, using processed stimuli for normal hearing individuals could enhance the spectral and temporal features in a way that could make the stimulus sound unnatural. It is also known that, altering the signal in any way could also introduce possible distortions (Oxenham et al., 2007) that might be detected by a normal hearing individual with normal auditory physiology, although it might not affect individuals with CHL. Therefore, even though signal enhancement strategies can be significant within the group, the pros and cons of the strategies that act differently on each group can level the amount of improvement across groups and hence, might not be significant.

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

Under each of the three signal conditions- unprocessed, enhanced and companding, the consonant identification scores were compared across the five SNRs considered.

5.4.1 Individuals with normal hearing

A significant difference in consonant identification scores was noticed for lower SNR pairs but not for higher pairs. In unprocessed and companding condition, there existed significant differences for SNR pairs between 0 to 10 dB SNR. In consonant enhanced condition, a significant difference was seen for all pairs except the pair of quiet, 10 dB SNR and 15 dB SNR. That is, with a decrease in noise levels from 0 dB SNR to up to 10 dB SNR, there was a significant improvement in the consonant identification scores. A further decrease in noise levels did not increase their speech performance scores significantly in all three signal conditions.

This improvement seen with increasing SNR can be correlated with the study done by Beattie, Barr and Roup (1997). They noted an improvement in the monosyllabic word identification scores as the level of multitalker background decreased from 5 dB SNR to quiet condition in individuals with normal hearing and CHL. There are several studies that showed that normal hearing individuals were able to extract spectral and temporal cues better than hearing impaired population even in noisy situations (Beattie et al., 1997; Pekkarison et al., 1990; Heifer & Huntley, 1991). This is attributed to the normal physiology in these individuals which is capable of differentiating the speech signal from competing background noise. Several mechanisms like medial olivocochlear bundle (MOCB) mediated suppression, two tone suppression and other nonlinearities of normal cochlea could aid in this process (Kumar & Vanaja, 2004).

From the previous discussion, it is known that background noise has a tendency to reduce the spectral and temporal contrasts and hence deteriorate the speech performance. As the noise level decreases, the spectral and temporal cues available increase by a substantial amount which helps in improving the speech perception. Therefore, reducing the noise level up to 10 dB SNR from 0 dB SNR showed a significant improvement in consonant recognition scores in unprocessed stimuli. A further decrease in noise level did not provide a significant additional benefit.

In companding condition, the lost spectral and temporal cues are made available to the listeners through processing of speech stimulus (Turicchia & Sarpeshkar, 2005). Hence, at 10 dB itself, these individuals perform almost like in quiet conditions.

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

Therefore, due to the above reasons, normal hearing individuals could extract cues and perceive speech even in noise. Their consonant identification scores at 10 dB were similar to the scores in quiet situation.

5.4.2 Individuals with CHL

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

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

5.5.1 Individuals with normal hearing

There was a significant difference seen in consonant identification scores only at 15 dB SNR across enhanced- unprocessed and companding- enhanced signal conditions with consonant enhanced condition providing more benefit in both condition pairs.

The usefulness of consonant enhancement in individuals with CHL in quiet conditions has been supported in literature (Bunnell, 1990; Summerfield et al., 1985; Stone & Moore, 1992). The improvement in individuals with CHL has been attributed to enhancing the otherwise unavailable spectral cues. However, these studies have shown minimal or no benefit for individuals with normal hearing in consonant enhanced condition. The normally functioning auditory system is already capable of extracting spectral and temporal cues even in the presence of noise. An additional enhancement of these cues therefore doesn’t always significantly improve speech perception in these individuals.

Another explanation is that, it is highly unlikely that a signal enhancement strategy would provide a significant improvement at only one out five SNRs for normal hearing individuals. It is evident from the descriptive analysis that the scores of these individuals at 15 dB SNR are high in all conditions. Therefore, it can be concluded that, the occurrence of a significant difference between conditions at 15 dB SNR in normal hearing individuals could be attributed to less sample size and chance factor as well.

5.5.2 Individuals with CHL

There was a significant difference seen in consonant identification scores only at 0 dB SNR across companding- enhanced signal conditions with better scores obtained in consonant enhanced condition.

These results are an indication of higher benefit from consonant enhancement technique than companding. They are in agreement with previously existing literature which has shown a significant improvement in speech identification scores with envelope enhanced signal, in the presence of a competing signal (Apoux, Tribut, Debruille & Lorenzi, 2004; Baer et al., 1993; Bunnel, 1990; Clarkson & Bahgat, 1991; Franck et al., 1999; Lyzenga et al, 2002). This could be because of a large amount of deterioration of spectral cues at such high levels of noise. However, the amount of spectral cues available increases as the level of background noise decreases. These spectral cues available in unprocessed signal at higher SNRs might be similar to that available in processed signal. Hence, a further enhancement with processed signal does not significantly improve the consonant identification performance at higher SNRs.

The benefit from consonant enhancement technique being more than companding can be attributed to the spectral enhancement provided by consonant enhancement. On the other hand, companding improves both spectral and temporal aspects of the signal. It has been reported in the literature that widened auditory filters mainly cause deterioration of spectral cues to an extent dependent on the amount of cochlear damage (Baer & Moore, 1993). Hence, a strategy, like consonant enhancement, that would compensate for this by making the spectral peaks and contrasts more available would benefit these individuals (Bunnell, 1990; Summerfield et al., 1985; Stone & Moore, 1992). When a strategy like companding is used, since the processing of signal is more complex with a series of compression and expansion, the process could have altered the spectral and temporal cues more than required for these individuals. Hence, it could have added distortion to the signal for individuals with CHL. Due to the above reasons; there was a significant difference in scores at 0 dB SNR between companding- enhanced signal conditions with better scores in consonant enhanced condition.

5.6 Sequential Information Transfer Analysis (SINFA)

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

5.6.1 Place of articulation

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

This was because vowels are higher in energy as compared to consonants (Fletcher, 1953).

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

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

5.6.2 Voicing

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

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

5.6.3 Manner of articulation

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

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

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

To conclude, processed speech was found to improve the performances of both the groups, although the amount of improvement seen across the groups was similar. Individuals with CHL benefitted from processed speech at 0 dB SNR while normal hearing listeners showed a benefit at 15 dB SNR which also could be due to chance factor. Individuals with CHL seem to have benefitted more with consonant enhanced stimulus than companded stimulus. Results of SINFA indicated that the best transmitted parameter in both the subgroups of CHL (CHL with flat and sloping configuration) was MOA followed by voicing and POA which were equally transmitted. Also, there was no benefit seen from the processed signal in the information transmitted for voicing and POA. Individuals with flat hearing loss showed benefit for MOA cues with consonant enhancement condition at 0 dB SNR.

Chapter 6

Summary and Conclusion

Cochlear hearing loss is a type of hearing loss that is most commonly found. It results from damage to the inner ear or the cochlea. Consequences of these, are many kinds of perceptual consequences including, impaired frequency and temporal resolution (Thibodeau & Van Tasell, 1987; Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006). Although they are able to perform reasonably well in terms of speech perception in quiet, their performance drastically decreases with the addition of background competing signal. This has been a challenge that is yet to overcome in the field of rehabilitation for individuals with CHL.

Over the years, several signal processing strategies have been tried to improve the speech perception in noise. Two such techniques are companding and consonant enhancement. The former technique makes use of the concept of two tone suppression and alters both spectral and temporal in the speech signal while the latter aims to increase the spectral contrast by enhancing the spectral peaks. Although there have been studies demonstrating equivocal results with respect to the amount of improvement from the strategies independently, there has not been any study that has compared these two strategies in the same population to compare the benefit obtained. Also, there is a dearth of studies to establish the consistency in terms of improvement provided by these signal enhancement strategies in individuals with cochlear hearing loss. Hence, this study was taken up with careful selection of the type of stimulus and competing signal.

Two groups, control (N=14) group having normal hearing and clinical (N=16) group with cochlear hearing loss ranging from mild to moderately severe were considered. The testing was done in a total of five SNR conditions- quiet, 15 dB SNR, 10 dB SNR, 5 dB SNR and 0 dB SNR for each of the signal conditions used, namely, unprocessed, companded and consonant enhancement. The background noise used was six talker babble developed by Konadath, Vimal and Suresh (2014). Nineteen consonants in the context of vowel /a/ were used to see the effects of these signal conditions. The clinical group was again subdivided into subjects with flat (N=11) and sloping hearing loss (N=5) based on their audiograms. The signal was presented through MATLAB 7.8 and the responses were obtained in the form of stimulus matrices in order to further carry out SINFA. The following is a summary of the results obtained.

1. A trend of reduction in speech scores with increase in noise in both groups was observed. CHL had lesser scores when compared to normal in all conditions with greater effect of noise. Subgroups of CHL had similar performances.
2. There was a significant difference between clinical and control group but not between the subgroups of CHL. No significant differences between the improvements by processed signal across clinical and control group.
3. There was a significant difference seen across lower SNRs within all conditions in normal. It was present in all SNRs except 15 vs. quiet in individuals with CHL.
4. Significant effect across conditions within SNR with better consonant identification scores in consonant enhanced condition was seen in 15 dB SNR and 0 dB SNR in normal and CHL listeners respectively.

5. Results of SINFA indicated that the best transmitted parameter in both the subgroups of CHL (CHL with flat and sloping configuration) was MOA followed by voicing and POA which were equally transmitted.
6. There was no benefit seen from the processed signal in the information transmitted for voicing and POA. Individuals with flat hearing loss showed benefit for MOA cues with consonant enhancement condition at 0 dB SNR.

Conclusion

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

Clinical implications:

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

References

American National Standards Institute (1987), Specifications for instruments to measure aural acoustic impedance and admittance (aural acoustic immittance). *ANSI S3. 39-1987-R1996*. New York: American National Standards Institute.

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

American National Standards Institute (2004), specification for audiometers. *ANSI S3.6-1987- R1996*. New York: American National Standards Institute.

Apoux, F., Tribut, N., Debruille, X., & Lorenzi, C. (2004). Identification of envelope-expanded sentences in normal-hearing and hearing-impaired listeners. *Hearing research*, *189*(1), 13-24.

Arehart, K. H., Hansen, J. H., Gallant, S., & Kalstein, L. (2003). Evaluation of an auditory masked threshold noise suppression algorithm in normal-hearing and hearing-impaired listeners. *Speech Communication, 40*(4), 575-592.

Baer, T., & Moore, B. C. (1993). Effects of spectral smearing on the intelligibility of sentences in noise. *The Journal of the Acoustical Society of America*, *94*(3), 1229-1241.

Baer, T., Moore, B. C., & Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality, and response times. *Journal of rehabilitation research and development, 30*, 49-49.

Beattie, R. C., Barr, T., & Roup, C. (1997). Normal and hearing-impaired word recognition scores for monosyllabic words in quiet and noise. *British journal of audiology*, *31*(3), 153-164.

Bhattacharya, A., & Zeng, F. G. (2007). Companding to improve cochlear-implant speech recognition in speech-shaped noisea). *The Journal of the Acoustical Society of America*, *122*(2), 1079-1089.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *27*(1), 134-144.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, *112*(1), 272-284.

Bronkhorst, A. W., & Plomp, R. (1989). Binaural speech intelligibility in noise for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *86*(4), 1374-1383.

Brungart, D. S., Chang, P. S., Simpson, B. D., & Wang, D. (2009). Multitalker speech perception with ideal time-frequency segregation: Effects of voice characteristics and number of talkers. *The Journal of the Acoustical Society of America*, *12* (6), 4006-4022.

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, *110*(5), 2527-2538.

Bunnell, H.T., (1990). On enhancement of spectral contrast in speech for hearing-impaired listeners. *Journal of the Acoustical Society of America, 88,* 2546-2556.

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

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, *25*(5), 975-979.

Clarkson, P. M., & Bahgat, S. F. (1991). Envelope expansion methods for speech enhancement. *The Journal of the Acoustical Society of America*, *89*(3), 1378-1382.

Coughlin, M., Kewley-Port, D., & Humes, L. E. (1998). The relation between identification and discrimination of vowels in young and elderly listeners. *The Journal of the Acoustical Society of America*, *104*(6), 3597-3607.

Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, speech, and hearing services in schools*, *31*(4), 362-370.

Deepthi M. (2012). *Perception of Spectrally Enhanced Speech through Companding in Individuals with Cochlear Hearing Loss*, unpublished masters dissertation, University of Mysore, Mysore.

Dorman, M. F., Studdert-Kennedy, M., & Raphael, L. J. (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics*, *22*(2), 109-122.

Dubno, J. R., & Dirks, D. D. (1982). Evaluation of hearing-impaired listeners using a nonsense-syllable test I. Test reliability. *Journal of Speech, Language, and Hearing Research*, *25*(1), 135-141.

Dubno, J. R., & Schaefer, A. B. (1995). Frequency selectivity and consonant recognition for hearing-impaired and normal-hearing listeners with equivalent masked thresholds. *The Journal of the Acoustical Society of America*, *97*(2), 1165-1174.

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

Dubno, J. R., Dirks, D. D., & Schaefer, A. B. (1989). Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. II: Articulation index predictions. *The Journal of the Acoustical Society of America*, *85*(1), 355-364.

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

Erber, N. P. (1972). Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. *Journal of Speech, Language, and Hearing Research*, *15*(2), 413-422.

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *112*(1), 259-271.

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

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, *12*(1), 47.

Fletcher, H., 1953. *Speech and Hearing in Communication*. Van Nostrand Company, Inc., New York.

Florentine, M., Buss, S., Scharf, B., Zwicker, E. (1980). Frequency selectivity in normally-hearing and hearing impaired observers. *Journal of Speech and Hearing Research, 23,* 646-669.

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

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America, 115*(5), 2246-2256.

Ghitza, O. (2001). On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception. *The Journal of the Acoustical Society of America*, *110*(3), 1628-1640.

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

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

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.

Hall, J. W. (2006). *New handbook of auditory evoked responses*. New Jersy: Pearson publishers.

Hazan, V., Simpson, A., & Huckvale, M. (1998). Enhancement techniques to improve the intelligibility of consonants in noise: speaker and listener effects. In *ICSLP*.

Hazrati, O., Sadjadi, S. O., Loizou, P. C., & Hansen, J. H. (2013). Simultaneous suppression of noise and reverberation in cochlear implants using a ratio masking strategy. *The Journal of the Acoustical Society of America*,*134*(5), 3759-3765.

Helfer, K. S., & Huntley, R. A. (1991). Aging and consonant errors in reverberation and noise. *The Journal of the Acoustical Society of America*,*90*(4), 1786-1796.

Hopkins, K., & Moore, B. C. (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *The Journal of the Acoustical Society of America*, *122*(2), 1055-1068.

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

Hornsby, B. W., Ricketts, T. A., & Johnson, E. E. (2006). The effects of speech and speechlike maskers on unaided and aided speech recognition in persons with hearing loss. *Journal of the american academy of audiology*, *17*(6), 432-447.

Humes, L. E., & Roberts, L. (1990). Speech-Recognition Difficulties of the Hearing-Impaired ElderlyThe Contributions of Audibility. *Journal of Speech, Language, and Hearing Research*, *33*(4), 726-735.

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

Johnson, D., Whaley, P., & Dorman, M. F. (1984). Processing of cues for stop consonant voicing by young hearing-impaired listeners. *Journal of Speech, Language, and Hearing Research*, *27*(1), 112-118.

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

Khanna, S. M., & Leonard, D. G. (1982). Basilar membrane tuning in the cat cochlea. *Science*, *215,* 305-306.

Kiang, N. Y. S., & Moxon, E. C. (1974). Tails of tuning curves of auditory-nerve fibers. *The Journal of the Acoustical Society of America*, *55*(3), 620-630.

Kiang, N. Y. S., Moxon, E. C., & Levine, R. A. (1970). Auditory-Nerve Activity in Cats with Normal and Abnormal Cochleas. In *Ciba Foundation Symposium-Sensorineural Hearing Loss* (pp. 241-273). John Wiley & Sons, Ltd..

Kluk, K., & Moore, B. C. (2005). Factors affecting psychophysical tuning curves for hearing-impaired subjects with high-frequency dead regions. *Hearing research*, *200*(1), 115-131.

Kompis, M., & Dillier, N. (1994). Noise reduction for hearing aids: Combining directional microphones with an adaptive beamformer. *The Journal of the Acoustical Society of America*, *96*(3), 1910-1913.

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

LaRiviere, C., Winitz, H., & Herriman, E. (1975). The distribution of perceptual cues in English prevocalic fricatives. *Journal of Speech, Language, and Hearing Research*, *18*(4), 613-622.

Lee, L. W., & Humes, L. E. (1993). Evaluating a speech-reception threshold model for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *93*(5), 2879-2885.

Lippmann, R. P., Braida, L. D., & Durlach, N. I. (1981). Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss. *The Journal of the Acoustical Society of America*, *69*(2), 524-534.

Liu, C., & Kewley-Port, D. (2004). Vowel formant discrimination for high-fidelity speech. *The Journal of the Acoustical Society of America*, *116*(2), 1224-1233.

Loeb, G. E., White, M. W., & Merzenich, M. M. (1983). Spatial cross- correlation. *Biological cybernetics*, *47*(3), 149-163.

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

Lorenzi, C., Debruille, L., Garnier, S., Fleuriot, P., & Moore, B. C. (2009). Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. *The Journal of the Acoustical Society of America*, *125*(1), 27-30.

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

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

McMahon, C. M., & Patuzzi, R. B. (2002). The origin of the 900 Hz spectral peak in spontaneous and sound-evoked round-window electrical activity.*Hearing research*, *173*(1), 134-152.

Moore, B. C., & Glasberg, B. R. (1987). Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hearing research, 28*(2), 209-225.

Moore, B. C., & Glasberg, B. R. (1993). Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech. *The Journal of the Acoustical Society of America,94*(4), 2050-2062.

Moore, B. C., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, *45*(4), 224-240.

Moore, B. C., Glasberg, B. R., & Simpson, A. (1992). Evaluation of a method of simulating reduced frequency selectivity. *The Journal of the Acoustical Society of America*, *91*(6), 3402-3423.

Moore, B. C., Peters, R. W., & Glasberg, B. R. (1993). Detection of temporal gaps in sinusoids: Effects of frequency and level. *The Journal of the Acoustical Society of America*, *93*(3), 1563-1570.

Moore, B.C.J. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*. England: John Wiley & sons Ltd.

Nabelek, A. K., & Pickett, J. M. (1974). Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners. *Journal of Speech, Language, and Hearing Research*, *17*(4), 724-739.

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

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.

Nejime, Y., & Moore, B. C. (1997). Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America*, *102*(1), 603-615.

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

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

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.

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, *63*(2), 533-549.

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

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *The Journal of the Acoustical Society of America*, *83*(6), 2322-2327.

Preminger, J., & Wiley, T. L. (1985). Frequency selectivity and consonant intelligibility in sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *28*(2), 197-206.

Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *The Journal of the Acoustical Society of America*, *114*(1), 446-454.

Revoile, S., Pickett, J. M., Holden, L. D., & Talkin, D. (1982). Acoustic cues to final stop voicing for impaired-and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *72*(4), 1145-1154.

Rhode, W. S. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *The Journal of the Acoustical Society of America*, *49*(4B), 1218-1231.

Richie, C., Kewley-Port, D., & Coughlin, M. (2003). Discrimination and identification of vowels by young, hearing-impaired adults. *The Journal of the Acoustical Society of America*, *114*(5), 2923-2933.

Robinson, J. D., Baer, T., & Moore, B. C. (2007). Using transposition to improve consonant discrimination and detection for listeners with severe high-frequency hearing loss: La utilización de la transposición para mejorar la discriminación consonántica y la detección en oyentes con hipoacusia severa para frecuencias agudas. *International Journal of Audiology*, *46*(6), 293-308.

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

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

Schooneveldt, G. P., & Moore, B. C. (1987). Comodulation masking release (CMR): Effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band. *The Journal of the Acoustical Society of America*, *82*(6), 1944-1956.

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

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. (1989).Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners. *Acta Oto-Laryngology, 469,* 101-107.

Skinner, M. W. (1980). Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation. *The Journal of the Acoustical Society of America*, *67*(1), 306-317.

Sperry, J. L., Wiley, T. L., & Chial, M. R. (1997). Word recognition performance in various background competitors. *Journal-American Academy of Audiology*, *8,* 71-80.

Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., & Mott, J. (1985). Speech perception ability and psychophysical tuning curves in hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *77*(2), 620-627.

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

Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M., & Kurowski, K. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *The journal of the acoustical society of America*, *91*(5), 2979-3000.

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.

Stone, M. A., & Moore, B. C. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *The Journal of the Acoustical Society of America*, *114*(2), 1023-1034.

Subtelny, J.D. (1983). Patterns of performance in speech perception and production. In Hochberg trving, Levitt and Osberger. *Speech of the Hearing Impaired: Research Training and Personnel Preparation.* Baltimore: University Park Press.

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.

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

Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream functional differences between vowels and consonants. *Psychological Science*, *19*(2), 137-144.

Trees, D. E., & Turner, C. W. (1986). Spread of masking in normal subjects and in subjects with high-frequency hearing loss. *International Journal of Audiology*,*25*(2), 70-83.

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

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

Van Tasell, D. J., Fabry, D. A., & Thibodeau, L. M. (1987). Vowel identification and vowel masking patterns of hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *81*(5), 1586-1597.

Vandana S. (1998) *Speech Identification Test For Kannada Speaking Children*, unpublished masters dissertation, submitted to University Of Mysore, Mysore.

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

Walden, B. E., & Montgomery, A. A. (1975). Dimensions of consonant perception in normal and hearing-impaired listeners. *Journal of Speech, Language, and Hearing Research*, *18*(3), 444-455.

Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: A study of perceptual features. *The Journal of the Acoustical Society of America, 54*(5), 1248-1266.

Wang, M. D., Reed, C. M., & Bilger, R. C. (1978). A comparison of the effects of filtering and sensorineural hearing loss on patterns of consonant confusions. *Journal of Speech, Language, and Hearing Research, 21*(1), 5-36.

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

YH Tsui, Valter Ciocca, I. (2000). Perception of aspiration and place of articulation of Cantonese initial stops by normal and sensorineural hearing-impaired listeners. *International Journal of Language & Communication Disorders*, *35*(4), 507-525.

Yost, W. A. (1997). The cocktail party problem: Forty years later. In *Binaural and Spatial Hearing in Real and Virtual Environments,* Gilkey, R. A., & Anderson, T. R., (eds), 329-348. Erlbaum: Mahwah, New Jersy.

Zurek, P. M., & Delhorne, L. A. (1987). Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment. *The Journal of the Acoustical Society of America*, *82*(5), 1548-1559.