Critical Band based Frequency Compression and Speech Perception in Noise in Individuals with Cochlear Hearing Loss

ARF project SH/CDN/ARF/4.06/2011-12

Principal investigator
Dr. Ajith Kumar U

Co- investigator
Mr. Arivudai Nambi

Personnel
Mr. Praveen H. R.

Table of Contents

Abstract

Critical band based compression has been proven to be useful in improving speech intelligibility in quiet. However, performance of the algorithm in the presence of noise is a concern, especially when the competing speech itself is a noise. Purpose of this study was to evaluate usefulness of critical band compressed speech on speech perception in noise in individuals with cochlear hearing loss. Specifically, this study measured signal to noise ratio required to identify 50% of the speech presented (SNR-50) in individuals with cochlear hearing loss with and without critical band compression. SNR-50 was assessed with 20% and 50% compression factors. Related to this, the study also investigated frequency resolution and stream segregation abilities. Frequency resolution was assessed by measuring auditory filter widths using notched-noise method and stream segregation abilities were evaluated by measuring concurrent vowel identification scores. Furthermore, individuals' ability to differentiate fundamental frequency was evaluated by measuring difference limen for fundamental frequency (F0DL) of a harmonic complex. Results revealed that participants in the cochlear hearing loss group had higher F0DL and wider auditory filters when compared to that with the normal hearing group. On concurrent vowel identification task, normal hearing participants showed increase in feature information transmitted as the differences in pitch between target and reference vowels increased. However, in individuals with hearing impairment, this improvement was seen only when the pitch difference was four semitones. Critical band based frequency compression improved speech identification ability of individuals with hearing impairment in noise. Critical band

based frequency compression might be compensating for the deleterious effects caused by auditory filter widening. Implementing this algorithm in real time in digital hearing aids may be of great benefit for individuals with hearing loss.

INTRODUCTION

Human cochlea can be thought of as a series of band pass auditory filters with slight overlap between the tails of the adjacent filters (Moore, 2003). When a complex broadband signal such as speech or other environmental sounds are given as input to cochlea, they are split into several band-passed signals by the auditory filters. One evident role of these auditory filters is to resolve spectral peaks in the speech signal. Each of these narrow band auditory filters corresponds to one position on the basilar membrane with an area of around 0.89mm (Moore & Glasberg, 1990). Generally each auditory filter has different band widths; low frequencies have narrow band width and high frequencies have broad band width (Moore & Glasberg, 1983; Moore & Glasberg, 1990). Frequency selectivity or the ability to separate the frequency components of complex stimuli largely depends on the filtering mechanism of cochlea (Evans, Pratt, & Cooper, 1989). Several studies have reported that frequency selectivity is one of the crucial factors for speech intelligibility (Bear, Moore, & Glassberg, 1999; Sommers & Humes, 1993).

Individuals with hearing impairment often report of inability to understand speech especially in the presence of background noise. Turner, Fabry, Barrett, and Horwitz (1992) reported that individuals with cochlear hearing loss required larger signal to noise ratio (SNR) to understand target speech when compared to individuals with normal hearing. Furthermore, it has been shown that individuals with cochlear hearing loss demonstrate poor identification of consonants and vowels in both quiet and noisy conditions compared to individuals with normal hearing. One of the factors that may be contributing to the speech perception problems of individuals with hearing impairement is their reduced frequency resolution. It has been reported that auditory filters are widened in individuals with cochlear hearing loss and thus resulting in poor frequency selectivity (Carney & Nelson, 1983; Festen & Plomp, 1983; Florentine, Buus, Scharf, & Zwicker, 1980; Glasberg & Moore, 1986). Poor consonant and vowel identification could be because of inability to resolve formant transitions and formant frequencies due to widened auditory filters. Identification of target signal in the presence of background noise depends on SNR at each auditory filter. Widening of auditory filters reduces the SNR at each auditory filters and thereby reducing the speech intelligibility. Reduced spectral resolution results from masking of frequency components in an auditory filter by the components in adjacent filter.

One approach to compensate for the reduced frequency selectivity is enhancing spectral contrast in the input signal itself. Investigators have tried variety of techniques to achieve this. Enhancing the spectral contrast by increasing the difference between spectral peaks and dips resulted in limited benefit in speech intelligibility (Bear, Moore, & Gatehouse, 1993). However, this technique did not reach the clinical platform due to lack of positive outcome with this algorithm (Simpson, Moore, & Glasberg, 1990). Kulkarni, Pandey and Jangamashetti (2006) proposed the binaural processing method to address the problem of reduced frequency resolution in individuals with cochlear hearing loss. In this method the input speech was divided into series of band passed signals using auditory critical band function. Then the set of odd-numbered bands was presented to the participant's right ear and the even numbered bands to the left ear. This technique improved speech clarity and authors attributed this improvement to reduced spectral masking (Kulkarni, Pandey, & Jangamashetti, 2006). However this technique would be useful only for the individuals with bilateral hearing loss with similar auditory filter characteristics and also for the individual who wish to use hearing aids in both the ears.

Yasu et al. (2002) proposed a novel method called critical band based frequency compression to overcome the effect of spectral masking. In this technique, the speech signal is divided into a number of bands, and the spectrum of each band is compressed towards the centre frequency using a constant factor. Listeners reported that critical band compressed speech was more clear than the original speech, which could be because of the reduced interference between the adjacent filters. Series of studies have shown that critical band based frequency compression improves the speech intelligibility (Kulkarni, Pandey, & Jangamashetti, 2009; Yasu et al., 2002).

Need for the Study

Critical band based compression has been proven to be useful in improving speech intelligibility in quiet (Kulkarni et al., 2009; Yasu, Hishitani, Arai, & Murahara, 2004; Yasu et al., 2002). However, performance of the algorithm in the presence of noise is a concern, especially when the competing speech itself is a noise. When the individual is forced to listen to target speech in the presence competing speech, one acoustic cue which helps the individual to perceptually segregate is the difference in fundamental frequency (F0) or the difference in pitch (Qin & Oxenham, 2003). Essential cues for the perception of pitch lies in both resolved lower harmonics and unresolved higher harmonics (Micheyl & Oxenham, 2004). Whether the critical band based frequency compression preserves the essential cues for the perception of speech in presence of competing speech babble is not known. Individual's ability to differentiate two stimuli based on F0 can be assessed by i) measuring difference limen of F0 (listener's ability to differentiate two harmonic complexes that differ in F0) and ii) concurrent or double vowel identification (istener's ability to identify the monaural simultaneous presentation of two vowels). These two measures assess the listener's ability to use F0 cues in extracting the relevant speech cues and may determine the benefit obtained by the critical band compressed speech.

Yasu et al. (2004) studied the effectiveness of critical band based frequency compression in two individuals with profound hearing loss. One participant performed well with a compression ratio of 20% but another participant obtained maximum scores with a compression ratio of 40% . According to Kulkarni et al. (2009), a compression factor of 60% is required for good intelligibility on a critical band based frequency compression when compared to speech perception at a compression ratio of 0%. This study was carried out on 11 individuals with moderate to severe hearing loss. This discrepancy between the results of studies needs to be resolved which would help in the prescription of optimal compression factor. Whether the requirement for the different compression factor by different individuals is related to the auditory filter characteristics of the participants is the research interest.

Statement of the Problem

Purpose of this study was to evaluate usefulness of critical band compressed speech on speech perception in noise in individuals with cochlear hearing loss. Specifically, this study measured signal to noise ratio required to identify 50% of the speech presented (SNR-50) in individuals with cochlear hearing loss with and without critical band compression. SNR-50 was assessed with 20% and 50% compression factors. Related to this, the study also investigated frequency resolution and stream segregation abilities. Frequency resolution was assessed by measuring auditory filter widths using notched-noise method and stream segregation abilities were evaluated by measuring concurrent vowel identification scores. Furthermore, individuals' ability to differentiate fundamental frequency was evaluated by measuring difference limen for fundamental frequency (F0DL) of a harmonic complex.

Objectives of the Study

- To measure the F0DL in individuals with cochlear hearing loss for a harmonic complex and compare that to age matched individuals with normal hearing.
- To investigate the concurrent vowel identification ability by individuals with cochlear hearing loss and compare that with the identification of age matched individuals with normal hearing.
- To investigate the SNR-50 with 20% and 50% compression factor in individuals with cochlear hearing loss and compare that with the identification of age matched individuals with normal hearing.

- To investigate the possible relationship between signal to noise ratio required to obtain 50% correct speech identification scores (SNR-50) and other psychophysical measures.

Review of Literature

The literature relevant to current study is reviewed under the following headings:

- Fundamental frequency difference limen (F0DL) for a complex tone in individuals with normal hearing and individuals with cochlear hearing loss
- Auditory filter shapes in individuals with normal hearing and individuals with cochlear hearing loss
- Concurrent vowel identification (CVI) ability in individuals with normal hearing and individuals with cochlear hearing loss
- Critical band based frequency compression and its relevance in speech perception in noise in individuals with normal hearing and individuals with cochlear hearing loss.

F0DL for a complex tone in individuals with normal hearing and in individuals with cochlear hearing loss

The smallest detectable frequency difference between two pure tones is often referred to as the 'frequency difference limen' (FDL). Similarly, the smallest detectable difference in fundamental frequency (F0) between two complex tones is sometimes called the fundamental frequency difference limen (Plack & Oxenham, 2005). Majority of the sound signals in our environment have acoustic waveforms which repeat over time. These sounds are often perceived as having pitch that corresponds to the repetition rate of the sound. Vowel sounds associated in speech are 'voiced' and can be associated with pitch. The auditory system combines information across cochlear location in order to derive the pitch of complex tones

(Plack & Oxenham, 2005). A complex tone can be defined as any sound with more than one frequency component that evokes a sensation of pitch. A periodic complex tone consists of a series of harmonics with frequencies at integer multiples of F0 (Hartmann, 1997). Discrimination of fundamental frequency in individuals with cochlear hearing loss has been studied by several investigators. In most of the studies, the minimum difference in fundamental frequency that is required to differentiate otherwise same harmonic complexes has been studied. In general individuals with cochlear hearing loss require larger differences in fundamental frequency than normal hearing participants although there are considerable variations in the participants with hearing impairment.

The variations for fundamental frequency difference limen for complex tones in individuals with hearing impairment are reported by many authors. Moore, Glasberg, and Flanagan (2006) studied the role of component resolvability and temporal fine structure in discrimination of fundamental frequency for complex tones in normal hearing individuals. F0DLs were measured using two-interval two alternative forced-choice procedure with three down one up tracking to estimate 79% correct point on psychometric function. They found that the F0DLs were small when the lower harmonics were present but increased when the harmonic components increased to more than 8 for the centre frequency of 2000 Hz. These results suggested that the phase locking was important for obtaining the low F0DLs with resolved harmonics. Moore and Moore (2003) studied the effect of fixed and shifted spectral envelopes on F0DL. Two sets of stimuli were used for the study. One set of stimuli included fixed three harmonics. Hence, the changes in the F0 were associated with spectral cues. Other set of stimuli included shaped stimuli tones thus minimizing the spectral cues. Both the stimuli were administered to participants with normal hearing and participants with hearing impairment in three different conditions namely the tones containing resolved harmonics (RES), harmonics with intermediate resolvability (INT) and high unresolved harmonics (UNRES). The nominal F0s used were 100 Hz, 200 Hz and 400 Hz. Normally hearing participants had smaller difference limens in the fixed harmonic condition when compared to shaped stimuli in RES condition. However, difference limens were similar for the fixed and shaped stimuli for the INT condition and the UNRES condition at 100 Hz F0, suggesting that spectral cues were not used in these conditions. Participants with hearing impairment had smaller difference limens for the fixed than for the shaped stimuli, for both INT and UNRES suggesting that they used spectral cues. For the shaped stimuli, difference limens were similar in the INT and UNRES conditions for the participants with hearing impairment, but were smaller in the INT than the RES condition for the normally hearing participants. Thus, it was hypothesized that normally hearing participants used temporal fine structure cues to perform the task. The Participants with hearing impairment appeared to use only temporal envelope cues.

Gockel, Moore, Carlyon, and Plack (2007) studied the effect of duration on frequency discrimination in complex tones and on the discrimination of the fundamental frequency. F0DLs and FDLs for individual partials within a complex tones (F0=250 Hz, harmonics 1-7) were measured for stimulus duration of 200, 50 and 16ms on four normal hearing participants (18-27 years) with various degrees of musical experiences. They found that for each duration, the F0DL was lower than the FDL of any single harmonic implying that F0DLs depend upon information being combined across harmonics. Overall, large variability has been observed among the listeners with hearing impairment on the complex tone fundamental frequency discrimination task.

Auditory filter shapes in individuals with normal hearing and individuals with cochlear hearing loss

Although different methods are available to evaluate the auditory filter shapes, the majority of the studies have documented auditory filter shapes using the notched-noise masking technique (Baker & Rosen, 2006). Broadened auditory filters due to reduced frequency selectivity were reported in moderate cochlear hearing loss individuals when compared to normal hearing individuals (Bernstein & Oxenham, 2006; Leek & Summers, 1993; Moore & Carlyon, 2005).

Sommers and Humes (1993) studied the effect of age on auditory filter shapes in normal hearing individuals and individuals with cochlear hearing loss. Equivalent Rectangular Bandwidths (ERBs) were estimated for normal hearing individuals, normal hearing elderly participants and elderly participants with hearing impairment. Filter shapes were measured at 2000 Hz by modified version of the notched-noise procedure given by Glasberg and Moore (1990). Also the filter shapes for simulated cochlear hearing loss was measured for normal hearing individuals. No significant differences were seen in the ERBs measured for young normal hearing individuals and normal hearing elderly participants. The results revealed greater degrees of asymmetry in individuals with hearing impairment when compared to normal hearing individuals. Baker and Rosen (2002) evaluated the effect of sensori-neural hearing loss on frequency selectivity. Notched-noise masking method was used to derive the shape of auditory filters in individuals with mild to moderate degree of hearing loss (PTA=20-50 dB). They reported broadened auditory filter widths in individuals with hearing impairment when compared to normal hearing individuals. Peters and Moore (1992) studied the auditory filter shapes in young and elderly individuals with hearing impairment for centre frequencies of 100 Hz, 200 Hz, 400 Hz and 800 Hz using modified notched-noise method given by Glasberg and Moore (1990). Some of the individuals with hearing impairment with mild degree of hearing loss had normal auditory filters. The auditory filters tended to broaden with the increase in hearing loss. No significant differences were observed between the filter characteristics of young and elderly individuals with hearing impairment. The signal-to-noise ratio at the outputs of the auditory filters required for threshold (k) tended to be lower than the normal for young participants with hearing impairment, but were not significantly different from normal for the elderly participants with hearing impairment. It was hypothesized that the lower k values may be because of broadened auditory filters which may reduce the deleterious effect on signal detection of fluctuations in the noise. Overall, broadened auditory filters were reported in individuals with cochlear hearing loss when compared to individuals with normal hearing.

Concurrent vowel identification (CVI) ability in individuals with normal hearing and with cochlear hearing loss

In case where two voices are presented together, such as identification of mixed synthetic vowels having different fundamental frequency (F0), the differences in the F0 helps the listeners to identify the presented stimuli (Assmann & Summerfield, 1990; Culling & Darwin, 1993). When listening in quite situation, listeners rely mostly on spectral cues i.e., the energy peaks of formants coded by the place of activation along the basilar membrane for vowel identification. In the presence of competing voice, the spectral information is masked by one another and the ability to use the spectral information is diminished. In such situation temporal cues such as changes in F0 (fundamental frequency of the voice) may come into picture. Fundamental frequency was found to play an important role in the perceptual segregation of simultaneous as well as non-simultaneous sources (Bregman, 1990; Darwin & Carlyon, 1995). For normal hearing listeners, the perception of voice pitch and the ability to discriminate different F0s may depend upon the temporal fine structure (TFS) i.e., the information present in the resolved lower-order harmonics (Plomp, 1967).

de Cheveigne (1999) investigated the segregation of concurrent vowel based on the waveform interaction in terms of delta F0's. Improvements were observed in the segregation of concurrent vowel identification for the delta F0's as low as 0.4%. However, significant improvement in the identification of concurrent vowels was observed at higher delta F0's. They proposed that the reduced benefit of delta F0 for identification at smaller delta F0's more likely reflected the breakdown of the same F0-guided segregation mechanism that operates at larger delta F0's.

Vongpaisal and Pichora-Fuller (2007) studied the effect of age on F0DL and concurrent vowel identification in younger and older normal hearing individuals. They found significant improvement in the concurrent vowel identification with increase in the F0DLs up to 2 semitone differences and levelled off after an increase of more than 2 semitone difference. Arehart, Rossi-Katz, and Swensson-Prutsman (2005) studied the effect of F0DL on double vowel perception i.e., concurrent vowel perception in listeners with cochlear hearing loss and in individuals with normal hearing. The F0 was varied from 0 to 4 semitones. In all the conditions individuals with cochlear hearing loss performed significantly poorer than the normal hearing individuals. When delta F0 is 0 semitones, listeners in the hearing impaired group often perceived the presence of only 1 vowel, whereas listeners in the normal hearing group generally perceived the presence of 2 vowels. The results supported the idea of increased susceptibility to masking as the primary factor underlying the degraded perception of vowels in listeners with cochlear hearing loss. Arehart, Souza, Muralimanohar, and Miller (2011) studied the effects of age on concurrent vowel perception in acoustic and simulated electroacoustic hearing in normal hearing individuals. They reported that the perception of simultaneously presented vowels facilitated by delta F0 was more evident in unprocessed stimuli (acoustic) condition when compared to simulated electro acoustic stimuli.

Perception of monoaural presentation of concurrent vowels was studied by Assmann and Summerfield (1990). The authors tested the efficacy of place model, linear place-time model and non-linear place-time model for the perception of concurrent vowel identification. The models included the stages of frequency analysis using auditory filter, determination of the pitch, segregation of the competing speech sources and classification of derived patterns to predict the possibilities of listeners' vowel identification ability. Among the three models the concurrent vowel identification was closely matched with the non-linear place-time model of concurrent vowel identification.

Critical band based frequency compression

It is well established that the auditory filters/ critical bands are widened in individuals with cochlear hearing loss due to reduced frequency selectivity (Bernstein & Oxenham, 2006; Leek & Summers, 1993; Moore & Carlyon, 2005). This spread of energy to the neighbouring critical bands is greater with increase in the intensity of the stimulus. One of the methods which came up to address this issue is the dichotic presentation of the stimulus (Chaudhari & Pandey, 1998a, 1998b). According to this method, a speech signal was split into 18 critical bands and a set of odd-numbered bands was presented to the participant's right ear, while the rest was presented to the left ear. Although authors reported that the speech was perceived better both in individuals with normal hearing sensitivity and individuals with cochlear hearing loss, this method of speech processing would be useful to only those individuals with similar auditory characteristics in both ears. Hence a new method of speech processing was proposed by Yasu et al. (2002) for digital hearing aid called critical band compressed speech. In this technique, each critical band was compressed along the frequency axis. This was based on the notion that it would reduce the spread of energy to the neighbouring critical bands in individuals with cochlear hearing. This was supposed to restrict the information within a critical band, thus reducing the spread of information to the neighbouring auditory filters/ critical bands i.e., reducing the effect of spectral masking. Authors reported an improvement in terms of naturalness, clarity and intelligibility when the degree of compression was between 50% and 90%.

Kulkarni et al. (2009) proposed that the effect of spectral masking can be reduced by multi-band frequency compression. They evaluated the benefit of multi-band based frequency compression in individuals with sensorineural hearing loss using the modified rhyme test. Mean opinion score (MOS) test was used for evaluating the quality of the frequency compressed speech. The MOS were obtained at four different conditions, i.e., unprocessed speech stimuli, and processed stimuli with compression factor of 0.4, 0.6 and 0.8 at different signal-to-noise ratio (SNR) levels. Results revealed that six normal hearing listeners, with hearing loss simulated at different levels of broadband masking noise, showed improvement in the recognition scores for the SNR values below 3 dB. However, a monotic decrease in the recognition score were seen at the SNR levels of ∞ and 6 dB with the reduction of the compression factor from 0.8 to 0.6. At the SNR levels of below 0 dB, the improvements with compression factor of 0.6 were significantly better when compared to those with the compression factor of 0.8 and 0.4. MOS obtained at 0.6 compression factor at -6 dB SNR is similar to that obtained for unprocessed speech at 0 dB SNR. Thus the authors concluded that the compression factor of 0.6 can lead to an SNR advantage of 6 dB. The authors evaluated 11 individuals with moderate to severe sensorineural hearing loss. The participants showed 2-8%, 3-16% and 12-14% improvement in the MOS for a compression factor of 0.8, 0.4 and 0.6 conditions respectively. The authors concluded that the processing of speech signal with multi-band compression improved speech perception and the maximum improvement was observed for the compression factor of 0.6. The authors recommended the further evaluation of the multi-band based frequency compression with different types of speech material and on larger number of population.

Kulkarni, Pandey, and Jangamashetti (2012) evaluated the benefit of multi-band frequency compression in six normal hearing adults (35-45 years) and eight adults with moderate to severe sensorineural hearing loss (32-66 years). They assessed the MOS and response time as an evaluative measure at different SNR levels (∞, 6, 3, 0, -3, -6, -9, -12, and -15) and at three different compression factors (0.4, 0.6, and 0.8). For individuals with normal hearing, the processed (compressed) speech resulted in a reduction of the recognition scores at SNR values of ∞, 6 and 3 dB. However, a significant increase ($p<0.01$) in the recognition scores was reported for SNR values below 0 dB. The improvement with compression factor of 0.6 was significantly better ($p<0.001$) when compared to compression factors of 0.4 and 0.6. Comparing with the scores with unprocessed signal at -9 dB SNR approximated the recognition scores at -15dB SNR for a compression factor of 0.6, thus indicating that a compression factor of 0.6 results in SNR advantage of 6 dB. The response time for the unprocessed speech increased with decrease in SNR due to increased perceptual load. A reduction of the response time was observed for all the compression factors at lower SNR level below 1dB indicating reduced perceptual load with a maximum reduction observed at a compression factor of 0.6. In the second experiment of the study, authors evaluated the benefit of multi-band frequency compression on individuals with hearing impairment. Six out of eight had moderate increase (1%-8%) in the recognition scores at compression factor of 0.8. At a compression factor of 0.4, only four out of eight had an improvement and at the compression factor of 0.6, all of participants had improvement in the range of 9%-21% with a mean improvement of 16.5%. Compression factor of 0.6 resulted in the maximum improvement of the recognition scores. The reaction time measures showed maximum reduction at 0.6 compression factor indicating the reduced perceptual load.

Although the presentation of alternate critical band information to two ears simultaneously is reported to improve the speech intelligibility, it has its own limitations. Critical band based frequency compression seems to be a better option when compared to alternate presentation of the critical bands simultaneously to two ears based on the studies reported in the literature.

METHOD

Participants

Participants were assigned to one of the two groups based on the criteria established through audiometric evaluations. The first group consisted of 23 individuals with normal hearing sensitivity. All the participants in this group had normal hearing sensitivity with four frequency pure tone average below 15 dB HL. Their age ranged from 36 years to 58 years, with a mean age of 47.26 years and a standard deviation of 7.23 years. Immittance evaluation in the ear to be tested revealed 'A' type tympanogram with the presence of auditory reflex thresholds within normal limits. All the individuals were able to perform an open set speech identification testing and had a score of more than 80%. Participants had no other significant history of any neurological or cognitive deficits which was ascertained through a structured interview.

The second group consisted of 28 individuals with hearing loss. All the participants in this group had moderate to moderately severe sensori-neural hearing loss with a four frequency pure tone average between 32.5 dB HL and 56.6 dB HL. The participants' age ranged from 27 years to 64 years, with a mean age of 53.11 years and a standard deviation of 9.16 years. Immittance evaluation in the ear to be tested revealed 'A' type tympanogram with present/elevated/absent auditory reflex thresholds. All the individuals were able to perform an open set speech identification testing and had scores proportionate to their audiometric hearing loss (68%-100%). Speech identification scores, pure tone average, configuration of hearing loss and age of participants in group II are shown in Appendix 1. The participants had no other significant history of any neurological or cognitive deficits which was ascertained through a structured interview. Both the group of participants were selected through convenient sampling and hence the number of participants are slightly different in each group.

Auditory perception tasks

Difference limen for F0 (F0DL)

Signal processing. Stimuli were digitally generated using Matlab™ version 7 (The MathWorks.inc., USA). The stimuli were a harmonic tone complex, composed of first eight equal-amplitude harmonics with F0 of 220 Hz (standard $F0=F0_{std}$). Phase of the each harmonics was randomized on each trial. The harmonic complex tone had a duration of 500 ms with a 20 ms raised cosine onset/offset ramps. In each trial, the participant was presented with two tokens separated by a 250 ms pause. Both the tokens differed in F0. One token had a standard F0 ($F0_{std}$) and other token had a F0 which was equal to $df+F0_{std}$. 'df' was systematically varied to obtain the threshold of differential limen for F0. Both the tokens were scaled to have same intensity.

Procedure. F0DLs for the base frequency (220 Hz) was measured using a two interval two alternative forced choice paradigm (2IAFC). Standard interval contained two stimuli tokens with F0 of either $F0_{std}$ and $F0_{std}$ or df+ $F0_{std}$ and df+ $F0_{std}$. Variable interval also contained two stimuli tokens, one with the F0 of $F0_{std}$ and another with the F0 of df+ $F0_{std}$.Within the variable interval, position of the two tokens was randomized. Participant's task was to identify the interval containing the tokens having different F0. Value of the 'df' was varied in one-up two-down manner to track the threshold at 70.7% criterion point. At the beginning of the trial, 'df' was kept at 50 Hz because this difference is sufficiently large enough for all the participants discriminate easily. The stimuli were presented at most comfortable level (=45 dB HL). The value of 'df' was varied in 10% ratio step to obtain F0DL. This procedure was continued till seven reversals. The geometric mean of the mid-point of the last five reversals was considered as F0DL. Participants were given a practice trial before the actual responses were considered.

Concurrent vowel identification

Signal processing. Five vowels /a/, /e/, /i/, /o/, /u/ were synthesized at the sampling rate of 44,100 Hz, using Klatt synthesizer. Each vowel was synthesized with a base F0 of 220 Hz. Base F0 of 220 Hz was chosen to simulate the F0 of a female speaker. All the five vowels were once again synthesized with different F0 which corresponded to 1, 2, and 4 semitones increase from base F0 (Eg:- 1 semitone increase from 220 Hz corresponds to 238.3 Hz). Each vowel had duration of 290 ms with 20 ms raised cosine onset/offset ramps. All the synthesized vowels were scaled to have similar amplitude. The formant characteristics of the vowels used for the concurrent vowel identification task are given in Table 1. These values were considered based on the normative study by Hillenbrand, Getty, Clark, and Wheeler (1995).

Table 1. *Format frequencies used for synthesis of different vowels*

Vowel	F1 (in Hz)	F2 (in Hz)	F3 (in Hz)
/a/	936	1551	2815
/i/	437	2760	3372
/u/	459	1105	2735
/o/	555	1035	2828
/e/	536	2530	3047

Procedure. Identification task was carried out separately for each semitone F0 condition. Listeners were first familiarized with all the vowels in a single vowel identification task. In the familiarization task, listeners heard five repetitions of all the vowels with base F0 and with F0 shifted by different semitones. For the purpose of concurrent vowel identification, the vowels were paired with each other. Vowel /a/ was kept constant and other vowels were variable which were considered as target stimuli. Vowels /i/, /u/, /o/ and /e/ with F0 varying in four semitone steps were paired with vowel /a/ with a base F0 of 220 Hz. Vowels within the pair were presented simultaneously to one ear at a time. Concurrent vowel identification was carried out in four different conditions. In all the four conditions, vowel /a/ with a base F0 (220 Hz) was kept constant. Whereas, the variable vowels (/i/, /u/, /o/ and /e/) had F0 that were varied in semitone levels. In the first condition both the vowel /a/ and variable vowel were presented at base F0. This resulted in 0 semitone difference between the F0 of two concurrently presented vowels. In second condition, F0 of the variable vowels were increased by 1 semitone (238.3 Hz). In third and forth condition, F0 of variable vowels were increased to 2 semitone levels (256.6 Hz) and 4 semitone levels (292.2 Hz) respectively. In all the conditions, F0 of vowel /a/ was maintained at 220 Hz. APEX 3 software was used to present the randomized stimulus (Francart, Wieringen, & Wouters, 2008).

Participants were seated in a quiet room in front of a computer monitor at a distance of 1.5 feet. The four variable vowels (/i/, /u/, /o/, /e/) appeared on the screen. Stimuli were presented monaurally to the participants' at their most comfortable level using Sennheisser HD449 circumaural headphones. The participants were instructed to click on the corresponding vowel button on the computer screen after hearing the stimuli. Each of the four variable vowels was presented 10 times in a randomized order. This resulted in a total of 40 of presentations in each of the semitone condition and a total of 160 presentations for all 4 semitone difference conditions. Participants were encouraged to guess if they were not sure. The results were tabulated in confusion matrices for analysis.

Measurement of auditory filter's shape and width

Signal processing. Notched-noise method was used to estimate auditory filter shape. Auditory filter shape and width were estimated for the center frequency of 2000 Hz. We selected 2000 Hz to measure auditory filter shapes because speech intelligibility is influenced greatly from frequencies around 2000 Hz. A pure tone of 2000 Hz frequency was generated for the duration of 400 ms with 20 ms raised cosine onset/offset ramps. The detection threshold for 2000 Hz tone was measured in the presence of a broadband noise with spectral notch around center frequency of 2000 Hz. Notch width (*w*) used for the threshold estimation was 0.0, 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6. '*w*' is a normalized value expressed using the formula, $w = \Delta f/fc$. Where, fc is the center frequency and Δf is the deviation from fc. Many sinusoids except the frequencies $f \pm (w \times fc)$ were generated with random phase but equal amplitude. These random phase sinusoids were summed together to create the noise with spectral notch. The outer edge frequency of the noise bands was restricted to 0.8 times of *fc*. Noise had a total duration of 500msec with onset/offset cosine ramp of 20 ms.

Procedure. Thresholds were determined for the 2000 Hz sinusoidal tone in the presence of a noise masker with the above mentioned notch widths. The notch was placed symmetrically about the centre frequency. The level of the notched-noise was kept constant at 45 dB HL and the level of the target stimuli (2000 Hz sinusoid) was varied in order to track the threshold. The thresholds for the target signals, as a function of notch width, were tracked in a two interval two alternate forced choice paradigm. The level of the target signal was varied in a two-down one-up manner so that threshold corresponds to 70.7% point on the psychometric function. Two interval two alternate forced choice task was used to track the threshold. The starting level of the target stimuli was 30 dB above the level of the notched-noise masker. 30 dB above the notched noise was chosen which was easily discriminated from the braod band noise by the participants. 10 dB steps was used for initial two reversals, 5 dB step for next two reversals and then 2 dB step for the last 4 reversals. Mid-points of last six reversals were averaged to obtain the threshold. Thresholds obtained in the presence of notched-noise were fitted using rounded exponential (roex) model (Glasberg & Moore, 1990) to obtain the auditory filter shape. Filter parameters ‘p’, ‘r’ and ‘ERB’ were derived from the roex function.

SNR-50 with Critical band based compression

Stimuli. The speech stimuli used were standardized Kannada sentences developed by Methi, Avinash, and Kumar (2009). The test consists of seven list each containing seven sentences. The first sentence in the each list was at +20 dB SNR, second sentence was at +15 dB SNR, third sentence was at +10 dB SNR, fourth sentence was at +5 dB SNR, fifth sentence was at 0 dB SNR, sixth sentence was at -5 dB SNR and the last sentence was at -10 dB SNR. Each sentence had 5 key words contributing for a total possible score of 35 points per list. The sentences were spoken by a female native Kannada speaker. Sentences were digitally recorded in an acoustically treated room, on a data acquisition system using a 44,100 Hz sampling frequency and 16-bit analog to digital converter. Multi-talker (4 talkers) babble was used as the background noise.

Signal processing. Algorithm for critical band based compression was implemented in Matlab™ version 7 (The MathWorks.inc., USA). Input speech signal was passed through 32 band 4th order gamma tone filters. Low cut off for first band pass filter was 80 Hz and high cut off for last band pass filter was 8000 Hz. Centre frequencies and their corresponding bandwidths of gamma tone filters were determined based on parameters given by Glasberg and Moore (1990). From the output gamma tone complexes, envelopes and carriers were extracted in each band. Bandwidth of the carrier component was estimated and multiplied by the factors 1, 0.8 and 0.5. Multiplication by the factor '1' indicates no compression. Multiplication by the factors 0.8 and 0.5 indicate 20% and 50% compression respectively. Bandwidth compressed carrier was modulated using the original band specific envelope. Finally, outputs of all the filters were summed together.

Procedure. Six equivalent lists from the original test were selected for the present study. The sentences were presented through a personal computer (Lenovo Z372) at comfortable levels using Sennheisser HD449 circumaural headphones. The participant's task was to repeat the sentences presented and each correctly repeated key word was awarded one point for a total possible score of 35 points per list.

RESULTS

In the present research report, results are discussed under the following headings.

- F0DL in individuals with cochlear hearing loss for a harmonic complex and compare that to age matched individuals with normal hearing.
- Concurrent vowel identification ability by individuals with cochlear hearing loss and compare that with the identification of age matched individuals with normal hearing.
- SNR-50 with 20% and 50% compression factor in individuals with cochlear hearing loss and compare that with the identification of age matched individuals with normal hearing.
- Relationship between signal to noise ratio required to obtain 50% correct speech identification scores (SNR-50) and other psychophysical measures.

Difference Limen for F0

Figure 1 shows the F0DL values for individual participants in both the groups. Figure 2 shows mean F0DL for both the groups. The error bars indicates one standard deviation of error.

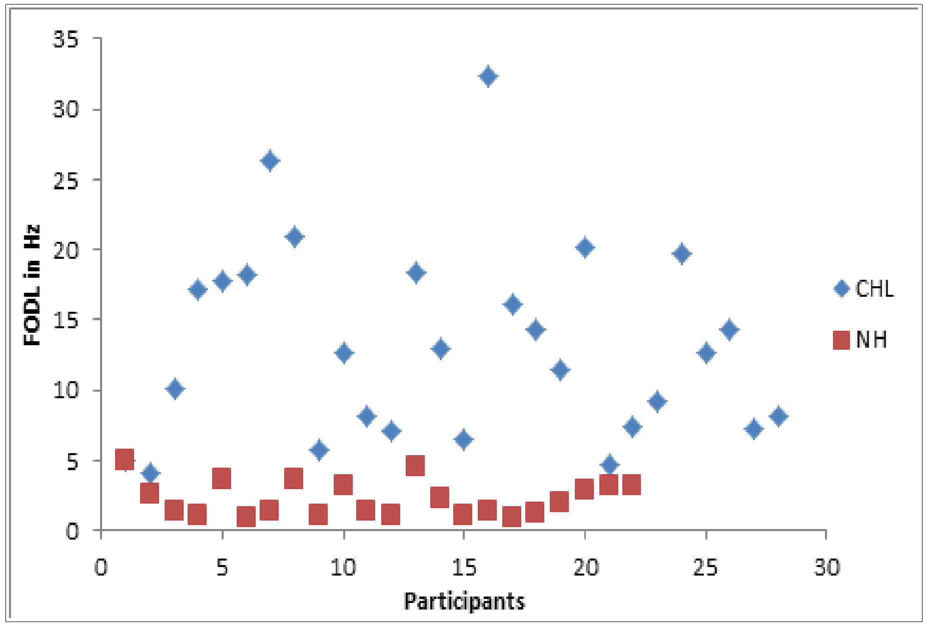

Figure 1. FODL for individual participants in both the normal hearing and cochlear hearing loss group. CHL = Cochlear hearing loss group, NH = Normal hearing group

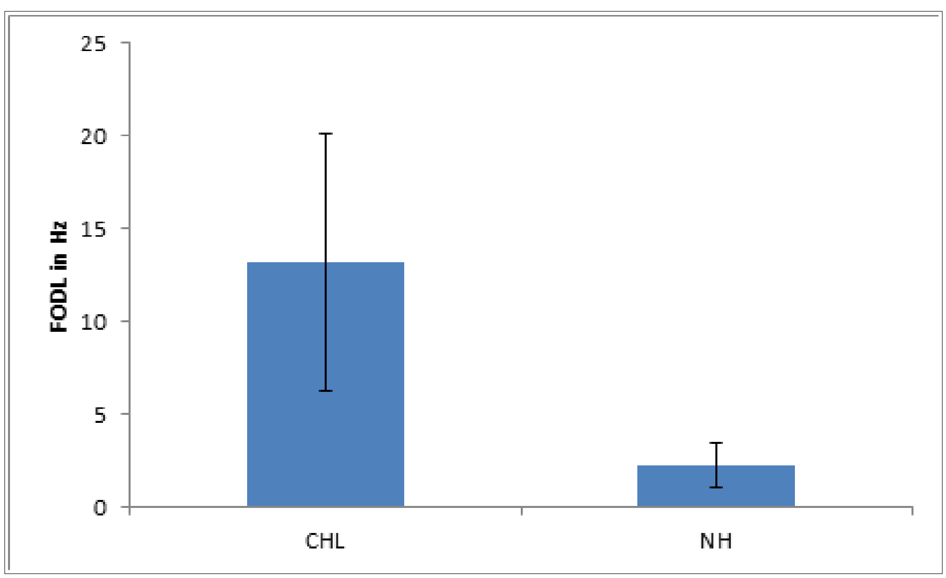

Figure 2. Mean FODL in both the normal hearing and cochlear hearing loss participants. CHL = Cochlear hearing loss group, NH = Normal hearing group

From Figures 1 and 2 it can be observed that the participants in the cochlear hearing loss group had higher F0DL values when compared to that with the normal hearing group. The mean F0DL value for the normal hearing group was 2.2 Hz whereas that of the cochlear hearing loss group was 13.17 Hz, which was almost 7 folds more. Furthermore, the cochlear hearing loss group showed larger inter-subject variations in FODL than the normal hearing group. This is evident as more scattered and larger standard deviations can be seen in the cochlear hearing loss group. An independent samples t-test showed a significant difference in F0DLs between the two groups ($t= 7.4$, $p<0.05$).

Auditory filter shape estimate

roex parameters, ERB, p and r values of individual participants in both the groups are depicted in Figures 3, 4 and 5 respectively. Arithmetic mean and standard deviations of these parameters are shown in Figure 6.

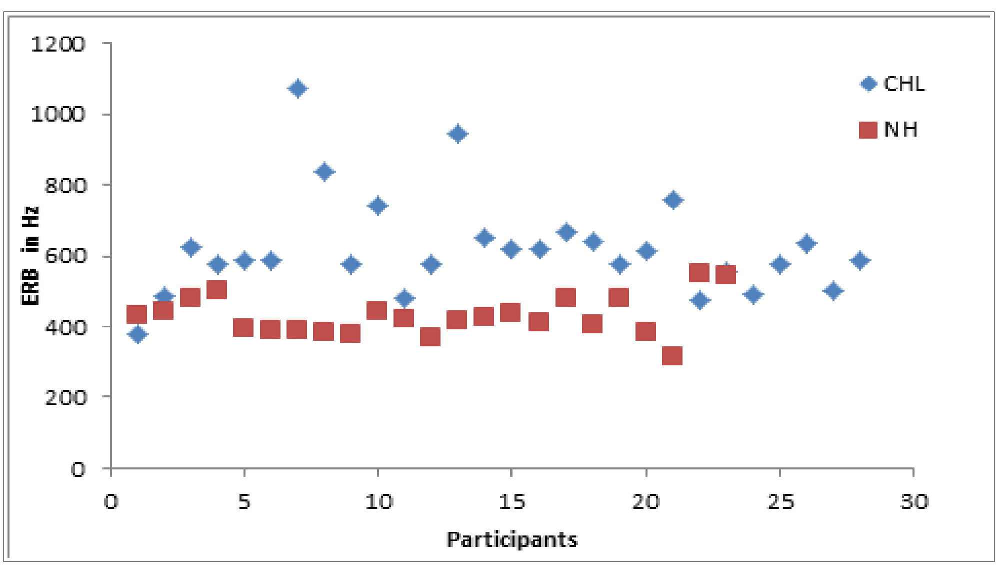

Figure 3. ERBs for individual participants in the normal hearing and cochlear hearing loss groups. CHL = Cochlear hearing loss group, NH = Normal hearing group

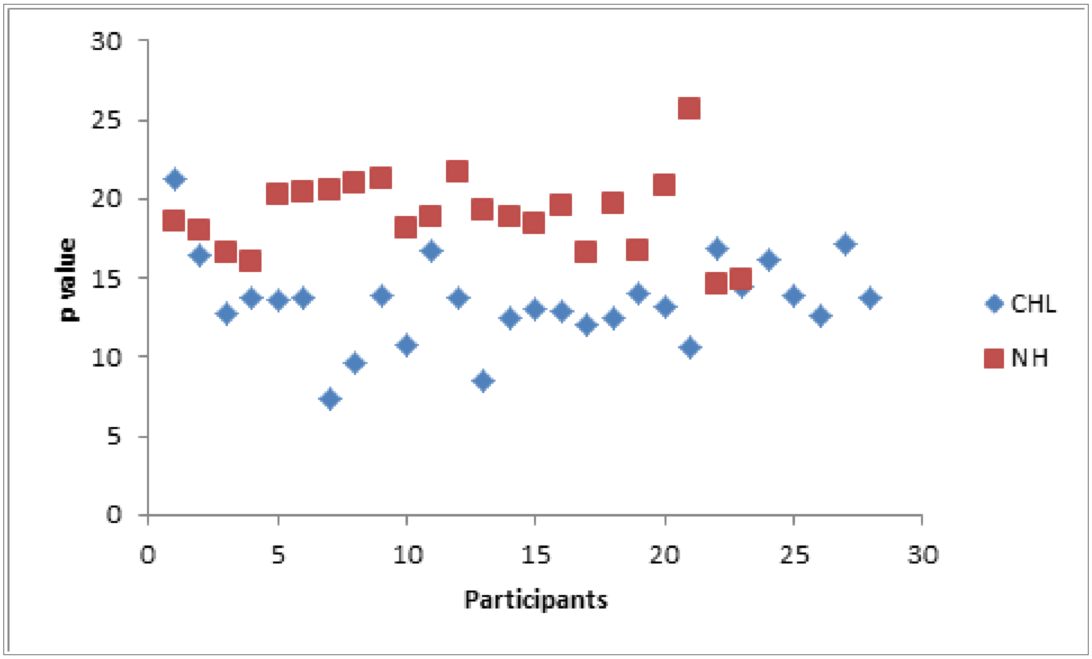

Figure 4. p values for individual participants in the normal hearing and cochlear hearing loss group. CHL = Cochlear hearing loss group, NH = Normal hearing group

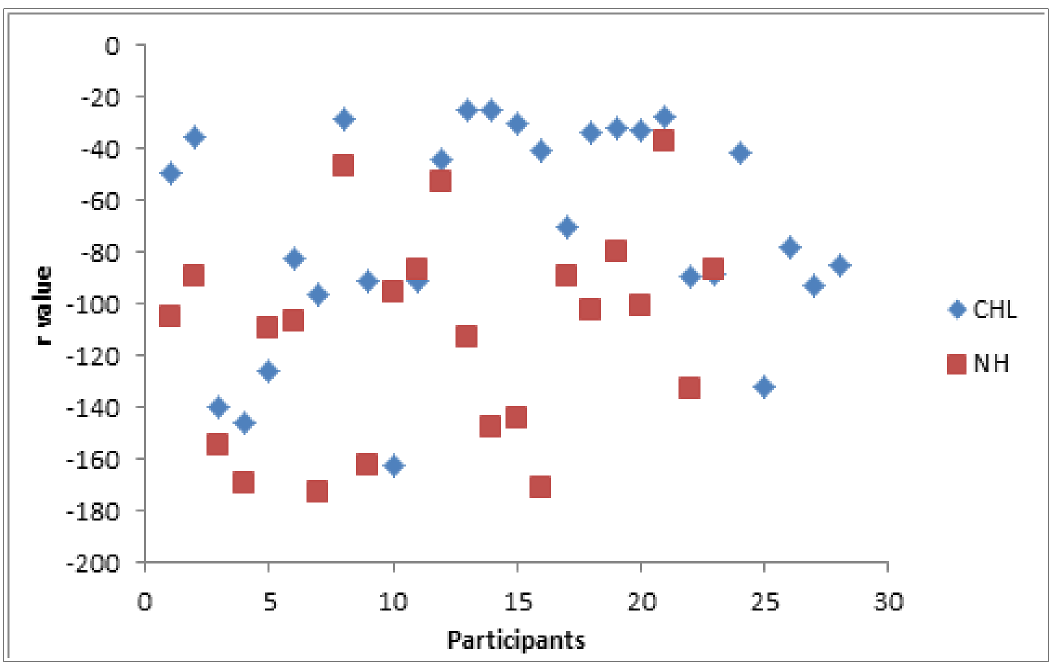

Figure 5. r values for individual participants in both the normal hearing and cochlear hearing loss participants. CHL = Cochlear hearing loss group, NH = Normal hearing group

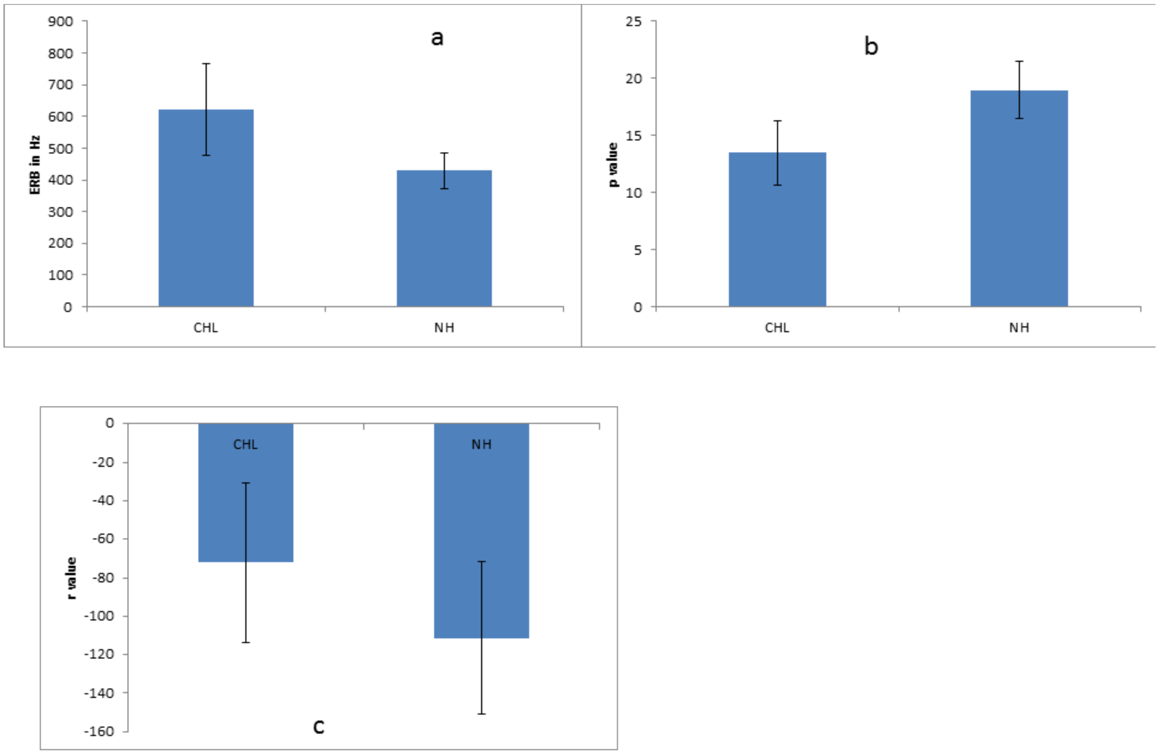

Figure 6. Mean ERB (a), p value (b) and r value (c) for normal hearing and cochlear hearing loss groups. CHL = Cochlear hearing loss group, NH = Normal hearing group

An independent samples t test revealed a significant difference in the mean ERB ($t = 6.07$, $p < 0.05$), p ($t = 2.2$, $p < 0.05$) and r ($t = 3.4$, $p < 0.05$) values between normal hearing and cochlear hearing loss group. Figure 7 shows the auditory filters drawn based upon the average values of p, r from of normal hearing and cochlear hearing loss group.

Concurrent vowel identification

Separate stimulus-response matrices were constructed for each semitone conditions for both the normal hearing and cochlear hearing loss group. In these matrices (tables 2-9), the number in each cell is the frequency with which each stimulus-response pair occurred. The numbers of correct responses can be obtained by totalling the frequency along the main diagonal (by adding the bold numbers). In each of the conditions, the stimulus concerned was presented 10 times. Concurrent vowel identification task was carried out on 15 participants in cochlear hearing loss group and 14 participants in normal hearing group as other participants did not follow instruction. Tables 2-5 shows the stimulus response matrices of vowel identification scores in normal hearing group for zero semitone, one semitone, two semitone and four semitone difference in fundamental pitch. Tables 6-9 depict the same information in cochlear hearing loss group.

Table 2. *Stimulus response matrix for 0 semitone difference in normal hearing group. The bold numbers indicate correct responses. The number in each cell reflects the combined responses from 15 participants*

	e	i	o	u
e	**41**	20	34	16
i	3	**110**	1	2
o	80	3	**89**	77
u	26	17	26	**55**
total	150	150	150	150

Table 3. *Stimulus response matrix for 1 semitone difference in normal hearing group. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from 15 participants*

	e	i	o	u
e	**91**	11	3	13
i	0	**132**	1	2
o	48	1	**135**	102
u	11	6	11	**33**
Total	150	150	150	150

Table 4. *Stimulus response matrix for 2 semitone difference in normal hearing group. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from 15 participants*

	e	i	o	u
e	**132**	3	8	9
i	0	**143**	1	4
o	15	2	**126**	46
u	3	2	15	**91**
Total	150	150	150	150

Table 5. *Stimulus response matrix for 4 semitone difference in normal hearing group. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from 15 participants*

	e	i	o	u
e	**116**	1	12	7
i	2	**147**	0	3
o	10	1	**95**	18
u	22	1	43	**122**
Total	150	150	150	150

Table 6. *Stimulus response matrix for 0 semitone difference in cochlear hearing loss group. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from 14 participants*

	e	i	o	u
e	**36**	26	32	24
i	16	**52**	38	27
o	50	29	**44**	52
u	38	33	26	**37**
Total	140	140	140	140

Table 7. *Stimulus response matrix for 1 semitone difference in cochlear hearing loss group. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from 14 participants*

	e	i	o	u
e	**28**	30	19	15
i	26	**50**	14	23
o	53	27	**69**	71
u	33	33	38	**31**
Total	140	140	140	140

Table 8. *Stimulus response matrix for 2 semitone difference in cochlear hearing loss group. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from 14 participants*

	e	i	o	u
e	**43**	27	26	28
i	32	**43**	9	17
o	43	30	**62**	51
u	22	40	43	**44**
Total	140	140	140	140

Table 9. *Stimulus response matrix for 4 semitone difference in cochlear hearing loss group. Bold numbers indicate correct responses. The number in each cell reflects the combined responses from 14 participants*

	e	i	o	u
e	**46**	25	26	32
i	35	**80**	25	18
o	22	11	**41**	29
u	37	24	48	**61**
Total	140	140	140	140

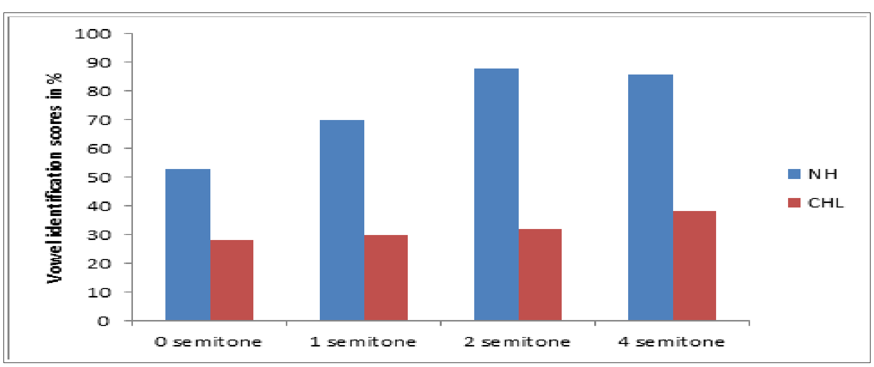

Figure 8. Mean vowel identification scores in each semitone difference conditions. NH= normal hearing group, CHL = cochlear hearing loss group.

From the confusion matrices it can be inferred that vowel /i/ was perceived better compared to other vowels in normal hearing listeners. It was also observed that in cochlear hearing loss group for zero semitone difference condition

i. Vowel /e/ was confused with /o/

ii. vowel /u/ was confused with /o/

Vowel identification scores improved and confusions decreased as the differences in the pitch between target and interfering vowel increased in both the groups. However, the amount of improvement was more in the normal hearing group than the cochlear hearing loss group. Figure 8 shows the mean vowel identification scores for both the normal and cochlear hearing loss groups. From the Figure 8 it is evident that the vowel identifications scores of cochlear hearing loss group is poorer than normal hearing group. One way repeated measures of ANOVA was performed to check the significance of difference in mean vowel identification scores. Hearing status served as between the subject factor and experimental conditions served as within the subject factor. Results showed a significant main effect of hearing status ($F(1,27)=101.1$, $p<0.05$) and experimental condition ($F(3,81)=38.4$, $p<0.05$). There was also significant interaction between the hearing status and experimental conditions ($F=(1, 27)=30$, $p<0.05$). Since there was a significant interaction between hearing status and experimental conditions separate repeated measures ANOVA was performed for each group along with Bonferroni's pair wise comparisons. Repeated measure ANOVA revealed significant main effect of experimental conditions in both the normal hearing ($F(3,39)=52$, $p<0.05$) and cochlear hearing loss group

(F(3,39)=4.1 p<0.05). Table 10 and 11 shows pair wise comparisons in different conditions in normal hearing and cochlear hearing loss group respectively. From the Table 10 and Figure 8 is evident that vowel identification scores improved significantly in normal hearing group as the pitch differences between the concurrent vowels increased up to 2 semitones. Further change in the pitch (from 2 semi tone to 4 semi tone) did not increase the vowel identification scores. However, pattern observed in cochlear hearing loss group was completely opposite. Vowel identification remained essentially same when pitch was raised till two semi tone. Further increase in the pitch to 4 semitone difference improved vowel identification scores (Table 11).

Table 10. *Results of pair wise comparisons in normal hearing group*

	0 semitone difference	1 semitone difference	2 semitone difference	4 semitone difference
0 semitone difference		Significant p<0.05	Significant p<0.05	Significant p<0.05
1 semitone difference			Significant p<0.05	Significant p<0.05
2 semitone difference				Not significant p>0.05
4 semitone difference				

Table 11. *Results of pair wise comparisons in cochlear hearing loss group*

	0 semitone difference	1 semitone difference	2 semitone difference	4 semitone difference
0 semitone difference		Not significant $p>0.05$	Not significant $p>0.05$	Significant $p<0.05$
1 semitone difference			Not significant $p>0.05$	Not significant $p>0.05$
2 semitone difference				Not significant $p>0.05$
4 semitone difference				

A sequential information transfer analysis (Wang & Bilger, 1973) was performed on group data for each experimental condition to assess the amount of information transfer from stimulus to response for a set of phonetic feature. This was done with 'Feature Information Xfer (FIX)' software from the Department of Linguistics, University College of London. Sequential Information Transfer Analysis (SINFA) is a method for determining the degree of information transfer from stimulus to response that is attributable to a particular feature. First, each feature's information transfer from stimulus to response is computed. Then a sequence of iterations is done in which one feature selected according to some criteria is partialled out per iteration by holding it constant. Typically, the feature that is transmitted to the maximum extent is held constant in subsequent iterations. Table 12 shows the different features that were assigned to the vowels. The vowels were classified based on the features of tongue height, lip rounding and place of articulation. Thus the vowels were either front or back (place), and rounded or unrounded (lip rounding) and high or low (height).

Table 12. *Features assigned to different vowels*

	e	i	o	u
Place	F	F	B	B
Lip rounding	-	-	+	+
Height	-	+	-	+

B= back vowels, F= front vowels, + = feature present, - = feature absent

Table 13 shows the relative information transmitted in bits per stimulus for each feature. Place and lip rounding features were equivalent in the vowels used and hence are not shown separately. Table 14 depicts the similar information in the cochlear hearing loss group.

Table 13. *Relative information transfer in the normal hearing listeners*

	Place/lip rounding	Height	Total information transferred
0 semitone difference	0.14	0.23	0.5
1 semitone difference	0.4	0.4	0.9
2 semitone difference	0.6	0.6	1.2
4 semitone difference	0.5	0.5	1.14

Table 14. *Relative information transfer in the cochlear hearing loss group.*

	Place/lip rounding	Height	Total information transferred
0 semitone difference	0.009	0.02	0.05
1 semitone difference	0.04	0.01	0.07
2 semitone difference	0.04	0.02	0.07
4 semitone difference	0.06	0.03	0.12

The numbers in the 2nd and the 3rd columns represent the extent of information transmitted. 0 indicates no transmission of that particular feature and a value of ‘1’ indicates maximum transmission of information. The numbers in the fourth column represent the total information transmitted. In general, the pattern of information transfer as shown by SINFA analysis was similar to the pattern found in the confusion matrices (tables 3 - 9). As the differences in pitch between target and reference vowels increased feature information transmitted improved in normal hearing listeners. However, in group hearing impairment, this improvement was seen only when the pitch difference was four semitones.

Speech perception with critical band compressed speech

Figure 9 shows the average speech identification scores obtained using three compression factors at different SNRs in normal hearing group. In Figure 10, similar information is depicted for cochlear hearing loss group.

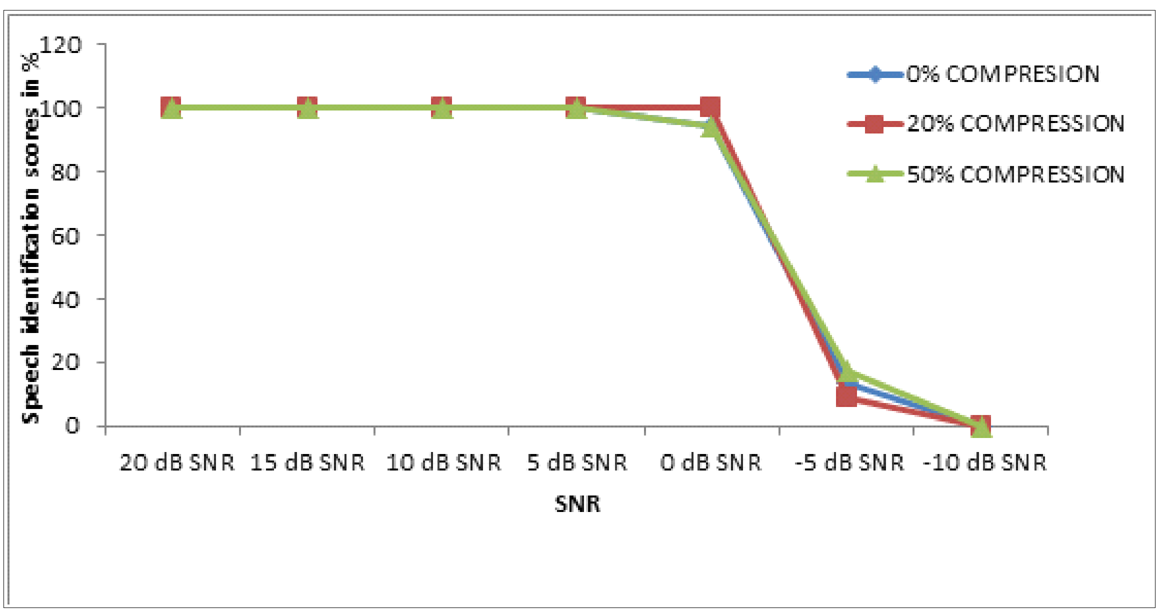

Figure 9. Speech identification scores obtained across 3 compression ratios for different SNRs in normal hearing group.

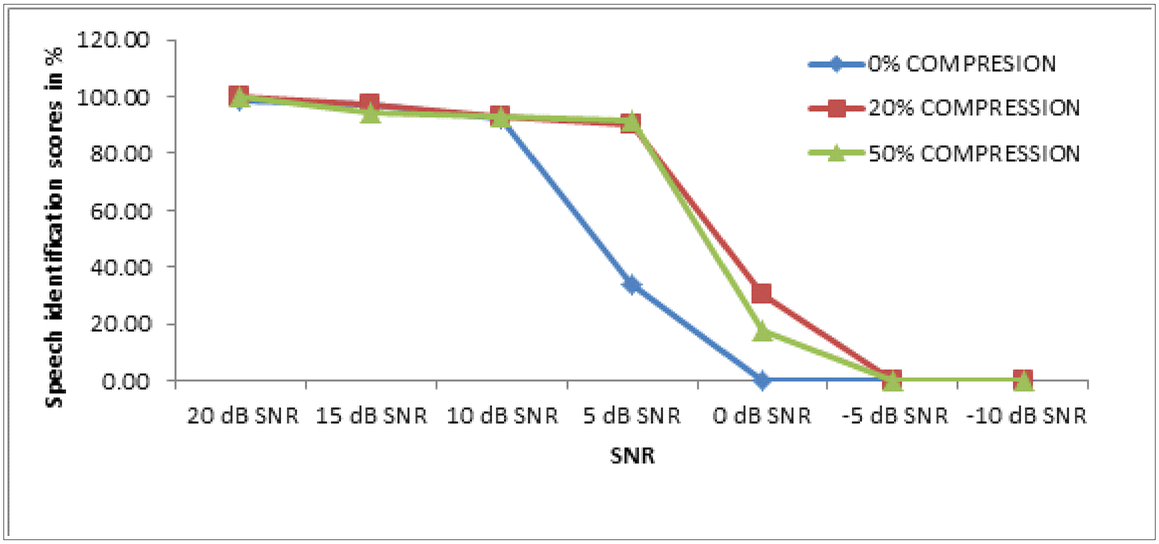

Figure 10. Speech identification scores obtained across 3 compression ratios for different SNRs in cochlear hearing loss group.

From the Figure 9 it can be seen that critical band compression did not influence the speech identification scores in the normal hearing group. However, in the cochlear hearing loss, the critical band compression (both 20% and 50% compression) enhanced the speech identification scores, especially, at 5 and 0 dB SNRs (Figure 10). Increasing the compression factor from 20% to 50% did not affect the speech identification scores much. From the above data, the threshold SNR required to obtain the 50% speech identification scores were calculated using Spearman and Karber equation (Finney, 1952):

$$50\% = i + ½(d) – (d)(\# \text{ correct})/(w)$$

Where,

i= the initial presentation level (dB S/N)
d= the attenuation step size (decrement)
w= the number of items per decrement
# = number of correctly identified words

Figure 11 shows Mean and standard deviation of SNR-50 values for different compression conditions in both the groups.

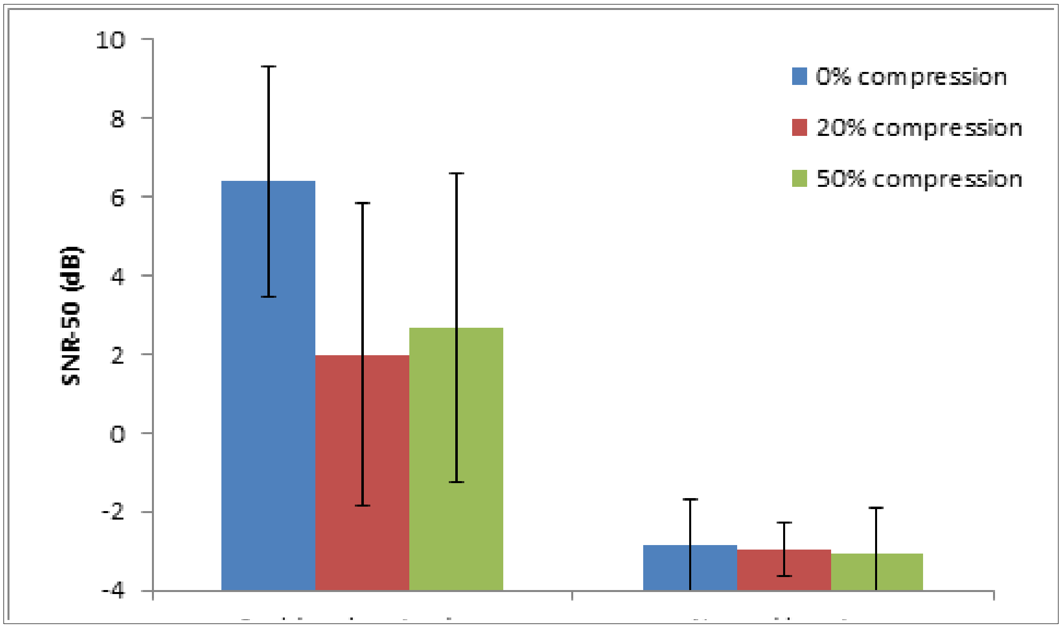

Figure 11. SNR-50 across different compression conditions in normal hearing and cochlear hearing loss group. Error bars indicate 1 standard deviation.

From the Figure 11, it is evident that the SNR-50 scores of cochlear hearing loss groups was poorer than normal hearing group. One way repeated measures of ANOVA was performed to check the significance of difference in SNR-50 values. Hearing status served as between the subject factor and compression condition served as within the subject factor. Results showed a significant main effect of compression factor ($F(2,98)=56.2$, $p<0.05$) and hearing status ($F(1,19)=83.1$, $p<0.05$) on SNR-50. There was also significant interaction between the hearing status and different compression ratios ($F=(2, 98)=50$, $p<0.05$). Since there was a significant interaction between hearing status and compression conditions separate repeated measures ANOVA was performed for each group. Repeated measure

ANOVA revealed a significant main effect of compression conditions only in cochlear hearing loss group ($F(2,54)=85.3$, $p<0.05$) but not in normal hearing group ($F(2,44)=0.3$ $p>0.05$). As normal hearing group did not show significant main effect of compression Bonferroni's pair wise comparisons were performed only for cochlear hearing loss group. Table 15 shows pair wise comparisons across different compression factors in cochlear hearing loss group. SNR-50 was significantly better with 20% and 50% compression factors compared to 0% compression condition. However, increase in the compression factor from 20% to 50% did not improve the SNR-50.

Table 15. *Pair wise comparisons of SNR-50 across different compression factors in cochlear hearing loss group*

	0% compression	20% compression	50% compression
0% compression		Significant $p<0.05$	Significant $p<0.05$
20% compression			Not significant $p>0.05$
50% compression			

Relationship between different psychophysical measures and speech perception abilities

To find the relationship between psychophysical measures and speech perception, a series of correlation analysis were performed with psychophysical measures as independent variable and SNR-50 obtained in 20% compression condition as dependent variable. For this purpose data from normal hearing and hearing impaired population were combined. SNR-50 for 20% compression condition was chosen as this condition resulted in best SNR-50 in cochlear hearing loss group and in normal hearing group compression did not have any significant influence on SNR-50. Table 16 shows correlation coefficients between SNR-50 for 20% compression and other psychophysical measures.

Table 16. Correlation coefficients between SNR-50 and other psychophysical measures

	CVI-0	CVI-1	CVI-2	CVI-4	ERB	F0DL
SNR-50	-0.5**	-0.8**	-0.9**	-0.8**	0.5**	0.6**

** $p<0.01$

CVI-0 = concurrent vowel identification with zero semitone pitch difference, CVI-1 = concurrent vowel identification with one semitone pitch difference, CVI-2 = concurrent vowel identification with two semitone pitch difference, CVI-4 = concurrent vowel identification with one semitone pitch difference, ERB = equivalent rectangular bandwidth, F0DL = difference limen for fundamental frequency

From the Table 16 it is clear that SNR-50 was significantly correlated with all other auditory measures. Negative correlations between concurrent vowel identification and SNR-50 indicate that individuals who had better concurrent vowel identification scores were able to identify the 50% of the speech presented at much lower (worse) SNRs. Positive correlations between SNR-50 and ERB indicates as the auditory filter bandwidth increases SNR required to identify 50% speech presented also increases. Similarly, positive correlation between F0DL and SNR-50 shows that individuals with better F0DL had lower SNR-50. To validate the correlations, scatter plot was drawn between SNR-50 and other auditory measures. Figure 12 shows scatter plot along with linear regression line and regression equation.

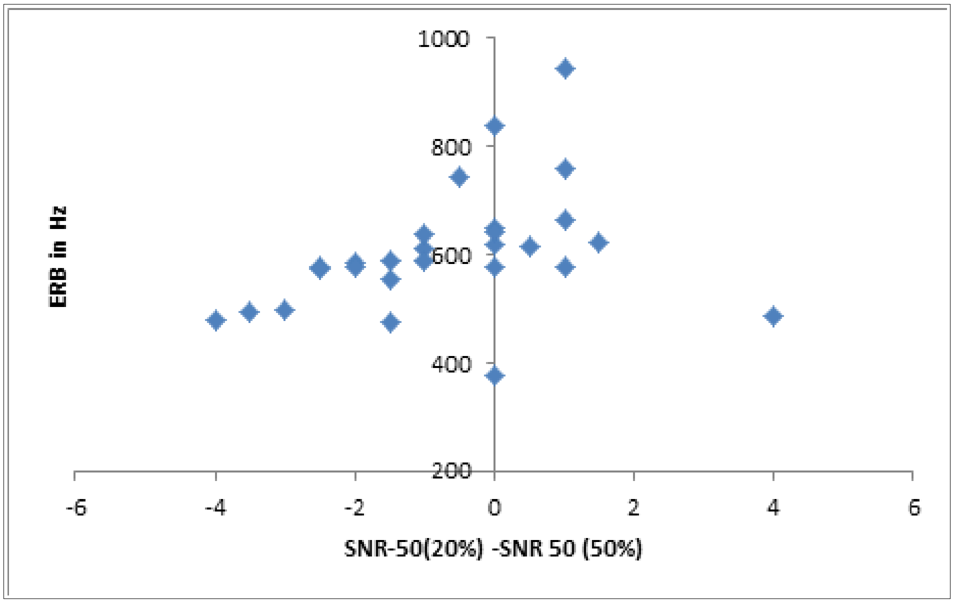

Figure 13. Scatter plot between ERB and SNR-50(20%) -SNR 50 (50%) in individuals with cochlear hearing loss. Positive values indicate SNR – 50 improved when compression factor was increased from 20 to 50% and negative values indicate SNR-50 was decreased when compression factor was increased from 20% to 50%.

Figure 13 shows scatter plot between improvements in SNR-50 when compression factor was increased from 20% to 50% for individuals with cochlear hearing loss. Positive values on the x-axis indicate that SNR-50 improved when compression factor was raised from 20 to 50% whereas, negative values indicate that SNR-50 deteriorated when compression factor was increased from 20 to 50%. From the Figure 13, it can be inferred that individuals who had wider ERBs benefited from increased compression while in individuals with relatively narrower ERB, increase in the compression factor did not change or some time even worsened SNR-50.

DISCUSSION

Psychophysical measures

The F0DLs obtained for individuals with cochlear hearing loss in our study was significantly poorer compared to F0DLs obtained for normal hearing individuals as shown in the Figure 2. The F0DLs measured for cochlear hearing group were almost 7 folds more than the normal hearing group. These results are in consensus with the results by Moore et al. (2006). This may be due to the fact that the phase locking is important for obtaining better F0DL which is reported to be lost in individuals with cochlear hearing loss. The results of the present study are also in agreement with the results by Moore and Moore (2003) who reported larger (poorer) F0DLs in individuals with hearing loss when compared to individuals with normal hearing. It can be hypothesized that the larger F0DLs in individuals with cochlear hearing loss may be due the fact that participants with hearing impairment use only temporal envelop cues or spectral cues unlike the hearing individuals who used temporal fine structure cues in the F0DL task. In the present study, a large variation in the F0DLs was observed for individuals with cochlear hearing loss when compared with individuals with normal hearing. These results are in consensus with the reports in the literature (Gockel et al., 2007; Oxenham, Micheyl & Keeber, 2009; Moore 2007; Moore & Glasberg, 2011).

The calculated auditory filter bandwidths (ERBs) obtained in cochlear hearing loss group was significantly higher when compared to normal hearing group in the present study as depicted in Figure 3 and Figure 6. Similar observations were also made by other investigators (Bernstein & Oxenham, 2006; Leek & Summers, 1993; Moore & Carlyon, 2005; Sommers & Humes, 1993). These results may be attributed to 1) reduced frequency selectivity, 2) greater spread of masking and 3) loss of capability to extract temporal fine structure cues due to reduced phase locking when compared to individuals with normal hearing.

Results of the present study (Table 2 – Table 9 and Figure 8) also showed that individuals with cochlear hearing loss had poorer ability to segregate two vowels using F0 cues compared to normal hearing group. Participants with hearing impairment were poor in identifying the two vowels especially in 0 semitone difference conditions. Similar results have been obtained by previous investigators too (Arehart et al., 2005). Various models have been developed to explain the concurrent vowel identification. In some, spectral information is used to match the most appropriate template while in others temporal cues are used for template matching. Assmann & Summerfield (1990) proposed that cochlear excitation pattern (spectral) for concurrent vowels are used for the template matching. F0 guided segregation models assume that listener will use two distinct periodicity information (temporal) to decide that two sound sources were present (de Cheveinge, 1997). These models base the template matching on pooled or average autocorrelation function. Spectro-temporal processing deficits in individuals with cochlear hearing loss may affect their ability to use combined excitation pattern or pooled auto correlation functions when two vowels are presented simultaneously.

In normal hearing listeners as the F0 differences between the concurrent vowels increased perception of both the place and tongue height information increased. In cochlear hearing loss group with the increase in the F0 difference perception of place feature increased marginally, but not the tongue height feature. This lack of improvement in vowel identification scores upon increase in the difference between F0 of concurrent vowels indicates inability of listeners with hearing impairment to use the F0 cues in stream segregation. On the other hand normal listeners can derive significant benefits when the F0 of concurrent vowels were varied. Inability of listeners with hearing impairment to use F0 frequency cues (as shown by poor performance on F0DL and concurrent vowel identification) in stream segregation may pose difficulty in understanding speech especially, when the competing signal is speech itself.

Perception of Critical Band Compressed Speech

Current study reveals that critical band based frequency compression improves speech identification ability of participants with hearing impairment in noise as shown in the Figure 10. This result is in consonance with the previous studies that evaluated speech identification quiet (Kulkarni et al, 2009; Yasu et al., 2004; Yasu et al., 2002). Critical band based frequency compression might be compensating for the deleterious effects caused by auditory filter widening. We put forth a few possible hypotheses for the mechanism by which critical band based frequency compression may be improving speech perception in noise by overcoming the adverse effect of auditory filter widening or poor frequency selectivity.

One possible mechanism by which impaired frequency selectivity might affect the speech identification in noise is related to spectral shape perception

(Moore, 2003). When compared to normal auditory filters, broader auditory filters produce relatively a smoothened representation of the spectrum. Due to this smoothening effect, spectral features become less prominent. Presence of background noise fills in the valleys between the spectral peaks and thus reduces their prominence and further aggravates speech perception difficulty. Critical band compression technique, compresses the sideband towards the centre frequency which might enhance the spectral feature contrast. This might have helped to overcome the background noise effect to some extent. It can be observed from the Figure 14 that spectral contrasts are clearly represented in excitation pattern for the compressed stimulus when compared to uncompressed stimulus. Loizou & Poroy (2001) also reported that individuals with poor frequency resolution require enhanced spectral contrast in the input stimulus. Various others studies also have shown improvement in speech intelligibility when the spectral contrast is enhanced moderately (Clarkson & Bahgat, 1991; Simpson et al., 1990). Simpson et al. (1990) implemented spectral contrast enhancement by convolving the spectrum with a Difference of Gaussians filter (DoF). In their algorithm, excitation pattern was calculated using formula given by Moore and Glasberg (1983) which resulted in a smoothed spectrum. Excitation spectrum was convolved with a DoF. DoF had a positive value in the presence of peaks and negative values in dips. Peaks were detected and multiplied by a specific factor which increased the peak to valley ratio. Yang, Luo, and Nehorai (2003) proposed a simple spectral contrast enhancement technique based on FFT method. In this method, transform of the input signal to frequency domain was obtained by performing FFT. Then the calculation of the enhancement magnitude spectrum was calculated as sum of “N” times the logarithm of spectral magnitude and the logarithm of spectral magnitude. “N” is the enhancement factor. Finally, processed speech was generated by taking the magnitude value enhanced spectrum, expressed in linear amplitude units, combined with the original phase values, and then by the reverse transformation using IFFT. Both the methods have shown to improve the speech recognition ability of participants with hearing impairment. Stone and Moore (1992) described a spectral contrast enhancement system using a 16-channel band-pass filter. Each channel generated an "activity function" that was proportional to the magnitude of the signal envelope in that channel, averaged over a short period of time. Positively weighted activity function from the ‘N’th channel was combined with negatively weighted functions from the adjacent channels, giving a correction signal used to control the gain of the band pass signal in the Nth channel. Recombining the band-pass signals resulted in a signal with enhanced spectral contrast. Results of the study revealed no improvement in speech intelligibility. Bunnell (1990) enhanced the spectral contrasts at middle frequencies, leaving high and low frequencies relatively unaffected. Spectral enhancement was performed on a spectral envelope that was calculated using a cepstral smoothing technique. However, the spectral enhancement did not improve speech performance in noise for participants with hearing impairment. This ambiguity in the literature leads to the following notions that, (i) there can be a differential effect of technique used to enhance spectral contrast on speech intelligibility as the earlier studies used different strategies, (ii) there can be large inter subject variability in realizing the benefit of spectral contrast enhancement.

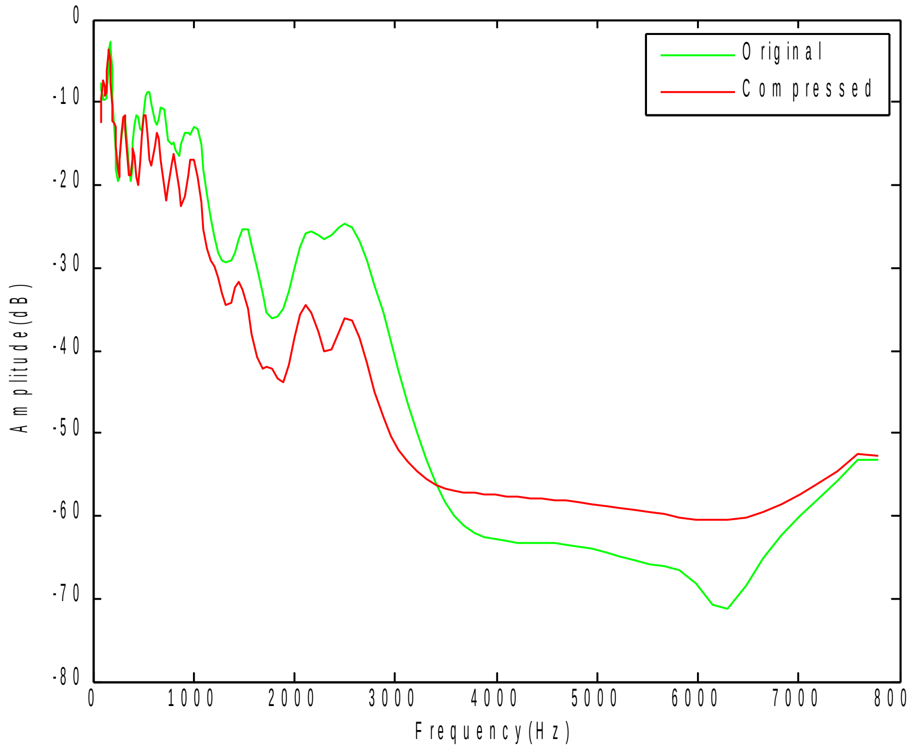

Figure 14: Excitation pattern for speech token /aba/. Green line represents excitation pattern for original unprocessed stimuli and red line represents excitation pattern for same stimuli with 20% compression.

Second possible mechanism by which speech perception ability is affected by background noise is related to increased overlap between the auditory filters. Intelligibility of speech in noise decreases as the auditory filter widens (Fletcher, 1953; Studebaker, Sherbecoe, McDaniel, & Gwaltney, 1999). Effective signal to noise ratio within the channel reduces due to widened auditory filter as a result of increased spread of excitation (Dubno, Horwitz, & Ahlstrom, 2005a; Studebaker et al., 1999). Decline in speech intelligibility due to spread of excitation may vary with the spectral content of the speech and masker (Dubno, Horwitz, & Ahlstrom, 2005b). For example, redundancy of the different spectral bands varies within the broadband speech. Some spectral bands convey more information than others. So, speech intelligibility may be largely determined by these spectral bands. When it comes to speech perception in noise, signal to noise ratio in the auditory filter which process this spectral band may be more important. In normal auditory filter, narrow band of background noise pass through the same auditory filter as that of the target band. When, auditory filter widens it allows additional spectral components which would normally pass through the adjacent auditory filter effectively in case of intact auditory filters. Generally, essential information for speech identification lies between 1000-2000 Hz and the background noise such as multi-talker babble has peak energy well below this region (Dubno et al., 2005a). Individuals with good frequency resolution are able to take advantage of this spectral separation. However, individuals with widened auditory filters are able to take less advantage due to spread of masking. This effect is graphically represented in Figure 15. It can be observed from the figure that, there is decrement in the SNR at target auditory filter, due to overlap of the adjacent filters. Because of the overlapping nature more amount of noise in the adjacent auditory filters pass the through the target auditory filter. Critical band compression approach compresses the noise in adjacent band away from the target auditory filter. This might result in release from masking. This effect is illustrated in Figure 16. When the critical bands are compressed, SNR at target filter increases as the overlap decreases. This reduced spectral masking by the adjacent bands has been proposed as the possible mechanism of improvement in speech perception in previous studies (Yasu et al., 2004; Yasu et al., 2002). To demonstrate this effect, a 1/3rd octave band noise with a peak energy at 1500 Hz (power spectrum the noise is represented in Figure 17) was sent through the 16 band ERB filter bank with 0.5 times more overlap than intact auditory filters. 0.5 times more overlap was used to simulate the effect of auditory filter widening. Input noise had maximum output at the 9th band and lesser output at higher bands. When the noise was compressed by 20%, output at the 9th band remained constant however output at the higher bands further reduced. Changes in the outputs at the higher bands are represented in Figure 18.

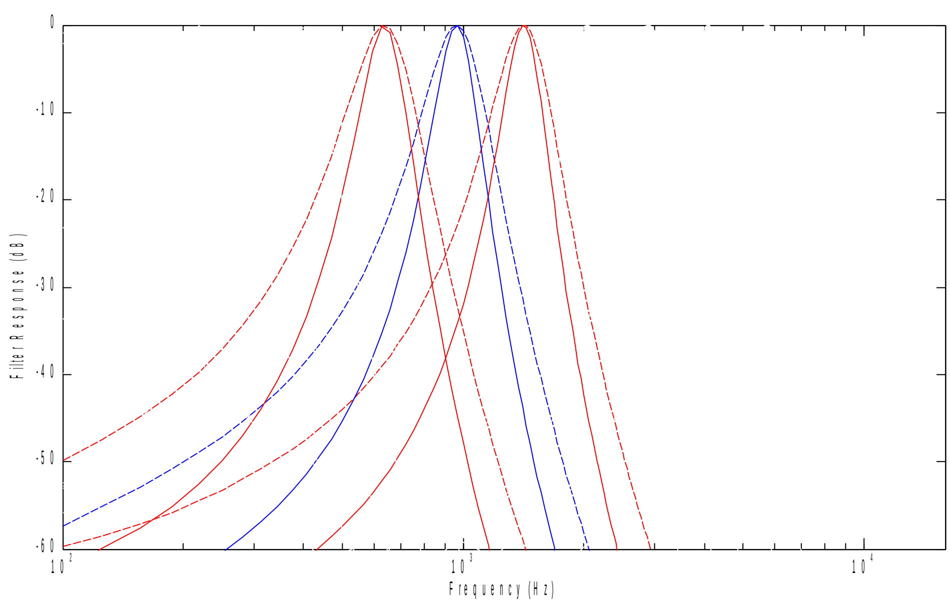

Figure 15: Simulated output of three auditory filters. Blue line indicate target auditory filter and the red lines indicate adjacent auditory filters. Continuous lines represent simulated response of intact auditory filters and broken lines represent responses of the auditory filters which are 0.5 times wider than the intact auditory filter.

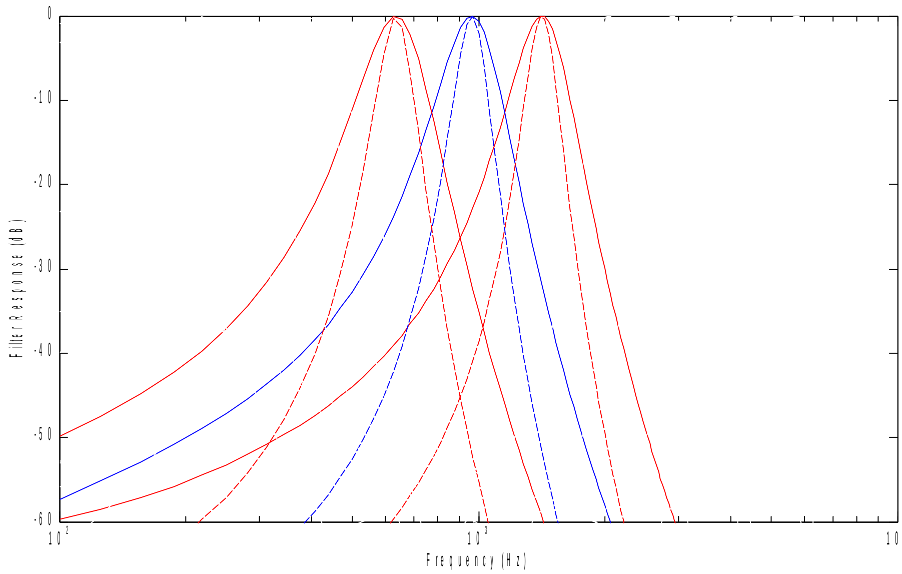

Figure 16: Simulated output of three auditory filters. Blue line indicate target auditory filter and the red lines indicate adjacent auditory filters. Continuous lines represent responses of the auditory filters which are 0.5 times wider than the intact auditory filter. Broken lines represent simulated response of auditory filters 20% percentage narrower than the intact auditory filter.

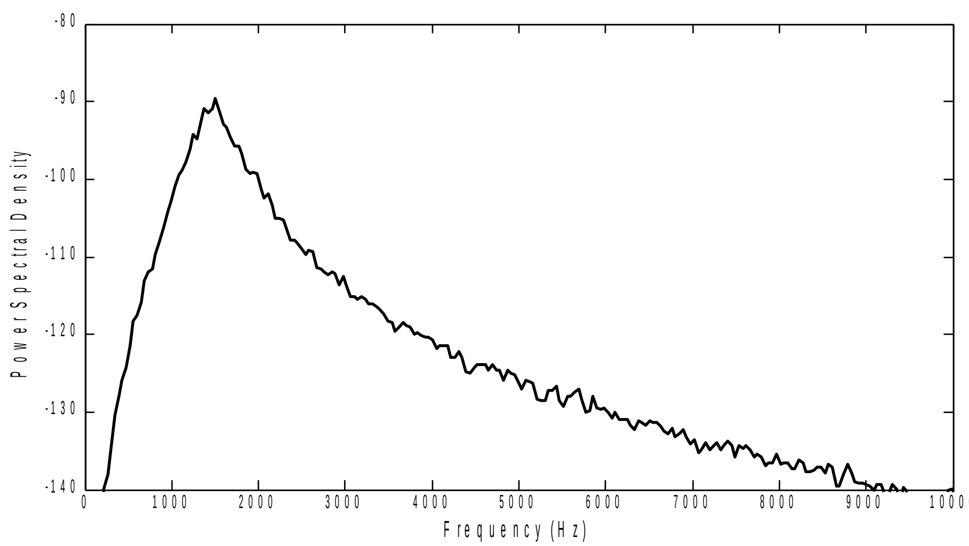

Figure 17: Power spectrum of 1/3rd octave band noise with peak energy at 1500Hz.

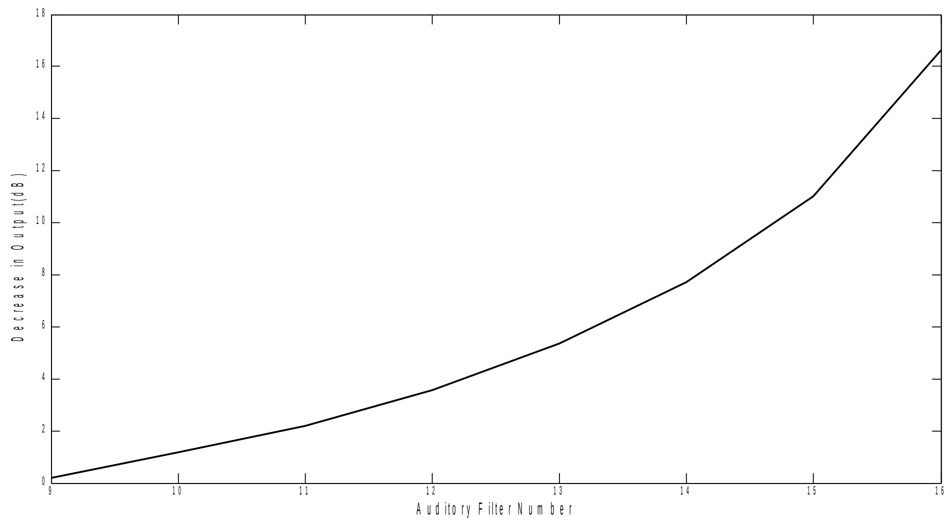

Figure 18: Simulated output of 16 band ERB filters with 50% more overlap than intact auditory filters. Line indicates relative change in the auditory filter output between unprocessed stimulus and stimulus with 20% compression at each auditory filter.

Third possible mechanism is related reduced temporal fine sensitivity associated with broader auditory filters. When complex broadband speech signals processed through the cochlea, they are broken into parallel band passed signals. Each band passed signal carries essential information in the form of envelope and fine structure (Hopkins & Moore, 2007). Envelope is a slow fluctuation in the amplitude which can produce sufficiently good speech recognition scores in quiet when it is presented alone (Shannon, 1995). Usually envelope processing is unaffected by cochlear hearing loss (Kale & Heinz, 2010; Lorenzi et al., 2006). Rapid fluctuation around each centre frequency is called as temporal fine structure (Hopkins & Moore, 2007). Temporal fine structure plays a major role perceiving speech in the presence of background noise (Lorrenzi, et al., 2006; Smith, Delgutte, & Oxenham, 2002). Temporal fine structure in the spectral region of around 1000-

2000 Hz provides the important and distinct phonetic cues. Cochlear hearing loss results in reduced ability to utilize the temporal fine structure cues (Hopkins & Moore, 2007; Moore & Sek, 2011). Broadened auditory filter could be one possible reason for impaired temporal fine structure sensitivity (Hopkins, Moore, & Stone, 2008). For a broadband sound with many components, the waveform at the output of a broader filter is much more complex than that at the output of a narrower filter centred at the same frequency. The TFS information at the output of such broad filter may be un-interpretable by the central auditory system because of its complexity (Hopkins et al., 2008). Critical band compression reduces the bandwidth at each spectral band. As a result of compression, number of frequency components processed in a single auditory filter will be reduced thus resulting in less complex temporal fine structure at auditory filter output. This less complex temporal fine structure may be interpreted with ease when compare to complex temporal fine structure. However, this hypothesis may be rudimentary as the exact relationship between auditory filter widening and temporal fine structure changes are unclear.

Summary and Conclusions

Critical band based compression has been proven to be useful in improving speech intelligibility in quiet (Kulkarni et al., 2009; Yasu, Hishitani, Arai, & Murahara, 2004; Yasu et al., 2002). However, performance of the algorithm in the presence of noise is a concern, especially when the competing speech itself is a noise. Purpose of this study was to evaluate usefulness of critical band compressed speech on speech perception in noise in individuals with cochlear hearing loss. Specifically, this study measured signal to noise ratio required to identify 50% of the speech presented (SNR-50) in individuals with cochlear hearing loss with and without critical band compression. SNR-50 was assessed with 20% and 50% compression factors. Related to this, the study also investigated frequency resolution and stream segregation abilities. Frequency resolution was assessed by measuring auditory filter widths using notched-noise method and stream segregation abilities were evaluated by measuring concurrent vowel identification scores. Furthermore, individuals' ability to differentiate fundamental frequency was evaluated by measuring difference limen for fundamental frequency (F0DL) of a harmonic complex. Results revealed that critical band based frequency compression improves speech identification ability of individuals with hearing impairment in noise. Critical band based frequency compression might be compensating for the deleterious effects caused by auditory filter widening. Implementing this algorithm in real time in digital hearing aids may be of great benefit for individuals with hearing loss.

Acknowledgment: This project was supported by the award of ARF 4.06 form the All India Institute of Speech and Hearing, Manasagangothri, Mysore. Investigators thank Director, AIISH for the support. Investigators also thank HOD, Audiology, AIISH and HOD, Speech Language Pathology and Audiology, KMC (a unit of Manipal University), Mangalore for their support.

References

Arehart, K. H., Rossi-Katz, J., & Swensson-Prutaman, J. (2005). Double-vowel perception in listeners with cochlear hearing loss: differences in fundamental frequency, ear of presentation, and relative amplitude. *Journal of Speech, Language and Hearing Research, 48* (1), 236-252.

Arehart, K. H., Souza, P. E., Muralimanohar, R. K., & Miller, C. W. (2011). Effects of Age on Concurrent Vowel Perception in Acoustic and Simulated Electroacoustic Hearing. *Journal of Speech, Language and Hearing Research, 54* (1), 190-210.

Assmann, P. F., & Summerfield, Q. (1990). Modelling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of the Acoustical Society of America, 88* (2), 680-697.

Baker, R. J., & Rosen, S. (2002). Auditory filter nonlinearity in mild/moderate hearing impairment. *Journal of the Acoustical Society of America, 111* (3), 1330-1339.

Baker, R. J., & Rosen, S. (2006). Auditory filter nonlinearity across frequency using simultaneous notched-noise masking. *Journal of Acoustical Society of America, 119* (1), 454-462.

Bear, 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* (1), 49-72.

Bear, T., Moore, B. C., & Glasberg, B. R. (1999). Detection and intensity discrimination of Gaussian-shaped tone pulses as a function of duration. *Journal of the Acoustical Society of America, 106* (4), 1907-1916.

Bernstein, J. G., & Oxenham, A. J. (2006). The relationship between frequency selectivity and pitch discrimination: effects of stimulus level. *Journal of the Acoustical Society of America, 120*(6), 3916-3928.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge: MIT Press.

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.

Carney, A. E., & Nelson, D. A. (1983). An analysis of psychophysical tuning curves in normal and pathological ears. *Journal of the Acoustical Society of America, 73* (1), 268-278.

Chaudhari, D. S., & Pandey, P. C. (1998a). Dichotic presentation of speech signal with critical band filtering for improving speech perception. *Proc. IEEE-Int. Conf. Acoustics Speech Signal Proc. Seattle,* 3601-3604.

Chaudhari, D. S., & Pandey, P. C. (1998b). Dichotic presentation of speech signal using critical band filtering for bilateral sensorineural hearing impairment. *Proc. 16th Int. Congress Acoustic,* 213-214.

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

Culling, J. F., & Darwin, C. J. (1993). Perceptual separation of simultaneous vowels: Within and across-formant grouping of F0. *Journal of the Acoustical Society of America,* 93 (6), 3454-3467.

Darwin, C. J., & Carlyon, R. P. (1995). Auditory Grouping. In B. C. Moore*, Hearing.* San Diego: CA: Academic Press.

de Cheveigne, A. (1997). Concurrent vowel identification. III. A neural model of harmonic interference cancellation. *Journal of the Acoustical Society of America, 101,* 2857-2865.

de Cheveigne, A. (1999). Waveform interactions and the segregation of concurrent vowels. *Journal of the Acoustical Society of America, 106* (5), 2959-2972.

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2005a). Word recognition in noise at higher-than-normal levels: Decreases in scores and increases in masking. *Journal of the Acoustical Society of America, 118,* 914–922.

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2005b). Recognition of filtered words in noise at higher-than-normal levels: Decreases in scores with and without increases in masking. *Journal of the Acoustical Society of America, 118,* 923–933.

Evans, E. F., Pratt, S. R., & Cooper, N. P. (1989). Correspondence between behavioural and physiological frequency selectivity in the guinea-pig. *British Journal of Audiology, 23,* 151-152.

Festen, J. M., & Plomp, R. (1983). Relations between auditory functions in impaired hearing. *Journal of the Acoustical Society of America, 73* (2), 652-662.

Finney, D. J. (1952). *Statistical Method in Biological Assay.* London: C. Griffen.

Fletcher, H. (1953). *Speech and hearing in Communication.* New York: Van Nostrand.

Florentine, M., Buus, S., Scharf, B., & Zwicker, E. (1980). Frequency selectivity in normal-hearing and hearing-impaired observers. *Journal of Speech and Hearing Research, 23* (3), 646-669.

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

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

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

Gockel, H. E., Moore, B. C., Carylon, R. P., & Plack, C. J. (2007). Effect of duration on t he frequency discrimination of individual partials in a complex tone and on the discrimination of fundamental frequency. *Journal of the Acoustical Society of America, 121* (1), 373-382.

Hartmann, W. M. (1997). Signals, Sounds and Sensation. New York: Springer-Verlag.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97* (5), 3099-3111.

Hopkins, K. & Moore, B (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *Journal of the Acoustical Society of America, 122*(2), 1055-68.

Hopkins, K., Moore, B. C., & Stone, M. A. (2008). Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *Journal of the Acoustical Society of America , 123* (2), 1140-1153.

Kale., S. & Heinz, M.G. (2010) Envelope Coding in Auditory Nerve Fibers Following Noise-Induced Hearing Loss. *Journal of the Association for Research in Otolaryngology,* 657-673.

Kulkarni, P. N., Pandey, P. C., & Jangamashetti, D. S. (2006). Perceptually balanced filter response for binaural dichotic presentation to reduce the effect of spectral masking (abstract). *Journal of the Acoustical Society of America, 120,* 3253.

Kulkarni, P. N., Pandey, P. C., & Jangamashetti, D. S.(2012). Binaural dichotic presentation to reduce the effects of spectral masking in moderate bilateral sensorineural hearing loss. *International Journal of Audiology, 51*(4), 334-44.

Kulkarni, P. N., Pandey, P. C., & Jangamashetty, D. S. (2009). Multi-band frequency compression for reducing the effects of spectral masking. *International Journal of Speech Technology, 10* (4), 219-227.

Leek, M. R., & Summers, V. (1993). Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise. *Journal of the Acoustical Society of America, 94*(6), 3127-3137.

Loizou, P. C., & Poroy, O. (2001). Minimum spectral contrast needed for vowel identification by normal hearing and cochlear implant listeners. *Journal of the Acoustical Society of America, 110* (3), 1619-1627.

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

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

Micheyl, C., & Oxenham, A. J. (2004). Sequential F0 comparisons between resolved and unresolved harmonics: no evidence for translation noise between two pitch mechanisms. *Journal of the Acoustical Society of America, 116* (5), 3038-3050.

Moore, B. C. (2003). *An introduction to psychology of hearing* (5th ed.). San Diego: Academic Press.

Moore, B. C. J., & Carlyon, R. P. (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In C. J. Plack, A. J. Oxenham, R. R. Fay, & A. N. Popper, *Pitch: Neural coding and perception* (p. 252). New York: Springer.

Moore, B. C. J., & Sek, A. (2011). Effect of level on the discrimination of harmonic and frequency-shifted complex tones at high frequencies. *Journal of the Acoustical Society of America, 129,* 3206–3212.

Moore, B. C., & Glasberg, B. R. (1983). Growth of forward masking for sinusoidal and noise maskers as a function of signal delay; implications for suppression in noise. *Journal of the Acoustical Society of America, 73* (4), 1249-1259.

Moore, B. C., & Glasberg, B. R. (1990). Frequency discrimination of complex tones with overlapping and non-overlapping harmonics. *Journal of the Acoustical Society of America, 87* (5), 2163-2177.

Moore, B. C., & Glasberg, B. R. (2011). The effect of hearing loss on the resolution of partials and fundamental frequency discrimination. *Journal of the Acoustical Society of America , 130* (5), 2891-2901.

Moore, B. C., & Moore, G. A. (2003). Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelops by normally hearing and hearing-impaired subjects. *Hearing Research, 182* (1-2), 153-163.

Moore, B. C., Glasberg, B. R., & Flanagan, H. J. (2006). Frequency discrimination of complex tones; assessing the roleof component resolvability and temporal fine structure*. Journal of the Acoustical Society of America, 119* (1), 480-490.

Oxenham, A. J., Micheyl, C., & Keebler, M. V. (2009). Can temporal fine structure represent the fundamental frequency of unresolved harmonics*? Journal of the Acoustical Society America, 125* (4), 2189-2199.

Peters, R. W., & Moore, B. C. (1992). Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects. *Journal of the Acoustical Society of America, 91* (1), 256-266.

Plack, C. J., & Oxenham, A. J. (2005). The Psychophysics of Pitch. In C. J. Plack, A. J. Oxenham, & R. R. Fay, *Pitch: Neural Coding and Perception.* New York: Springer.

Plomp, R. (1967). Pitch of complex tones*. Journal of the Acoustical Society of America, 41,* 1526-1533.

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

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

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

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature, 416*, 87-90.

Sommers, M. S., & Humes, L. E. (1993). Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listeners. *Journal of the Acoustical Society of America, 93* (5), 2903-2914.

Stone, M. A., Moore, B. C. J. (1992). Syllabic compression: Effective compression ratios for signals modulated at different rates. *British Journal Audiology, 26,* 351–361.

Studebaker, G., Sherbecoe, R., McDaniel, D., & Gwaltney, C. (1999). Monosyllabic word recognition at higher-than-normal speech and noise levels. *Journal of the Acoustical Society of America, 105,* 2431–2444.

Turner, C. W., Fabry, D. A., Barrett, S., & Horwitz, A. R. (1992). Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners. *Journal of Speech and Hearing Research, 35 (4),* 942-949.

Vongpaisal, T., & Pichora-Fuller, M. K. (2007). Effect of age on F0 difference limen and concurrent vowel identification. *Journal of Speech, Language and Hearing Research, 50* (5), 1139-1156.

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

Yang, J., Luo, F-L., & Nehorai, A. (2003). Spectral contrast enhancement: algorithms and comparisons. *Speech Communication, 39* (1-2), 33-46.

Yasu, K., Hishitani, M., Arai, T., & Murahara, Y. (2004). Critical Band based frequency compression for digital hearing aids. *Acoustical Science and Technology, 25* (1), 61-63.

Yasu, K., Kobayashi, K., Shinohara, K., Hishitani, M., Arai, Y., & Murahara, Y. (2002). Frequency compression of critical band for digital hearing aids. *Proc. China-Japan Joint conf. Acoustics,* 159-162.