Speech recognition of spectrums with 'holes' by children

Register Number: 05AUD010

MANASA (R.P)

A dissertation submitted in part fulfillment for the degree of
Master of Science (Audiology)
University of Mysore, Mysore

ALL INDIA INSTITUTE OF SPEECH & HEARING,
MANSAGANGOTHRI, MYSORE-570006
APRIL 2007.

Dedicated to
My Parents and
the All Mighty

CERTIFICATE

This is to certify that this dissertation entitled **"Speech recognition of spectrums with 'holes' by children"** is a bonafide work in part fulfillment for the degree of Master of Science (Audiology) of the student Registration no: 05AUD010. This has been carried under the guidance of a faculty of this institute and has not been submitted earlier to any other university for the award of any diploma or degree.

Dr. Vijayalakshmi Basavaraj,

Director,

All India Institute of Speech & Hearing,

Mansagangothri, Mysore-570006

Mysore
April 2007

CERTIFICATE

This is to certify that this dissertation entitled **"Speech recognition of spectrums with 'holes' by children"** has been prepared under my supervision and guidance. It is also certified that this dissertation has not been submitted earlier to any other university for the award of any diploma or degree.

Asha Yathiraj

Prof. Asha Yathiraj,
Guide,
Professor of Audiology,
All India Institute of Speech & Hearing,
Mansagangothri, Mysore-570006

Mysore
April 2007

DECLARATION

This is to certify that this master's dissertation entitled **"Speech recognition of spectrums with 'holes' by children"** is the result of my own study and has not been submitted earlier to any other university for that award of any degree or diploma.

Register Number. 05AUD010

Mysore
April 2007

ACKNOWLEDGEMENT

First among the very many I express my sincere and deep sense of gratitude to my teacher and guide , **Prof. Asha Yathiraj**, Prof. of Audiology, for constant guidance and support during the course of this study and for putting up with all my short comings. Thank you ma'am for planting in me an interest for research in speech perception and cochlear implant.

I am thankful to **Prof. V. Basavraj**, Director, AIISH, for permitting me to undertake this study.

My sincere thanks to **Dr.K, Rajalakshmi**, Reader and HOD, Department of Audiology, AIISH, for permitting me to use the department facilities for my data collection.

My sincere thanks to **Mr Animesh Barman, Ms P. Manjula, Dr. C. S. Vanaja,**, for their moral support through out my academic years and for enlightening my knowledge in the field of Audiolgy.

I am also, very much thankful to **Mr. Vijaya Narne** and **Mrs. Basantalakshmi** for their timely and needful help to complete this project.

I would like to acknowledge the immeasurable love, affection and inspiration that I have got from my sisters, brothers and brother-in-laws that made my life pleasurable.

To all my classmates, thanks a lot for your love and affection that made me to be lively in class. My special thanks to Dayalda, Deepashree, Kishan and Swapna for helping me to complete this project.

I would also like to extend my thanks to first and second B. Sc guys to make my stay in hostel happier. I am grateful to Hemraj, Keerthi, Pawan, Sindhusree, Dhatree Pragati for helping me in need.

I also extend my thanks to Sudipto, Udit and Mohan for helping me in their own ways

I am thankful to Ravi, Chandan, Bimba, Sanjay, Pankajda, Sharmila and Bijoya to be with me as a pillar of support in the worst period of my life. Yours love and inspiration have taught me to struggle through out the life.

To my dear Dhana, your absence is as inspirational as your presence, as a part of life.

To my **Grand parents** for their love and faith in me.

My deepest gratitude to all my **subjects** who participated in this study- thank you all for your patience during the experiment.

Above all I thank the **LORD Jagannath**, Who showered His abundant blessing on me in each steps of my life and for completing this study.ss

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

It is generally accepted that humans rely on cues that exist across several frequency bands to understand speech. The question of how listeners use and combine information across several frequency bands when understanding speech has puzzled researchers for many decades. Speech has proven to be a robust signal that is resistive to many forms of distortion and information reduction. It can be recognized even when the spectral information is reduced to three sinusoids that track the formant transitions over time (Remez, Rubin, Pisoni & Carrell, 1981). A high degree of speech discrimination and recognition is observed even under conditions of great reduction of spectral information. (Van Tasell, Greenfield, Logemann & Nelson, 1992; Shannon, Zeng, Wygonski, Kamath & Ekelid, 1995; Turner, Souza& Forget, 1995)

Understanding how speech is perceived after being processed through a cochlear implant is a challenge. In cochlear implants a relatively small number of electrodes activate tonotopic patches of neurons with a portion of the speech signal. Here all neurons activated by an electrode are driven in a highly deterministic fashion (Van den Honert & Stypulkowski, 1984, 1987). Even with abnormalities in physiologic process, Shannon et al. (1995) found high levels of phoneme, word and sentence recognition could be achieved by adults with just four bands of information. This observation indicates how little is understood about recognition of speech under conditions of distorted spectral information.

Cochlear implantation is based on the idea that there are surviving neurons in the vicinity where electrodes are placed in the cochlea. The lack of hair cells and/or surviving neurons within the areas of cochlea essentially creates 'hole(s)' in the spectrum. The influence of the 'holes' in the spectrum in speech understanding is not well understood. It is not known, whether the spectral 'holes' can account for some of the variability in the performance among cochlear implant users (Kasturi, Loizou, Dorman & Spahr, 2002). Hence, it is of interest to find whether recognition will be affected with the set of 'hole' pattern in speech.

Shannon, Galvin and Baskent (2001) assessed the impact of size and location of spectral 'holes' in cochlear implantees and normal hearing listeners by dropping 2-8 low, middle or high frequency bands in a 20-band cochlear implant simulation. Results showed that holes in the low frequency region are more damaging than holes in the middle or high frequency region, on speech recognition. Vidya, Rima and Yathiraj (2006) evaluated perception of seven lists of words having different band rejections and one list with no modification, on thirty normal hearing adults. They found that despite information being filtered from the speech signal, perception was not altered. Based on their findings they interpreted that as many as eight adjacent electrodes could be switched off in a cochlear implant without affecting speech identification. This gives an insight of how cochlear implant users perceive speech, even when specific electrodes are switched off.

Need for the study

Various research findings contradict each other regarding number of electrodes recquired for better speech performance. Holmes, Kemker and Mervin (1987), evaluated speech perception of a patient fitted with a multichannel processor under different electrode conditions. Their study suggested that by increasing the number of programmed electrodes, the subject's speech perception improved. This finding contradicts that obtained by others (Dorman, Loizou & Rainey 1997; Shannon et al. 1995; Turner, Souza & Forget, 1995). Dorman et al. (1997) processed vowels, consonants and sentences through software emulations of cochlear implant signal processors with 2-9 output channels. They found that high levels of speech understanding could be obtained using signal processors with a small number of channels. Despite many investigations the basic question of "how many electrodes for speech information" are still to be answered.

Knowing the relation between the numbers of electrodes activated and speech recognition is necessary for future designing devices as well as during counselling. Clinically a number of conditions may necessiate reducing the number of electrodes in a cochlear implant or not programming the electrodes. This could be because the electrodes may be damaged during surgical insertion; some electrodes with excessive impedance may not function well and may not be activated. The subjects and/or parents of the subjects also need to know how many electrodes will be suitable for significant speech recognition. Hence, knowing the number of electrodes and performance through reduced number of electrodes is also important while counselling the cochlear implantee.

The effect of spectral resolution on speech recognition has received considerable attention in last few years. However, this attention is mainly concentrated on adults (Dorman et al., 1997; Fu, Zeng, Shannon & Soli., 1998; Holmes et al., 1987; Shannon et al., 1995, 2001; Vidya et al., 2006). It is theoretically and practically important to understand whether the limited spectral resolution is a key factor for children also. Eisenberg, Shannon, Martnez, Wygonski and Boothroyd (2000) reported that young children did not have sufficient developed speech pattern recognition. Thus, children may require more spectral channels than adults to obtain similar speech recognition skills. Hence, there is a need to study to assess the effects of spectral 'hole(s)' in children using cochlear implants.

Aims of the study

The present investigation aims at studying the following, using a cochlear implant simulated condition:

- Speech recognition of spectrums with "holes" by children,
- The effect of different band stop filters in speech perception in children and
- The vowel and consonantal errors due to spectral 'hole' in the low, mid and high frequency regions.

REVIEW

Human listeners rely on cues that exist across several frequency bands to understand speech but the relative importance of the frequency band differs. French and Steinberg (1947) studied the relative importance of various frequency bands by systematically low-pass and high-pass filtering the spectrum and measuring speech recognition. This was done in order to calculate articulation index. They reported when all the frequencies above 1000 Hz were passed, 90% of the syllables were recognized correctly and when the frequencies below 1000 Hz were presented, correct identification of the items declined to 27%. Although the articulation index method was found to be very successful, it did not take account the fact that listeners might combine and utilize speech information from multiple disjoint bands.

Likewise many studies have investigated the intelligibility of high-passed, low-passed (Pollack, 1948; Kryter, 1962) and band-passed-filtered speech (Warren, Riener, Bashford & Brubaker, 1995; Stickeny & Assmann, 2001). However, not many studies have investigated the perception of band-stopped-filtered speech (i.e. speech with 'holes' in the spectrum). These band stop filters selectively attenuate (i.e. remove) all frequency components of a sound between two frequencies (the high-pass and low-pass cutoff frequencies) and retain (i.e. allow to "pass" through) all frequency components below the high-pass cutoff frequency and above the low-pass cutoff frequency. Such information has been used to study the relative importance of specific frequency bands and to create spectral 'holes' by removing information from certain frequency region of the speech spectrum. These band-stopped filters have also used to simulate various conditions like switching off electrodes and providing partial information in cochlear implantee by Shannon, Galvin and Baskent (2001).

Speech recognition by cochlear implantee with altered spectral information

It has been noted by von Bekesy (1960, cited in Moller, 2000) that in the normal-hearing ear there is a well-defined mapping of spectral information onto specific locations in the cochlea. However, in a prosthetic device like a cochlear implant, in which electrodes are organized in an array to roughly estimate with the tonotopic organization of a normal cochlea, there may be a frequency-place mismatch due in part to the signal processing of the device and in part to the pathology. Local regions of damaged neurons may create a 'hole', further distorting the frequency-to-place mapping. When this frequency-place map is shifted or distorted speech recognition reduces.

If the electrode array is not fully inserted, matching frequency information to the acoustic tonotopic place results in the loss of range of frequencies that is critical for normal speech recognition. Baskent and Shannon (2004) studied speech recognition as a function of insertion depth, varying from a deep insertion of 10 electrodes at 28.8 mm to a shallow insertion of a single electrode at 7.2 mm, in four Med-El Combi 40+ users. Short insertion depths were simulated by inactivating apical electrodes. Speech recognition increased with deeper insertion, reaching an asymptotic level at 21.6 or 26.4 mm depending on the frequency-place map used. They found a strong interaction between the optimal frequency-place mapping and electrode insertion depth. In their study, frequency-place matching produced better speech recognition than compressing the full speech range onto the electrode array for full insertion ranges (20 to 25 mm from the round window). For shallower insertions (16.8 and 19.2 mm) a mild amount of frequency-place compression was better than truncating the frequency range to match the basal cochlear location. It was also observed from their study that scores for both maps reached an asymptotic level with increasing insertion depth (and hence increasing number of electrodes) before the baseline condition of 10 electrodes. The improvement in the performance stopped at the insertion depth condition of nine electrodes for the compressed map and seven electrodes for the matched map. Their findings were also consistent with previous implant studies that showed little improvement in speech recognition as the number of electrodes was increased above seven (Fishman, Shannon & Slattery, 1997; Friesen, Shannon, Baskent & Wang, 2001).

Bredberg et al. (2006) compared open set speech recognition of twenty-one postlingually deafened adults with a Med-El Combi 40/40+GB split-electrode implant to patients using a Med-El cochlear implant with a standard electrode. Speech recognition was assessed over an 18-month period. Split-electrode patients improved significantly over time, but their scores were significantly lower and increased significantly slower than those of the controls. Their study showed that with partial insertion, performance was poorer than that with deep insertion.

Kirk, Sehgal and Miyamoto (1997) examined the speech perception performance of five children with partial insertions of the Nucleus 22-channel cochlear implant. The partial-insertion subjects' pre-implant and 1.5 years post-implant performance on a battery of speech perception tests was compared to the average performance of age-matched control subjects who received full electrode insertions. All the partial-insertion subjects were fitted with their Nucleus cochlear implants between ages of 2 and 5 years and had used their device for at least 1.5 years. In their study they found that the subjects with partial electrode insertions performed similarly to the control group at both the pre-implant and 1.5 years post-implant intervals. Furthermore, the partial-insertion subjects showed continued improvements in speech perception performance with increased device experience past 1.5 years, again similar to the full-insertion control group. This study suggested that partial insertion of a multichannel implant device is a feasible approach to auditory rehabilitation of children when full insertion is contraindicated.

The long-term outcome of seven children with partial insertion of the Nucleus electrode array was evaluated by Rotteveel, Snik, Vermeulen and Mylanus (2005). They compared the speech perception performance of the seven children with that of eighteen children with full insertion of the electrode array. Over a 3-year follow-up period, the children with partial insertion showed continuing progress, although there was wide variation in performance and the rate of progression. Some open-set comprehension was achieved with the insertion of only eight electrodes of a nucleus device. However, speech perception in the partial insertion children was poorer than that in the control groups and they showed slower progress. Four of the seven children acquired open-set word recognition. From their study it could be inferred that patients with partial insertion of the electrode array benefited from cochlear implant, although less than patients with complete insertion.

Dorman, Dankowski, McCandless and Smith (1989) investigated speech recognition of disjoint bands of low and high-frequency information. They assessed consonant recognition as a function of number of channels of stimulation by ten post lingual adult using Symbion cochlear implant. It consisted of six monopolar electrodes. In most patients, the four most apical electrodes were activated. Given the depth of insertion of the electrodes and their spacing, they assumed that the most apical electrode 1 was near the 1 kHz place, electrode 2 near the 2 kHz place, electrode 3 near the 4 kHz and electrode 4 near the 8 kHz place in the cochlea. They found that the addition of channel 2 to channel 1 produced little or no gain in the intelligibility for 9 or the 10 patients. Further, they also found that either channel 3 or channel 4, but usually not both, added greatly to the intelligibility allowed by channels 1 and 2. From their study they concluded that for most 'good' patients, consonants information was carried by one low frequency channel and one high frequency channel.

From the above studies it is evident that cochlear implant clients with a partial insertion of electrodes do perceive a considerable amount of speech. However, they do not perceive as well as those with a full insertion. Further, the way the way frequency information was provided to the electrodes of those with a partial insertion, did vary their speech perception.

Comparison of speech perception by cochlear implantees and simulated conditions

Besides evaluating the effect of spectral 'holes' on individuals using cochlear implant, it has also been evaluated on normal hearing individuals using simulated material. This has been a preferred method of study due to the ease at which the research can be conducted. It has also been found that controlling subject-related variables are a lot easier in a simulated condition.

Studies by Dorman and Loizou (1998) and Dorman, Loizou and Rainey (1997) validated the noise band simulation of cochlear implant by comparing the simulation results to performance of cochlear implant patient with the same number of channels. The results from the best performing implants listeners using CIS processor were similar to those from the normal hearing listeners for the same number of channels. It was opined that poor performing implant listeners might be using fewer effective spectral channels, or there may be other processing deficits underlying their poor performance. Also, Fu, Shannon and Wang (1998) in their study found the best cochlear implant user showed similar performance with CIS strategy in quiet and in noise as that of normal hearing listeners when listening to correspondingly, spectrally degraded speech.

Shannon et al. (2001) assessed the impact of the size and location of spectral 'holes' with six adult listeners fitted with Nucleus-22 cochlear implant using SPEAK strategy and six normal hearing adult listeners. 'Holes' in the tonotopic representation (from 1.5 to 6 mm in extent) were created by eliminating electrodes. Four conditions were created in which the spectral information that would normally have been presented to those electrodes were altered and presented. In the first condition the spectral information was dropped; in the second condition the spectral information was assigned to the electrode on the apical edge; in the third condition the spectral information was assigned to the electrode on the basal edge of the 'hole' and in the fourth condition the it was evenly splited between the apical and basal edge of the 'hole' (split). They found, in general, for both cochlear implantees and normal hearing subjects that speech recognition decreased as the 'hole' size increased and the decrease was larger for apical 'holes'. Overall there were few significant differences between the four conditions that distributed the spectral information around the 'holes' in the apical, middle and basal regions, but all caused performance to decrease significantly. They found 'holes' in the apical region were more damaging than 'holes' in the basal or middle regions for both group of subjects. Also rerouting spectral information around a 'hole' was no better than simply dropping it. Their study revealed that normal hearing listeners as a group had significantly higher overall performance as consonant and vowel recognition than cochlear implant listeners, but overall there was no significant difference between normal hearing and cochlear implant listeners once the scores were normalized to the baseline performance.

Further, Shannon et al. (2001) found consonant recognition generally decreased significantly as the size of the 'hole' increased. For normal hearing listeners, 4.5 mm 'holes' in the apical, middle and basal regions all caused performance to decrease significantly but there was no significant difference in recognition between the four conditions which distributed the spectral information from the 'hole' in different ways. For cochlear implant listeners, 4.5 mm 'holes' in the apical and middle regions significantly worsened consonant recognition, while performance was not affected in the basal region until the 'hole' was 6 mm in extent. For cochlear implant listeners, reassignment was significant, but no significant difference between any pair of reassignment conditions was found. Both normal hearing and cochlear implant listeners showed similar patterns of consonant recognition for basal and mid-cochlea 'holes', but cochlear implant listeners showed a smaller decrease in consonant recognition than normal hearing listeners for large apical 'holes'. Their study showed that normal hearing listeners as a group had significantly higher overall performance on consonant recognition than cochlear implant listeners. However, overall there was no significant difference between normal hearing and cochlear implant once scores were normalized.

Sentence recognition generally decreased significantly a the size of the 'hole' increased for both the groups. For normal hearing listeners there was no significant decrease in performance for the basal 'holes' even up to 6 mm, but the sentence recognition was significantly lower for 'holes' 4.5 mm or larger in the middle and apical cochlear regions. For cochlear implant listeners performance significantly decreased only when the basal 'holes' were 6 mm in extent and that sentence recognition was significantly lower for 'holes' 4.5 mm and larger in the middle and apical cochlear region.

Voicing and manner information were both virtually unaffected by the size of the 'hole' at all cochlear location. However, for both normal hearing and cochlear hearing listeners, information received on place of articulation decreased considerably as the 'hole' size was increased, particularly for the apical 'hole' location. This pattern of results was similar to that of observed for vowel recognition. Shannon et al. (2001) inferred that the similarity between vowel recognition and information received on consonants place of articulation demonstrated that 'holes' in the tonotopic representations primarily affected the spectral cues. Thus, the study by Shannon et al. (2001) revealed that once scores were normalized to baseline performance both the normal hearing and cochlear implant listeners showed similar performance. They also reported that better cochlear implant listeners had a pattern of performance more similar to that of the normal listeners.

The above studies indicate that simulated cochlear implant conditions can provide information similar to that seen in cochlear implantee. Hence it can be concluded that manipulating speech material to simulate perception through a cochlear implant, is a valid way to study perception through the device.

Speech recognition by normal hearing individuals with altered spectral information

Shannon et al. (1995) removed spectral information from speech and replaced the frequency specific information in broad frequency regions with a band limited noise. Thus, temporal and amplitude cues were preserved in each spectral band but the spectral details within each band was removed. All bands were then summed and presented to the normal hearing listeners. One, two, three or four band processors were used, each with envelope information low-pass filtered at 16, 50, 160 or 500 Hz for a total of 16 conditions. The surprising result was that speech was highly recognizable with only three or four bands of noise. However, from the study by Shannon et al. (1995), it was not clear whether the four frequency bands chosen by them were optimum for speech information content.

The effect of the location and size of the spectral 'holes' on vowel and consonants recognition in twenty normal hearing adults was investigated by Kasturi et al. (2002). Speech material was first low-pass filtered using a sixth order elliptical filter with a cutoff frequency of 6000 Hz. Filtered speech was passed through a pre-emphasis filter with a cutoff frequency of 2000 Hz. This was followed by band pass filtering into six different frequency bands using sixth-order Butterworth filters with center frequencies of 393, 639, 1037, 1685, 2736 and 4444 Hz. To create a 'hole' in freqeuency band, the amplitude of sinusoid corresponding to that frequency band was set to zero. With this technique they found that when a single 'hole' was introduced in the spectrum, vowel and consonant recognition decreased. The degree of performance depended on the location of the 'hole' or, equivalently, the frequency band removed. For vowels, there was a significant drop in performance when either of the frequency bands, 1, 3 and 4 centered around 393, 1037, and 1685 Hz, respectively were removed. For consonants, there was a modest, yet significant, drop in performance when either of the frequency bands 4, 5 and 6 centered around 1685, 2736, and 4444 Hz, respectively, were removed. In their study vowel recognition was affected most, with the lowest performance (60% correct) obtained when channel 3 which was centered around 1037 Hz, was removed. This was considered responsible for coding F2 information. Vowel recognition performance was dependent on the frequency location of the pairs of bands removed. In particular, removing pairs of band that contained Fl and/or F2 information caused a significant drop in performance. Consonant recognition was mildly affected by the location of the pair of frequency bands removed. Consonant recognition remained robust at 70% correct, even when the mid and high frequency speech information was missing. The manner and voicing features were not affected by the location of the 'hole' in the spectrum.

All the above mentioned studies were carried out with adult participants. A few studies have also been carried out on children. Studies on children are necessary since a large number of children are being implanted. In addition to investigating the number of spectral channels required to understand speech, studies on children also assess the developmental of speech pattern. Eisenberg et al. (2000) investigated this developmental time course of pattern recognition for speech degraded in spectral resolution. They measured performance on two groups, normally hearing children (5-7 and 10-12 years) and normal hearing adults. Age appropriate test materials were used. Spectral manipulation of speech signal was accomplished digitally using filtered noise band modulated by the speech amplitude envelope from the same spectral band (Shannon et al., 1995). For this purpose they utilized a set of 4, 6, 8, 16 and 32 noise bandpass filters. At first the acoustic signal was digitized and passed through a pre-emphasis filter. The frequency response of the pre-emphasis filter was uniform above 1200 Hz, with a -6-dB/octave roll off below 1200 Hz. The frequency spacing of the bank of filters was computed from the formula derived by Greenwood, (1990, cited in Rosen, Faulkner & Wilkinson, 1999). Assuming a total processor bandwidth of 300-6000 Hz, the distance corresponding to the difference of 5700 Hz, was divided by the number of processor channels determined for investigation. Each of the bandpass filters was implemented using a cascade of second-order elliptical infinite impulse. The output from each of filter was input to an envelope detector consisting of a full-wave rectifier followed by a low-pass filter (160 Hz). The extracted envelopes modulated a pseudorandom noise, which was input to corresponding bandpass filters with the same frequency bandwidths as those of the original filters. Then the signals were summed and converted to analog form.

Eisenberg et al. (2000) reported improvement of mean sentence scores as the number of noise bands increased up to 8 noise bands. Performance by the older children and adults were similar, with mean scores of 93% and 94%, respectively, for 8 noise bands. The younger children demonstrated lower scores than the adults and older children, recognizing only 82% of sentences correctly in the 8-band condition. Even with 32 noise bands, only two of the six younger children attained scores similar to those of adults and older children in the 8-band condition. This indicated that sentence scores generally did not improve after 8 bands for the younger children. When analyzed by the number of words correct within a sentence, the older children and adults approached an asymptote by 6 bands, with a mean word score of 96%. The younger group achieved a mean word score of 94% by 8 bands. A ceiling effect for words scores was noticed for 16 and 32 noise bands.

Further, Eisenberg et al. (2000) also studied manner, voice and place contrasts. The high performance scores achieved by the adults and older children were weighted primarily by vowels and consonants manner, with the older children and adults obtaining mean scores of 95% to 100% across the 4, 6 and 8 band conditions. They found that younger children yielded significantly lower scores than the older children and adults on the 4 and 8 band conditions. The younger children achieved mean scores of 90% or higher in the 6-band condition for vowels and 89% or higher for consonants manner across all noise-band conditions. For consonant place, the older children and adults obtained mean scores that exceeded 90% by the 6-band condition; the younger children did not reach this level of performance for any of the noise-band conditions. The consonant voicing contrast was clearly the most troublesome for all subjects. Also, it was observed that the younger children were more variable in their performance than were adults and older children. Eisenberg et al. (2000) attributed this in part to fatigue and partly due to lack of attention.

From the above studies, it is clear that even with altered spectral information speech can be recognizable to both adults and children. However, performance of younger children has been found to be slightly poorer to that of adults.

From the review of literature, it is evident that both the cochlear implantees and normal hearing individuals get a good amount of spectral information to understand speech even with altered spectral information and the presence of 'hole'. However, this speech perception performance varies depending on the cutoff frequencies and bandwidth of the filters i.e the location and size of the spectral 'hole' and the listeners' age. These studies also report that simulating cochlear implant is a good technique to study the speech perception performance in cochlear implantee.

METHOD

The study was designed to assess speech recognition of spectrums with 'holes' by children. The speech material that was developed simulated perception through a twenty-two channel cochlear implant. The 'holes' in the spectrum were created by filtering out specific frequency bands, which corresponded to the frequency bands in a Nucleus cochlear implant.

Participants

Thirty children with normal air conduction and bone conduction thresholds in frequency range of 250 Hz to 8 kHz and 250 Hz to 4 kHz respectively, participated in the study. The children were in the age range of 7-10 years. They were native speakers of Kannada and were able to read and write the language. They had no history of any neurological disorders and had normal middle ear functioning as measured by tympanometry and acoustic reflexes.

Instrumentation

A Pentium IV computer with Matlab software was used for the development of the material. A CD burner (Nero 7 Ultra Edition) was used to write the material on a CD. A calibrated two channel audiometer (Madsen OB 922) with TDH 39 earphone was used to evaluate as well as present the test items. A GSI Tympster immittance meter was used to evaluate status of middle ear and to obtain acoustic reflexes. The speech material was played through a Philips CD player.

Material development

The speech identification test material "The Kannada Phonemically Balanced words" developed by Yathiraj and Vijayalakshmi (2005) was used to simulate speech processed through a cochlear implant. The test consisted of four lists of bisyllabic words with each list having 25 words. The CD recorded version of the test was used. The below mentioned procedure was used to simulate speech processed through a twenty-two channel cochlear implant.

The words of each of the original lists were randomized to create eight lists. These speech signals were band pass filtered into twenty-two frequency bands using 6th order Butterworth filters. Crossover and center frequencies were calculated using the following equation relating the position on the basilar membrane to its characteristic frequency and assuming a basilar membrane length of 35 mm (Greenwood, 1990, cited in Rosen, Faulkner and Wilkinson, 1999):

$$\text{Frequency} = 165.4\ (10^{0.06x} - 1)$$

$$X = 1/0.06 \log (\text{Frequency}/165+1)$$

The envelope detection occurred at the output of each filter by full wave rectification and second order Butterworth low pass filter at 400 Hz. Forward and backward filtering were used to cancel the phase delays. These envelopes were then multiplied by signal correlated noise. Before being summed, the signal was passed through an output filter similar to the analysis filter (Figure 1). Two conditions were created, "no 'hole' condition and "dropped" condition. In the dropped condition, seven lists were created in which the output noise bands were simply omitted (band pass filtered with 6th order

Butterworth filter) from the processed signal. The 'hole' size was calculated corresponding to each of the dropped condition. The frequency bands of the band stop filters and 'hole' sizes are shown in table 1. These band rejections were created in the frequencies which corresponded to frequency bands of groups of adjacent electrodes in a Nucleus cochlear implant. The above procedure was carried out using a Matlab software. In the no 'hole' condition, no band pass filtering was carried out. The altered stimuli were recorded on a compact disc (CD). Prior to each test stimuli a 1 kHz calibration tone was also recorded.

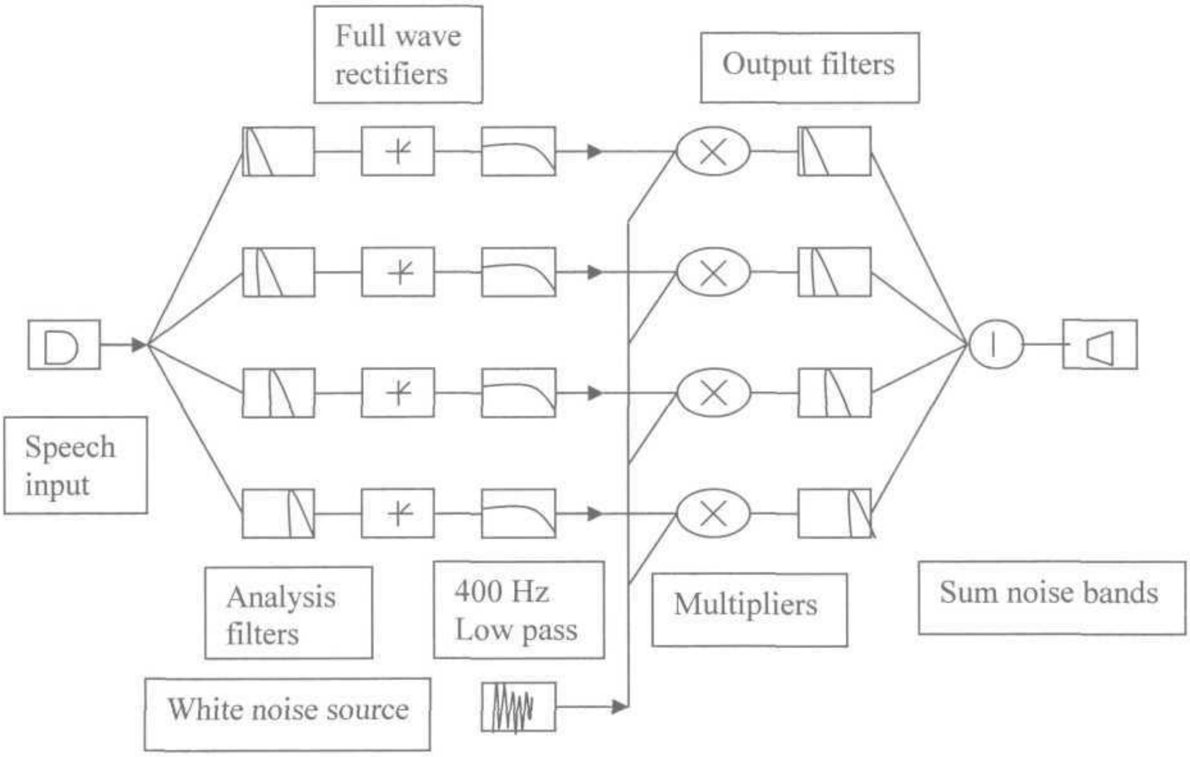

Figure 1: Block diagram of the processing used for transforming the speech signal

Table 1: Band rejection done for specific lists

LISTS	**FREQUENCY BANDS**	**HOLE SIZE**
LIST1	438-813Hz	3.4 mm
LIST 2	938-1313Hz	2.0 mm
LIST 3	1563-2313 Hz	2.6 mm
LIST 4	2688-4063 Hz	2.8 mm
LIST 1 A	4688-6938 Hz	2.8 mm
LIST 2 A	438-1063 Hz	5.1 mm
LIST 3 A	438-1313 Hz	6.4 mm
LIST 4 A	No filter	-

Test Environment

The test was carried out in a two-room audiometric set-up which was acoustically treated. The noise level was within permissible limits as recommended by ANSI (1991, cited in Wilber, 1999).

Procedure

The hearing sensitivity, of the participants was assessed using a calibrated two channel audiometer (Madsen OB 922) with TDH 39 earphone. A GSI Tympster immittance meter was used to evaluate status of middle ear and to obtain acoustic reflex. Thirty children who met the subject selection criteria participated in the study.

The developed material was played using a Philips CD player. The output from the player was routed to the tape input of the two-channel audiometer. Prior to the speech signals being presented, the recorded 1 kHz calibration tone was played. This was used to adjust the VU meter deflection to zero. The signal was presented at 40 dB HL. The output from the audiometer was heard by the participants through circumaural headphone. Half the participants heard the signal in the right ear and the other half in the left ear. The order in which the lists were presented was randomized to avoid any list order bias. The participants were instructed to write down the words they heard.

Scoring

The written responses of the participants were scored in terms of phonemes as well as words that were correctly identified. For word scoring each correct response was given a score of one and a wrong response a score of zero. The maximum possible word score for each list was 25. For phoneme scoring each correct response was scored as one and wrong as zero. The maximum possible scores for list 1/1A, 2/2A, 3/3A and 4/4A was 103, 103, 104 and 106 respectively

RESULTS AND DISCUSSION

To investigate the aims of the study, statistical analysis using SPSS software (version, 10.0) was carried out. Descriptive statistics and repeated measure ANOVA was carried out on the data obtained from the thirty children aged seven to ten years to check significance of difference of speech identification scores across different filter conditions. This was done across all the filter conditions as well as non-filter condition. Both word scores and phoneme scores were analyzed for each filter condition.

The results are discussed under the following two headings:

A) Effect of different band stop filters ('holes') on speech identification scores using
   a) Word scores
   b) Phoneme scores

B) Phoneme errors as a function of different band stop filters for
   a) Vowels
   b) Consonants

A) Effect of Different Band Stop Filters on Speech Identification Scores

The mean and standard deviation for the word scores and phoneme scores are shown in Tables 2. This information is provided for all eight lists that were evaluated. Details of the vowel and phoneme scores are discussed below.

Table 2: Mean and SD for the word and phoneme scores for different filter conditions.

Lists	Band Rejections (Hz)	Word Score		Phoneme Score	
		Mean scores	Standard deviation of raw scores	Mean scores	Standard deviation of raw scores
List I	438-813	49.08% (12.27)	5.56	80.30% (82.73)	9.01
List II	938-1313	52.52% (13.13)	6.13	84.36% (86.90)	8.28
List III	1563-2313	56.28% (14.07)	6.02	83.46% (86.80)	10.00
List IV	2688-4063	54.00% (13.50)	7.16	81.00% (85.86)	11.51
List IA	4688-6938	57.48% (14.37)	5.85	85.95% (88.53)	8.53
List IIA	438-1063	58.28% (14.57)	5.67	87.01% (89.63)	7.62
List IIIA	438-1313	52.00% (13.00)	6.56	83.71% (87.06)	9.93
List IVA	No Filter	57.20% (14.30)	6.34	85.08% (90.13)	9.03

Note: Maximum word score = 25.
Maximum phoneme scores vary from 103 to 106.
Values given in bracket refer to the raw score.

a) Word scores across filter conditions

The mean word scores were relatively low across all the filter conditions, including the unfiltered condition. This indicates that the material simulating speech processed through a twenty-two channel cochlear implant, resulted in reduced word scores, without and with band stop filters. This highlights that the simulated material, resulted in distortion, reducing the overall intelligibility of the developed material.

A repeated measure ANOVA revealed that there was a significant difference between the word scores across the different lists [F (7, 203) =3.957, p < .05]. The Bonferroni multiple comparison test further indicated that for the word scores there was a significant difference between List I and List IIIA; List IV and List IVA at the level of 0.05 level of significance. Surprisingly, though the 'hole' was larger for List IIIA, the performance was poorer for List I. Though both the lists were equal in terms of phonemic balance and difficulty, the variations in test items could have lead to the difference in performance. The participants may have been able to utilize coarticulated information to a greater extent in List IIIA than in List I.

A significant difference was also found between List IV and List IVA. List IVA was without any 'hole', whereas List IV had a 'hole' of 2.8 mm. The cutoff frequencies used to create List IV were in the frequencies region of the second format (F2) for several of speech sounds and F2 is a major cue to differentiate vowels (Carlson, Fant & Grantson, 1975). Hence, removal of F2 information in List IV might have lead to a significant difference to that of the no 'hole' condition. However, no significant differences were found for word scores across all other lists. Based on this finding it can be inferred that generally the listeners were able to combine the information across various frequency bands to perceive a "whole" signal.

b) Phoneme score across filter conditions

The repeated measure ANOVA also revealed that there was a statistically significant difference between the phoneme scores across the eight lists [F (7, 203) = 7.863, p < .05]. This significant difference was observed between List I and Lists IA, III, IIIA, IV and IVA. In addition List IV was significantly differed from List IVA. The difference in phoneme scores can be attributed to increase in 'hole' size. The 'hole' size for List I was larger (3.44 mm), when compared to List IA (2.8 mm), List III (2.6 mm) and List IV (2.8 mm).

The finding of the present study is in agreement with the findings of Shannon et al. (2001). They too reported that speech recognition decreased as the 'hole' size increased in normal hearing adults. They reported that a 4.5 mm 'hole' caused the performance to decrease significantly for consonants and vowel recognition. However, in the present study, it was found that a 'hole' size of more than 2.6 mm was able to reduce speech recognition. This difference of findings in both the studies might be due to test procedure and age of the participants. Shannon et al. (2001) carried out the experiment on adults in a free field condition while, the present study was carried out on children under head phones. The task in the study by Shannon et al. (2001) was to identify medial vowels and medial consonants. However, in the present study, consonants and vowels in the initial, medial and final position had to be identified. This might have also attributed to difference in findings.

Shannon et al. (2001) also found that decrease in speech recognition was larger for apical 'holes' than basal 'holes'. In the present study also the 'holes' representing the apical region of the cochlea resulted in the poorer scores (List I) when compared to the 'holes' representing the more basal regions. However, this was not seen for all 'holes' representing the basal region. No significant difference was found among other lists. It showed that listeners were able to effectively combine information from different frequency regions and perceive the speech signal despite the removal of certain frequency components.

In general, higher phoneme scores were obtained for phonemes in comparison to word scores. Similar results were also, found by Olsen, Van Tasell and Speak (1997) in a group of normal hearing adults. In their study, phoneme scoring yielded scores that were on the order of 20% higher than scores for whole words heard. Barick (2006) also found a significant difference in word and phoneme scores. He recommended that word scores be calculated rather than phoneme scores, since this scoring procedure depicts the perceptual problem better. However, he suggested that if the client was to be referred for auditory listening training, the phoneme scoring procedure should be used.

Phoneme scores as a function of age

Phoneme scores across filter conditions were also evaluated as a function of age of the participants. All the participants were grouped into four age group, 7-8 years, 8-9 years, 9-10 and 10-11 years. ANOVA results indicated that there was a significant difference between the three age groups [F (3, 22 = 9.173, $p < .05$]. The Duncan post hoc test revealed that the youngest group (7-8 year olds) performed significantly poorer than the rest of the groups.

The finding regarding the performance of different age groups is similar to that reported by Eisenberg et al. (2000). They too noted that the youngest group in their study (5-7 year olds) performed significantly poorer than their older children aged 10-12 years, on a speech perception task. They had used age appropriate test material and found this finding on a variety of age matched tasks. They reported in their study that, the younger children were more variable in their performance than the adults and older children, suggesting probable cognitive or task related factors playing a role.

Thus, the findings of the present study substantiate the presence of a developmental trend in the perception of spectral 'holes'. This indicates that unlike the older children, younger children are unable to carry out an auditory closure activity and guess the material that has been presented them.

B) Phoneme Errors as a Function of Different Band Stop Filters

In addition to obtaining word and phoneme scores, confusion of vowels and consonants were observed in each lists. Further, an error analysis was carried out for both the vowels and consonants to assess the error pattern. The analysis was done for three lists (List I, List III, and List IA). These three lists were analyzed as they represented low (438-813 Hz), mid (1563-2313 Hz) and high (4688-6938) frequency band-stop filters respectively. Overall less consonantal errors were noticed than vowel errors (Table 3). Probably due to the larger number of redundant segmental cues present in consonants, the participants were able to guess them despite the presence of 'holes' in the spectrum. Thus, if one cue is missed due to filtering, listeners can perceive the other cues and identify the consonants. Despites vowels being more robust, the number of redundant segmental cues present in them are less.

t

Table 3: Consonants and vowel errors across different filter condition

List	**Consonant errors (in** %)	**Vowel errors (in** %)
List!	40.0	60.0
List II	44.0	56.0
List III	41.1	58.8
List IV	42.1	58.8
List IA	40.6	60.0
List IIA	46.6	53.3
List IIIA	35.3	64.7
List IVA	46.6	53.3

Vowel Errors

Among the three lists, the least errors were observed in List IA, followed by list III and List I. In List IA, the cutoff frequency was between 4688-6938 Hz whereas the cues for vowel perception lie between 270 Hz and 2160 Hz (Peterson & Barney, 1954). Thus, in List IA, all the information for the perception of vowels was preserved, hence the spectral 'hole' in it did not adversely affect the perception of vowel. In contrast, the errors were maximum in List I as it removed the low frequency component in which first formats for all the vowels and second formants of many vowels lies (Peterson & Barney, 1954)

Further, the error analysis of these three lists revealed that there was confusion between short and long duration vowels. This was observed across all the three lists. This highlights that the simulated speech material affected the temporal cues required for the perception of long versus short vowels, resulting in this confusion. The error analyses for vowels also highlighted that in List I, which contained the low frequency spectral 'hole', /u/ was confused with *I'll* and /o/ was confused with /a/. Elimination of essential format information probably resulted in this confusion.

Consonant Errors

The consonantal error analysis was carried out in the same three lists as that done for vowels. In all the three lists voicing, place of articulation and manner was noticed to be affected. However, the majority of errors were observed for place of articulation followed by and manner and voicing errors. Manner and voicing cues have been found to be primarily temporal cues (Van Tassel, Soli, Kirby & Widin, 1987) and they require minimal spectral cues to be accurately perceived (Shannon et al., 1995). In contrast place cues require spectral cues which are affected by the spectral 'holes'. Similar results were also obtained by Shannon et al. (2001). They too reported that information received on place of articulation decreased considerably as the 'hole' size was increased, particularly when the 'hole' was located apically. In the present study maximum errors were obtained in List I which had a 'hole' located more apically. The 'hole' size in the List I was 3.4 which was larger than the other two lists. This apical location and larger size might have lead to more errors in List I. Also, the alveolar */d/* was confused with labial /b/ and velar /k/ in List I and in List IA respectively. Generally the major cues for the perception of place of articulation of stops are the bursts and second format transition (Cooper, Delattre, Liberman, Borst & Gerstman, 1952). The spectal 'holes' in the speech material probably eliminated some of the major cues, causing confusion in the place of articulation perception.

The various general findings obtained from the present study are:

- With increase in 'hole' size, there was a deterioration in speech recognition scores.
- The apical location of 'holes' or band-stop filters, affected speech perception more than the basal 'holes' or band-stop filters. However, not all the basal 'hole' had the similar had similar adverse affect.
- The phoneme scores were higher in comparison to word scores.
- The error analysis indicated that consonantal perception was better than vowels for all filter condition.
- Among vowels, maximum confusion was noticed among short versus long.
- For consonants more place of articulation confusion than manner and voicing confusion.

SUMMARY AND CONCLUSION

Shannon et al. (2001) have shown that a simulated cochlear implant conditions can provide information similar to that seen in cochlear implantee. Studies carried out by various other investigators on normal hearing individuals (Kasturi et al.,2002; Shannon et al., 1995) show that even with altered spectral information speech is recognizable. All these studies advocate using a simulated cochlear implant procedure to study speech perception through the device. Most of these studies have been carried out on adult listeners. However, the number of studies done on young children is relatively less. There is less information available regarding how children behave under conditions of spectrally degraded speech. Eisenberg et al. (2000) reported that performance of younger children had been found to be slightly poorer than that of older children and adults using cochlear implants. The participants in their study were of 5-7 year and 10-12 year old.

The present investigation was taken up to study speech perception in children using a cochlear implant simulation condition. The following were studied:

- Speech recognition of spectrums with "holes" by children,
- The effect of different band stop filters in speech perception in children and
- The vowel and consonantal errors due to spectral 'hole' in the low, mid and high frequency regions.

To carry out the study thirty normal hearing children (age range of 7-10 years) were taken as participants. The speech identification test material "The Kannada Phonemically Balanced words" developed by Yathiraj and Vijayalakshmi (2005) was processed to simulate cochlear implant using Greenwood (1990, cited in Rosen, Faulkner & Wilkinson, 1999) equation. Seven lists were created to provide 'hole' in the lists using band-stop filters which corresponded to frequency bands of groups of adjacent electrodes in a Nucleus cochlear implant. Stimuli were presented and the written responses of the participants were scored in terms of phonemes as well as words that were correctly identified.

It was found that there were less consonant errors than vowel errors for both word scoring and phoneme scoring. This error pattern was more with larger 'hole' size. However, this trend was not followed in each of the lists suggesting that listeners were able to combine the information from other frequencies to perceive the whole signal. It was also noticed that 7-8 year olds children showed significantly poorer performance than older children.

Implication

The present study will add to the current knowledge pool of understanding speech pattern recognition in young cochlear implantees and their perceptual differences in the speech recognition with adults.

Clinically, findings of this study may help in predicting speech perception as a function of the electrodes that are switched on.

The study will help to know the effect of band-stop filter and relative importance of various frequency bands in speech perception.

Information from the study would be useful in counselling parents of young cochlear implantees or cochlear implantees regarding the speech sounds that will be affected if specific electrodes would be switched off.

REFERENCES

Barick. S. K. (2006). High frequency-english speech identification test. Unpublished Master's dissertation submitted as part fulfillment for the degree of Masters of Science, to the University of Mysore, Mysore.

Baskent, D. & Shannon, R. V. (2004). Interactions between cochlear implant electrode insertion depth and frequency-place mapping. *Journal of Acoustical Society of America,* 77, 90-119. 117(3), Pt. 1, 1405-1416

Bredberg, G., Lindstrom, B., Baumgartner, WD., Farhadi, M., Goldberg, T., Gstottner, W., Pillsbury, H, Skarzynski, H., Sorri, M., Van de Heyning, P., Zaghis, A.. Graham. J., Williams, G. & D'haese, P. (2006). Open-set speech perception in adult cochlear implant users with ossified cochleae. Cochlear Implant International, 4, 2, 55-72

Carlson R., Fant G., & Grantson B. (1975). Auditory Analysis and Perception of Speech.

Cooper, F. S, Delattre. P. C, Liberman, A. M., Borst, J. M. & Gerstman. L. J. (1952). Some experiments on the Perception of Synthetic Speech Sounds. *Journal of Acoustical Society of America,* 24, 6, 597-606.

Dorman, M. F., Loizou, P. C, & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *Journal of Acoustical Society of America,* 102, 2403-2411.

Dorman, M. F., & Loizou, P. C. (1998). The identification of consonants and vowels by cochlear implant patients using a 6-channel continuous interleaved sampling processor and by normally hearing subjects using simulations of processors with two to nine channels. Ear and Hearing, 19, 162-166.

Dorman, M. F., Loizou, P. C, Fitzke, J., & Tu, Z. (1998). The recognition of sentences in noise by normally hearing listeners using simulations of cochlear-implant signal processing. *Journal of Acoustical Society of America,* 104, 3583-3585.

Dorman M. F, Dankowski K, McCandless G & Smith L. (1989). Consonant Recognition as Function of the Number of Channels of Stimulation by Patients Who Use the Symbion Cochlear Implant. Ear and Hearing, 10, 5, 288-291

Eisenberg L, Shannon, R. V., Martnez, A. S., Wygonski, J. & Boohroyd, A. (2000). Speech perception with reduced specral cues as a function of age, *Journal of Acoustical Society of America,* 107, 5, 2704-2710

Fishman, K. E., Shannon, R. V. & Slattery, W. H. (1997). Speech Lang, recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," Journal of Speech Language and Hearing, 40, 1201-1215.

French, N. R. & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sound. *Journal of Acoustical Society of America, 11,* 90-119.

Friesen, L. M., Shannon, R. V., Bas.kent, D. & Wang X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal of Acoustical Society of America,* 110, 1150-1163.

Fu, Q. J., Shannon, R. V. & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and Electric hearing. *Journal of Acoustical society of America,* 104, 3586-3596.

Fu QJ, Zeng FG, Shannon RV & Soli SD. (1998). Importance of tonal envelope cues in Chinese speech recognition. *Journal of Acoustical Society of America,* 104: 505-510

Holmes A. E., Kemker F. J. & Mervin GE. (1987). The effects of varying the number of cochlear implant electrodes on speech perception. *American Journal of Otology,* 8: 240-246

Kasturi, K., Loizou, P., Dorman, M. & Spahr, T. (2002). The intelligibility of speech with "holes" in the spectrum. *Journal of Acoustical society of America*, 112 (3), 1102-1111.

Kirk K. I., Sehgal M. & Miyamoto R. T. (1997). Speech Perception performances of nucleus cochlear implant users with partial electrode insertions, *Ear and Hearing* 18,6,456-71

Kryter, K. (1962). Validation of the articulation index. *Journal of Acoustical society of America.* 34, 1698-1702.

Moller, A. R. (2000). Hearing: Its Physiology and Pathology, (pp. 74) San Diego, Academic Press.

Olsen, W. O., Van Tassel, D. J. & Speaks C. E. (1997). Phoneme and Word Recognition for Words in Isolation and in Sentences, *Ear and Hearing,* 18, 3, 175-188

Peterson G. E. & Barney H.L. (1954) "Control methods used in a study of the identification of vowels. *Journal of Acoustical society of America,* 24, 183-

Pollack, I. (1948). Effects of high-pass and low-pass filtering on the intelligibility of speech in noise. *Journal of Acoustical society of America,* 20, 259-266.

Remez, R., Rubin, P., Pisoni, D., & Carrell, T. (1981). Speech perception without traditional cues, *Science,* 212, 947-950.

Rotteveel, L.J.C. Snik, AD.F.M. Vermeulen, A.M. & Mylanus, E.A.M. (2005). Three-year follow-up of children with postmeningitic deafness and partial cochlear implant insertion. *Clinical Otolaryngol,* 30, 242-248

Rosen, S., Faulkner, A., & Wilkinson, L. (1999). Adaptaion by normal listeners to upward spectral shifts of speech: Implications for cochlear implant. *Journal of Acoustical Society of America,* 106, 6 3629-3636

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

Shannon, R. V., Galvin, J.J., & Baskent, D., (2001): Holes in hearing, *Journal of the Association for Research in Otolaryngology,* 185-199.

Stickney, G., & Assmann, P. (2001). Acoustic and linguistic factors in the perception on bandpass-filtered speech, *Journal of Acoustical society of America,* 109, 1157—1165.

Turner, C. W., Souza, P. E., & Forget, L. N. (1995). Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *Journal of Acoustical society of America,* 97, 2568—2576.

Van den Honert, C. & Stypulkowski, P. H. (1984). Physiological propperties of the electrically stimulated auditory nerve. II. Single fiber recordings. *Hearing Research,* 14,225-243.

Van den Honert, C. & Stypulkowski, P. H. (1987). Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve. *Hearing Research.* 29, 195-206.

Van Tasell, D. J., Soli, S. D., Kirby, V. M. & Widin, G. P. (1987). Speech waveform envelope cues for consonant recognition. *Journal of Acoustical society of America*, 82, 1152-1161

Van Tasell, D. J., Greenfield, D. G., Logemann, J. J. & Nelson, D. A. (1992). Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of cochlear implants. *Journal of Acoustical society of America*, 92, 1247-1257.

Vidya, M, Rima, D. & Yathiraj A. (2005). Speech Identification of a Spectrum with Holes. Presented at ISHACON, 2006 held at Ahmedabad.

Von Bekesy (1960): Cited in Hearing: Its Physiology and Pathology; Aage R. Moller, 2000

Warren, R. M., Riener, K. R., Bashford, Jr., J. A., & Brubaker, B. S. (1995). Spectral redundancy: Intelligibility of sentences heard through narrow spectral slits. Percept. Psychophysics. 57, 175-182.

Wilber, L. A. (1994). Calibration, pure tone, speech & noise signal. In J. Katz, *Handbook of Clinical Audiology.* (pp. 73-96). Baltimore: Willliams & Wilkins.

Yathiraj, A. & Vijayalakshmi, C. S. (2005) The Kannada Phonemically Balanced Word Test developed in the department of Audiology, All India Institute of Speech and Hearing.