Int. Adv. Otol. 2011; 7:(3) 379-384

ORIGINAL STUDY

Perception of Spectrally Shifted Speech: Implications for Cochlear Implants

Pitchai Muthu Arivudai Nambi, Subramaniam Manoharan, Jayashree Sunil Bhat, Sangamanatha Ankmnal Veeranna

Department of Audiology and Speech Language Pathology Kasturba Medical College (A Unit of MANIPAL University) Mangalore, Karnataka, India (AN, JSB)
Consultant Audiologist Madras Audiology and speech therapy center Chennai (SM)

Objective: Present study used the normal hearing subjects to check the effect of spectral shifts in monoaural and binaural cochlear implant simulations. Cochlear implant signal processing was simulated using eight channel sine wave vocoders.

Materials and methods: The spectrally shifted sentences (HINT) were created to simulate 20 mm & 25 mm insertion depth of cochlear implants. Speech recognition scores were assessed in three mono aural conditions (No shift, 25 mm & 20 mm insertion) and two binaural conditions [asymmetrical spectral shift (20 & 25 mm insertions in each ear) & no spectral shift in both ears]. Greenwood (1990) function was used to decide the corner frequencies for the simulation of spectral shifts.

Results: It reveled that partial shift (20 mm) resulted in poor scores in monaural than 25 mm shift. Spectrally unshifted speech resulted in significantly higher scores when compared to spectrally shifted speech. In binaural asymmetrical condition, scores were equivalent to 25 mm shift in monaural condition. Even in symmetrical condition also binaural scores were equivalent to mono aural scores which implies that at least in initial stages there may not be binaural integration or interference may be present.

Conclusion: In binaural cochlear implantation there may not be any change in the speech perception in the initial stages even if the insertion depth varies both ears, at least in quiet condition.

Submitted : 29 December 2010 *Revised: 14 July 2011* *Accepted : 31 August 2011*

Introduction

Cochlear implant is device which converts acoustic stimuli to electrical stimulation. This implantable device helps hearing impaired to understand speech by making use of important cues in the speech signal [1]. The cues for speech recognition can be broadly classified as temporal envelope and temporal fine structure cues. The contribution of spectral, temporal envelope and temporal fine structure cues for speech recognition has been studied extensively. In 1995, Shannon, Zeng Kamath, Wygonski and Ekelid [2] reported that envelope cues from as few as four bands are sufficient for good speech recognition in quiet. Increasing the number of spectral channels improves the speech perception. Adding fine structure cues along with envelope significantly improves the speech perception in adverse listening conditions [3]. Currently, cochlear implants code essentially the spectral and temporal (envelope) information of the speech signals. The spectral information (tonotopic organization) is coded by number of stimulating electrodes and electrode position in the cochlea. Temporal information is coded by presenting the band specific envelope to each corresponding electrodes [4]. But in the cochlear implants the spectral information is presented to wrong place of the auditory nerve array, due to the fact that electrodes can only be inserted partially into the cochlea. This partial insertion causes spectral shift in the speech signal carried by the auditory nerves. The effect of this insertion depth is

Corresponding address:
Ankmnal Veersanna Sangamanatha
Lecturer Department of Audiology and Speech Language Pathology
Kasturba Medical College (A Unit of MANIPAL University) Mangalore, Karnataka, India
Phone: +91-9880777500
E-mail: sangamanatha@gmail.com

difficult to be studied in cochlear implantees due to the interaction of following variables such as duration of deafness, age of implantation, insertion depth across individuals, cognitive and linguistic performance and the amount of neuronal survival. So, efforts have been made to simulate the effect of insertion depth by shifting the spectrum towards high frequencies. There is a significant decrement in the performance when the spectrum is shifted towards the high frequencies to mimic the basal positioning of the electrodes in cochlea [5, 6]. Dorman, Louizou, Fitzke & Tu Z [7] have simulated 22-25 mm insertion depth of a six channel cochlear implants. Similarly Shannon, Zeng & Wygonski [5] have simulated 26 mm insertion depth using 4 channel noise wave processors. The average normal insertion depth is 25 mm and shallow insertion depth of 20 mm is also common [8]. Most of the cochlear implants use more than 4 or 5 electrodes, the speech perception performance increases with increase in the number of electrodes and reaches saturation beyond 8 electrodes [9].

Recently bilateral cochlear implants have also gained popularity. The insertion depth is also likely to vary in both ears, which will produce asymmetrical spectral shift in both ears. This may also interact with varying patterns of nerve survival in the two ears. To what extent are listeners able to adapt when presented with frequency-place maps that differ between two ears is still not known. Dorman and Dahlstrom [10] reported binaural advantage for speech recognition in two bilateral cochlear implant patients who had different implant in both ears, which may support the hypothesis that mismatch between the frequency to place map can be combined. Tested patients showed improvement of 32% - 34% on HINT sentence with addition of the second implant over performance of the better ear alone. However the study included only two subjects and the method of determining mismatch between the ears and pitch ranking of electrodes was not conclusive. Evidence from dichotic listening studies shows that information presented in compliment across the ears can be integrated easily [11]. This process, later termed as spectral fusion by Cutting [12] is robust to differences in level and fundamental frequency, but not relative onset time; the majority of listeners appear able to integrate (tonotopically matched) acoustic cues efficiently [13]. Binaural advantage in cochlear implant can be expected if there is similar amount of shift in both ears and if the listener able to adapt to different degrees of shift in each ear, then they might still be able to combine spectral cues from both the ears. Hence, the effect of bilateral asymmetrical spectral shifts on speech perception needs to be studied especially when there is a shallow insertion on one side. So, current study aimed to investigate sentence recognition scores in spectrally shifted and unshifted condition(i.e. spectrally shifted conditions are produced to simulate 20 mm and 25 mm insertion depth) and also to investigate the effect of asymmetrical spectral shift (one ear 20 mm and another 25 mm) on sentence recognition scores.

Method

Subject

Twelve normal hearing subjects, in the age range of 18 to 25 years, participated in the present study. Subjects whose hearing thresholds were better than 15dBHL at audiometric test frequencies from 250 to 8,000Hz and who were exposed to English language at least for 5 years were selected.

Speech material

Three sentence lists of each containing 10 English sentences with 30 key words were taken from HINT sentences [14]. The speech stimuli were spoken by female speaker with Indian English accent. The stimuli were recorded digitally on a data acquisition system at 44.1 kHz sampling frequency and using a 16-bit A/D converter in a sound treated room. The recorded speech material was scaled for the same intensity.

Signal processing

Eight channel sine wave vocoders were used to simulate cochlear implant speech processing. Speech signals were band pass filtered into 8 frequency bands with a slope of 24dB/octave. The temporal envelopes were extracted from each sub band by full wave rectification and low pass filtering at 400 Hz with a slope of 24 dB/octave. The envelope was used to modulate a fixed sine wave carrier. In spectrally unshifted condition the envelope is extracted from 200 -7000 Hz at 24dB/octave and modulated on a same carrier frequencies that is 200 – 7,000 Hz at 24 dB/octave. For spectral shift conditions the envelope is extracted from 200 - 7,000 Hz at 24 dB/octave and modulated on a different carrier. For 25mm insertion conditions 512 to 5,084 Hz at 6 dB/octave served as a

carrier and for 20mm condition 1168 to 10290 Hz frequencies were served as carriers (see Fig 1). These corner frequencies were calculated based on Greenwood's map [15]. Following is the greenwood's equation which describes the relationship between cochlear place and its corresponding frequency.

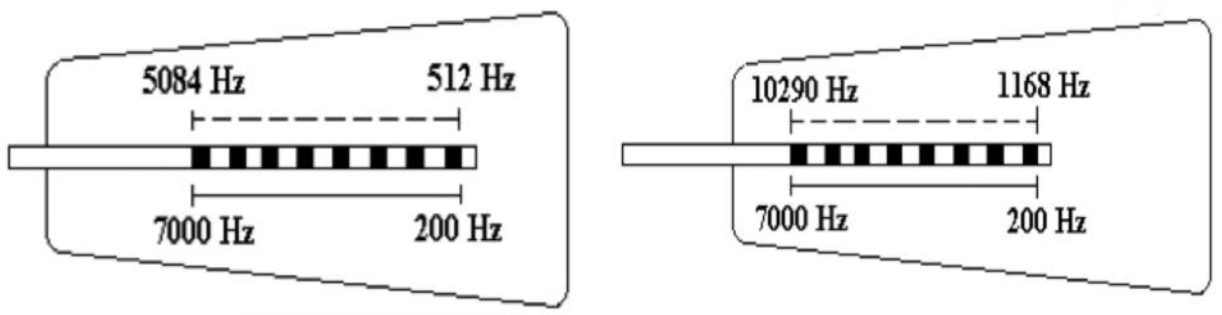

Figure 1. Schematic diagram representing spectral shift that occur in cochlear implants. Continuous line represents analysis band from which envelope is extracted. Dotted line represents the place of stimulation for 25 mm insertion (left) and 20 mm (right) insertion.

$f = 165.4 \times 10^{dx2.1} - 1$

In the equation "f" is frequency in Hertz and "d" is cochlear distance in millimeter. Distance is measured from apex to base.

Procedure

The processed stimuli were recorded on a CD and presented through stereo headphones, which was routed through MA.53 diagnostic two-channel audiometer. For monoaural condition [(i)No shift (ii) 25 mm insertion and (iii) 20 mm insertion] the stimuli were presented through one headphone for and binaural condition [(i) asymmetrical shift, 25 mm insertion in one ear & 20 mm insertion in another ear (ii) No shift in both ears] the stimuli were presented through two headphones dichotically at 0 msec delay between two ears. Written responses were taken from the subject on open set task. Only the keywords were scored and converted into rationalized arcsine units.

Result

Sentence recognition was assessed using loose method [8] in which only keywords were scored. Maximum possible in score in condition was 30 as 10 sentences were used for each condition and each sentence has three keywords. Prior to conducting statistical analysis the open set scores were transformed using rationalized arcsine transformation to account for the critical differences that are inherent in conventional scoring methods [16]. Rationalized arcsine units (RAU) were preferred rather than percentages as RAUs even account for floor and ceiling effects [16]. However there was no floor or ceiling effect found in any of the conditions in the present study. Repeated measure of ANOVA was done to check the effect of different simulation condition on speech perception. It showed significant main effect ($F(4, 28) = 56$, $p < 0.05$) of spectral shift between various spectral shift condition (No shift, 25 mm insertion, 20 mm insertion, binaural asymmetrical 25 & 20 mm insertion and no shift in both ears). Bonferroni's post hoc pair wise comparison showed significant difference ($p < 0.01$) between all the conditions except between monaural 25mm insertion and binaural 25 mm & 20 mm insertion ($p > 0.01$) and between monoaural no shift & binaural no shift condition ($p > 0.01$).

Figure 2 shows the mean performance of sentence recognitions scores in terms of rationalized arcsine units. From the Figure 2 it is clear that unshifted condition yields highest score, 20mm insertion condition yields lowest score. Monaural 25 mm insertion and binaural 25 mm & 20 mm insertion conditions produced similar scores.

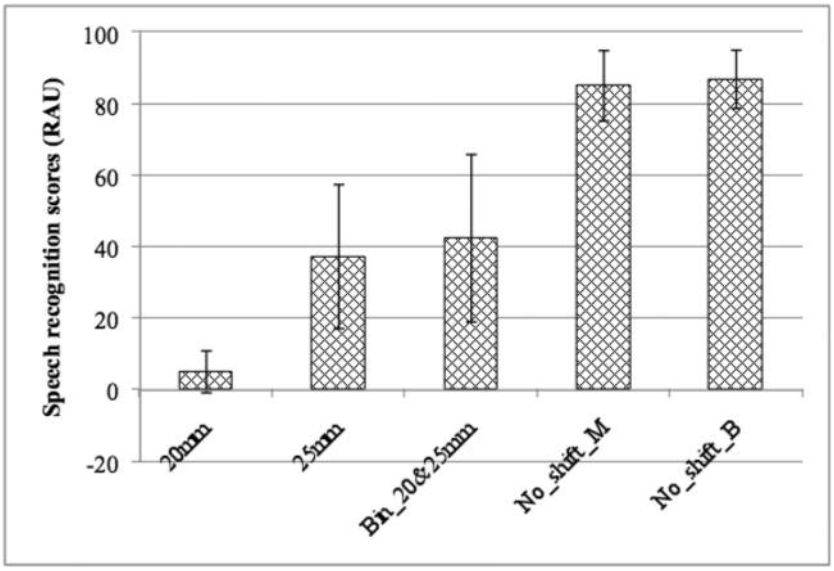

Figure 2. Mean and standard deviation values of speech recognition scores (arcsine transformed) obtained for different testing conditions

Correlation analysis was performed to investigate whether speech recognition scores under spectrally shifted condition can predict speech recognition scores under spectrally unshifted condition. It is important to study their relationship as previous investigations have shown that training improves the speech recognition scores of spectrally shifted speech to the level of unshifted speech [17]. So, predictability of speech recognition scores of spectrally unshifted speech from spectrally shifted speech would facilitate the training process. However Pearson's correlation analysis revealed no significant correlation between any of the condition (no shift and 25mm insertion, $r=0.34$, $p>0.05$; no shift and 20mm insertion, $r=0.65$, $p>0.05$).

Discussion

This study evaluated the effect of spectral distortion on speech recognition scores in different simulated conditions. Results of the current study indicate that envelope cues from 8 bands can yield high level of performance, if there is no spectral shift. Rosen, Faulkner & Wilkinson [8] reported that speech material such as sentences and vowels requires effective transmission of spectral information for good performance when compared to consonants. In our study, since the spectral information is effectively transmitted in no shift condition, the sentence recognition scores were found to be significantly high. The spectral information is not effectively transmitted in 25 mm and 20 mm insertion conditions. So there was significant decrement in the performance was observed. The spectral degradation is much higher in 20 mm condition when compared to 25 mm condition which resulted in worse performance in 20 mm condition when compared to 25 mm condition. Before transforming these results of simulated study with normal hearing to actual cochlear implant one should be cautious about confounding factors such as difference in acoustic and electric hearing. In electric hearing, parameters such as exact stimulation of the electrode array is not known or the length of the cochlea that are not known precisely unless imaging techniques such as CT scans are used [18]. Other major factors which can contribute to speech perception are the proximity of the electrodes to the modiolus and the surviving neurons near the electrode [19, 20]. In present study we used Greenwood equations, which was used to determine the acoustic input frequencies in the experimental condition, might not be applicable to electrical stimulation of the auditory nerve because it was originally formulated for healthy cochlea. Therefore findings of the stimulation studies should be considered as a preliminary data and be cautiously interpreted for real-life applications with implant users.

In binaural asymmetrical condition the sentence recognition scores were similar to 25 mm insertion condition and in binaural symmetrical condition the

scores were similar to monoaural no shift condition, which implies that there is no binaural integration and binaural interference is present at least in the initial stages. The listener attends only to the better ear. Several studies have shown synergetic improvement in speech recognition when using both the implants [10,21]. Yet not all bilateral cochlear implant users have shown binaural advantage [22, 23] over their best ear alone and substantially mismatched frequency place map underpin this lack of advantage. Benefits of binaural hearing are realized in noisy condition, where differences in the sound signal at each ear can be used to obtain a better representation of what has been said. Dorman and Dahlstrom [10] have shown binaural benefit over better ear benefit in noisy condition. Current study revealed that there is no one to one relationship between speech recognition scores for spectrally shifted and unshifted speech. It may not be possible to predict speech recognition scores of spectrally unshifted from spectrally shifted speech.

Conclusion

Current study investigated the effect of spectral shift speech recognition scores in monoaural and binaural asymmetrical conditions. There was a decrement in speech recognition scores when spectral shift is present. Thus, partial insertion may result in poor performance at least in the initial stage, when there is no training is given. Binaural asymmetrical spectral shift condition yielded similar scores that of 25 mm insertion simulation which implies that, in binaural cochlear implants there may not be improvement or decrement in the performance even if the insertion depth varies in both ears at least in the quiet condition in initial stages.

References

1. Hong JC, Yu TH, Lee KD, Heo SD, & Kang, MK. Changes in the threshold of the electrically evoked compound action potential with waveforms analysis after cochlear implant. Int. Adv. Otol. 2010; 6; 347-352.

2. Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Sci.1995; 270; 303–04.

3. Nie KB, Stickney GS, & Zeng FG. Encoding frequency modulation to improve cochlear implant performance in noise. IEEE Trans Biomed Eng. 2005; 52:64-73

4. Nie K, Barco A & Zeng FG. Spectral and temporal cues in cochlear implant speech perception. Ear Hear. 2006; 27:208-17.

5. Shannon RV, Zeng FG & Wygonski J. Speech recognition with altered spectral distribution of envelope cues. J Acoust Soc Am. 1998; 104: 2467-76.

6. Dorman MF, Loizou PC, & Rainey D. Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. J Acoust Soc Am. 1997; 102:2403-11.

7. Dorman MF, Loizou PC, Fitzke J, & Tu Z. The recognition of sentences in noise by normal hearing listeners using simulations of cochlear implant signal processor with 6-20 channels. J Acoust Soc Am. 1998; 104: 3583- 85

8. Rosen S, Faulkner A, & Wilkinson L. Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. J Acoust Soc Am. 1999; 106:3629-36.

9. Lu T, Carroll J, & Zeng FG. On acoustic simulations of cochlear implants, in Conference on Implantable Auditory Prostheses; Lake Tahoe, CA. 2007.

10. Dorman MF, & Dahlstrom L. Speech understanding by cochlear-implant patients with different left- and right-ear electrode arrays, Ear Hear. 2004; 25:191–94.

11. Broadbent DE & Ladefoged P. On the fusion of sounds reaching different sense organs, J Acoust Soc Am. 1957; 29:708–10.

12. Cutting JE. Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening, Psychol Rev. 1976; 83:114–40

13. Rubin BA, Uchanski RM, & Braida LD. Integration of acoustic cues for consonants across frequency bands. J Acoust Soc Am. 1992; 91: 2473–74.

14. Nilsson M, Soli S, & Sullivan J. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. The J Acoust Soc Am. 1994; 95:1085–99.

15. Greenwood, DD. A cochlear frequency-position functions for several species–29 years later. J. Acoust. Soc. Am. 1990; 87:2592–95.

16. Studebaker GA. A "rationalized" arcsine transform. J. Speech Lang. Hear. Res. 1985. 28, 455-62.

17. Stacey PC, Summerfield AQ. Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. J. Speech Lang. Hear. Res. 2008; 51: 526-28.

18. Ketten DR, Skinner MW, Wang G, Vannier MW, Gates GA, & Neely JG et al. In vivo measures of cochlear length and insertion depth of Nucleus cochlear implant electrode arrays, Ann Otol Rhinol Laryngol. 1998; 107:1–16.

19. Saunders E, Cohen L, Aschendorff A, Shapiro W, Knight M., Stecker M. et al. Threshold, comfortable level and impedance changes as a function of electrode-modiolar distance. Ear Hear, 2002; 23: 28S–40S.

20. Khan AM, Whiten DM, Nadol JB, & Eddington DK. Histopathology of human cochlear implants: Correlation of psychophysical and anatomical measures. Hear Res. 2005; 205:89–93.

21. Litovsky R, Parkinson A, Arcaroli J, & Sammeth C. Simultaneous bilateral cochlear implantation in adults: A multicenter clinical study, Ear Hear. 2006; 27:714–31.

22. Wackym PA, Runge-Samuelson CL, Firszt JB, Alkaf FM, Burg LS. More challenging speech-perception tasks demonstrate binaural benefit in bilateral cochlear implant users, Ear Hear. 2007; 28: 80S–85S.

23. Tyler RS, Dunn CC, Witt SA, Noble WG. Speech perception and localization with adults with bilateral sequential cochlear implants, Ear Hear. 2007; 28: 86S–90S