Speech Perception under Spectral Shift and Compression-Expansion Conditions: A Cochlear Implant Simulation Study

Ganesh A.C[1]., Ajith Kumar U[2] & Subba Rao T.A[1]

Abstract

The effects of spectral shift and spectral distortions (individual and combined) on speech recognition were measured using acoustic simulations.The present study included 30 participants with normal-hearing (NH), listening to acoustic cochlear implant (CI) simulations to measure speech recognition under spectrally matched, shifted, compression and expansion conditions for Kannada Quick Sin sentences, generated by 16-channel, noise-band vocoder. In the shift-only condition, the frequency ranges of the carrier bands were shifted up or down relative to the analysis bands. In the compression- and expansion-only conditions, the analysis band-range was made larger or smaller, compared to the carrier band range. Measurements were made in noise (with four talker babble at 0 dB SNR). A statistically significant difference was present for both, compression and expansion conditions when compared to the perfect frequency and place matched condition. Scores were significantly poorer in the expansion compared to compression condition. Information from the present study will help in understanding the effect of spectral manipulations on speech recognition ability in noise.

Key words: *Cochlear implant simulation, speech recognition, MATLAB, spectral distortions*

Cochlear hearing loss is the most common form of hearing loss in adults and is typically associated with the loss of sensory hair cells within the cochlea. Cochlear Implants (CI) are one of the treatment options available for individuals with severe to profound cochlear hearing loss. There is a frequency-to-place mapping along the basilar membrane in cochlea that is known as tonotopic organization. This organization is maintained throughout the auditory system upto the level of the primary auditory cortex. CI speech processing attempts to replace the function of the cochlea that is relevant for speech understanding. Modern cochlear implants have multiple electrodes that are inserted 20 to 30 mm into the scala tympani. In a CI device, the processor controls how the spectral content of the acoustic input is assigned onto the electrodes. The active length of the electrode array and its insertion depth primarily determine the tonotopic range of stimulation in the cochlea.

In CIs, there may be a discrepancy between the input acoustic frequency range assigned onto the electrodes and the tonotopic stimulation range, due to spectral distortions like spectral shift, spectral compression, and spectral expansion (Baskent & Shannon, 2003; Baskent & Shannon, 2004; Baskent & Shannon, 2007; Fu & Shannon, 1999). In spectral shift there will be mismatch between the acoustic input frequencies and the cochlear stimulation range. For example, shift might be the result of an electrode that is not fully inserted, while in other cases the shift may be due to the assignment of default signal processing parameters. Other types of spectral distortions can be caused due to compressive or expansive mapping of the acoustic input spectrum onto the stimulation range of the electrodes.

When the electrode array is fully inserted, the most apical contact is usually 20-30 mm from the round window, depending on electrode type. The active stimulation range of the electrode array ranges from 13.5 mm to 26.4 mm for devices like Nucleus 24, MED EL Maestro and Harmony. For example, an array of 16 mm would stimulate a cochlear region that corresponds to an acoustic frequency range of 500-6000 Hz for a 25 mm insertion depth according to Greenwood's frequency-place function (Greenwood, 1990). Most of the cochlear implant manufacturers allocate a wider range of frequencies in spite of the limited cochlear region. For example, the Harmony device assigns an acoustic range of 100-8000 Hz while the default frequency allocation for the Nucleus CI-24 implant assigns a frequency range of 150 Hz-10 kHz to the electrodes. As a result, the acoustic frequency range assigned to the stimulation region in the cochlea can be wider or narrower than the acoustic characteristic frequency range of that region, resulting in compression or expansion of the frequency-to-place mapping, respectively.

The results of several investigations in both CI listeners and young normal hearing individuals listening to CI simulations, suggest that a frequency-to-place mismatch can cause significant decreases

1. Dr M.V.Shetty College of Speech and Hearing, Mangalore
2. All India Institute of Speech and Hearing, Mysore

in speech perception abilities (Dorman, Loizou, & Rainey, 1997; Fu & Shannon, 1999; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). Conditions like frequency mismatch can negatively affect the perception of various types of speech stimuli such as phonemes, words, and sentences (Baskent & Shannon, 2003; Shannon, Zeng, & Wygonski,1998). Speech recognition performance deteriorates when the acoustic input range is made wider or narrower when simulated using noise band vocoder processing, and measured as a function of the acoustic input range on normal hearing people and implant users (Baskent & Shannon, 2003; Baskent & Shannon, 2004). In CI patients, both types of spectral distortions may occur simultaneously (Baskent & Shannon, 2004; Baskent & Shannon, 2007; Fu & Shannon, 1999).

There is a large variance in speech performance among implant recipients due to factors like duration of deafness, age of onset of deafness, age at implantation, and duration of CI use. Other factors that may affect auditory performance include: (1) number of surviving spiral ganglion cells, (2) electrode placement and insertion depth, (3) electrical dynamic range, and (4) signal processing strategy. Unfortunately, it is not easy to assess the significance of individual factors on speech perception due to the interaction among the above mentioned factors. Acoustic simulations have proven very useful in studying the effects of manipulations in electrode insertion depth, and the role of spectral or temporal cues on speech performance (Dorman et al.,1997; Shannon et al.,1995; Shannon et al., 1998).

Till date only a few attempts have been made in the Indian context to study spectral distortions on speech recognition. Nambi and Anusha (2008) studied the perception of spectrally degraded speech and its implications for CIs on eight normal hearing participants using an 8-channel sine-wave vocoder to simulate a CI. They investigated the effect of spectral shift by mimicing different electrode insertion depths in the following conditions: no spectral shift, monaural 25 mm insertion, monaural 20 mm insertion, and binaural asymmetrical shift. Results showed a significant main effect of spectral shift on speech recognition in all four conditions.

The effect of spectral compression-expansion and shifted sentences (e.g., the ability to correctly recognize and identify spectrally degraded stimuli without training or adaptation) has yet to be thoroughly investigated in the Indian context. The number of cochlear implant recipients in India is rapidly increasing, therefore research on various parameters that affect speech recognition performance is necessary. Baskent and Shannon, who contributed several research studies to this topic, used vowel and consonant identification tasks in quiet condition to assess the effect of spectral shift and the resulting distortion. However in the current study, a sentence recognition task in the presence of multi-talker babble noise was used to reflect a more natural acoustic environment. Hence the present study aimed at measuring the effects of spectral shift and spectral distortions individually and combined, on speech recognition in normal hearing participants using acoustic simulation.

Method

Participants

Thirty (10 males and 20 females) young adults with normal hearing sensitivity in the age range of 20 to 30 years participated in the study. All participants were native speakers of Kannada and none had participated previously in any other experiment using CI simulations.

Test materials and procedure

The test material consisted of 90 sentences which had been developed for the Kannada Quick Speech in Noise test (Avinash, Meti, & Kumar, 2010). The sentences could be identified with 100% accuracy in quiet situation by normal hearing individuals. Each sentence had five key words. The sentences were uttered by a male speaker at a comfortable level and recorded in a sound-attenuated room with a Zebronics microphone held at a distance of about 7-10 cm from the mouth. Sentences were recorded using the PRAAT (Boersma & Weenink, 2008) software with a sampling frequency of 44.1 kHz. Four-talker babble at 0 dB SNR was added to all the recorded sentences prior to the processing of the stimuli.

Mapping Conditions

Sixteen-band noiseband vocoder processing was implemented in Matlab (Math works, Natick, MA, 2007). The simulation of the CI processing has been shown graphically in Figure 1 where the analysis bands represent the input frequency range of the CI-speech processor and the noise carrier band represents the electrode array of the CI.

The process of creating the simulations was as follows. The signal was first processed through a band-pass filter (Table 1 explains the exact frequency ranges and cut-off frequencies for the specific experimental conditions) with a slope of 18 dB per octave. Each input frequency range was further band-passed into sixteen spectral bands using sixth-order Butterworth filters. Higher order Butterworth filters

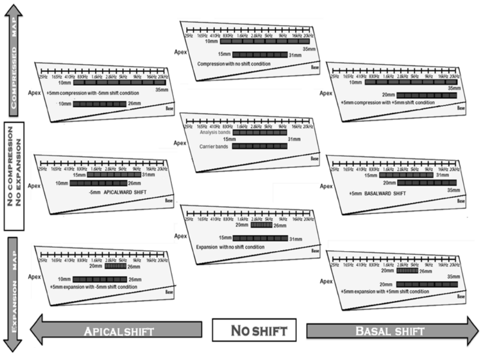

Fig.1. Frequency-place mapping conditions (8-channel processor) with simulated 20-mm electrode insertion depth. The different conditions are A) Matched condition (Middle), B) Shift-only map with no compression or expansion, where the carrier bands were shifted basally to higher frequencies (middle right), C) Shift-only map with no compression or expansion, where the carrier bands were shifted apically to lower frequencies (middle left), D) Compression with no shift, where the analysis band range was wider than the carrier band range (middle top), E) Apically shifted maps combined with compression, F) Basally shifted maps combined with compression (top right), G) Expansion maps with no shift, where the analysis band range narrower than the carrier band range (middle bottom), H) Apically shifted maps combined with expansion, I) Basally shifted maps combined with expansion (bottom right).

can be unstable and this limitation should be kept in mind while interpreting the results of the present study (Gaydecki, 2004). The cochlear locations of the two end frequencies for each input frequency range were calculated according to Greenwood's formula (Greenwood, 1990). The estimated cochlear distance was evenly divided into sixteen spectral channels. The estimated cochlear location for each spectral channel was then transformed back into frequency, again using Greenwood's formula. The Greenwood function is species-dependent and has shown to be preserved in mammals when normalized to the species-dependent range of auditory frequencies and cochlear spiral length (Greenwood, 1990). For humans, it recommends the value f = 165.4(102.1x − 1) for the constants. According to Greenwood's paper, x is relative to the cochlear length, and a = 0.06 if x is calculated in mm.

The corner frequencies (3 dB down) of the analysis filters are listed in Table 1. The temporal envelope was extracted from each frequency band by half-wave rectification using a third-order Butterworth filter with a cut-off frequency of 160 Hz at –3 dB and a filter slope of –18 dB per octave. For each channel, a noise band carrier was generated using the same set of band-pass filters as the analysis bands. The modulated carrier bands were summed to produce the processed sentences. All the analysis and carrier bands had the same overall root mean square (RMS) energy. .

The stimulation region covered by the simulated electrode array was fixed at 16 mm for all the shift, expansion, and compression conditions (refer to Figure 1). The 20 mm insertion depth condition simulated an electrode array located between 4 and 20 mm from the round window. Because the stimulation region was fixed at 16 mm, electrode locations represented by noise bands were 1 mm wide for the 16- band condition.

The frequency of the noise wave carrier was equal to the center frequency of the analysis band. This served as a baseline condition in which frequency ranges of the corresponding carrier and analysis bands were the same. For the spectrally shifted condition, the output carrier bands were basally shifted or apically shifted by 5 mm. The compressed

Table 1. *Analysis (first row) and carrier band (second row) ranges for different cochlear positions (in mm) and their respective frequency ranges (Hz).*

	Shift in carrier bands		
Compression-expansion	-5 mm Apical shift	0 mm No shift	+5mm Basal shift
Compression (+5 mm)	10-35(513–20677) 10-26(513–5860)	10-35(513–20677) 15-31(1168–11837)	10-35(513–20677) 20-35(2476–20677)
No compression- expansion (0 mm)	15-31(1168–11837) 10-26(513–5860)	15-31(1168–11837) 15-31(1168–11837)	15-31(1168–11837) 20-35(2476–20677)
Expansion (-5 mm)	20-26(2476–5860) 10-26(513–5860)	20-26(2476–5860) 15-31(1168–11837)	20-26(2476–5860) 20-31(2476–20677)

Table 2. *Means and standard deviations of the speech recognition scores for different spectral mismatch conditions.*

	Shift in carrier bands		
Compression-expansion	-5 mm Apical shift	0 mm No shift	+5mm Basal shift
Compression (+5 mm)	63.26 (8.92)	67.92 (6.4)	69.12 (7.62)
No compression-expansion (0 mm)	87.6 (5.78)	92.26 (4.62)	95.26 (2.94)
Expansion (-5 mm)	6.06 (1.67)	1.66 (0.91)	3.72 (1.42)

map without shift was produced by manipulating the analysis bands only; the analysis band range was made wider by 5 mm in cochlear distance at each end. The expanded map without shift was produced by manipulating the analysis bands only; the analysis band range was made narrower by 5 mm in cochlear distance at each end. All the stimulus manipulation was implemented in Matlab, with the help of a Matlab code (Baskent & Shannon, 2007). With this stimulus manipulation CI speech was simulated to mimic the following conditions 1) Matched mapping condition without any shift or compression, 2) 5 mm apical shift, 3) 5 mm basal shift, 4) 5 mm compression without any shift, 5) 5 mm apical shift with 5 mm compression, 6) 5 mm basal shift with 5 mm compression, 7) 5 mm expansion without any shift, 8) 5 mm apical shift with 5 mm expansion, and 9) 5 mm basal shift with 5 mm expansion. Figure 1 shows the frequency-place maps used in the experiment.

Procedure

The stimuli were presented from a Compax CQ-45 personal computer connected to high fidelity Tech-Com Digital Sound stereo headphones (SSD-HP 202). The volume was adjusted to the participant's comfortable level. The testing was done in a sound-treated booth. The presentation order of the conditions was randomized. All participants were tested once in each experimental condition. They were instructed to repeat what they had heard. Stimuli were not repeated. Each correctly repeated key word received a score of '1' while an incorrectly repeated word received a '0'. The score of correctly repeated key words for each sentence (out of 5) and a total score (out of 50) was calculated. This was then converted into percent-correct score.

Results

Percent-correct scores for sentences were obtained with 16 channel processing at simulated insertion depths of 20 mm as a function of shift, compression or expansion in the frequency-place mapping. (refer to Figure 2) The means and standard deviations of shift and compression-expansion for the speech recognition scores are presented in Table 2.

As evident from Figure 2 and Table 2, speech recognition in the matched conditions was generally better than for any condition of frequency shift or place expansion-compression. The frequency-shift-only condition yielded better responses (mean value 91.43%), compared to the shift and compression (66.50%) condition. The Shift and Expansion (3.75%) produced poor responses.

A repeated measures ANOVA test was used to assess the significance of differences in speech recognition scores between different conditions. Speech recognition for the matched conditions was generally better than for any condition of frequency shift and place expansion-compression except for basal shift. Analysis of the speech recognition data revealed that the assumption of sphericity was violated, hence the Greenhouse-Geisser correction was applied. Results indicated a significant main effect of condition [$F(3.86, 111.99) = 2263.81$, $p<0.0001$]. For each condition, the result was compared to that of the perfect frequency-place-match (carrier and the analysis bands were perfectly matched) with the eight other conditions (frequency shift, compression and expansion individually, and combined) using pair-wise comparisons with Bonnferroni's adjustments.

Post-hoc test results showed significant difference ($p <0.05$) between the perfect frequency-place-match condition and each of the other eight conditions. There was statistical significance ($p <0.05$) The scores for the combined shifted and compressed condition were significantly different ($p < 0.05$). Further, the scores for the combined shift and expansion condition were significantly different ($p < 0.05$) from those for the shift-only condition. There was no significant statistical difference when the compression-alone condition was compared with the combined shift and compression condition ($p >0.05$).

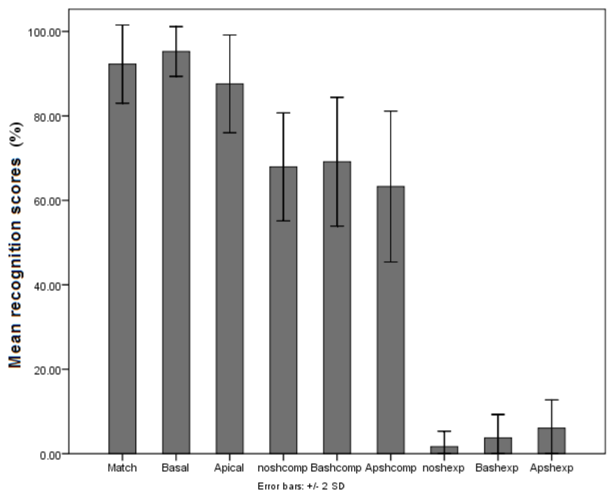

Fig 2. Mean sentence recognition scores (%) for noise carrier bands simulating a 20 mm electrode insertion depth, as a function of shift (basal, apical), compression (noshcomp: no shift and only compression; bashcomp: basal shift and compression; apshcomp: apical shift and compression) & expansion (noshexp: no shift and only expansion; bashexp: basal shift and expansion, apshexp- apical shift and expansion) in the frequency-place mapping.

Discussion

Results revealed a statistically significant difference in performance for the perfect frequency-place-match condition versus both compression and expansion conditions. The reduction in scores due to expansion was more drastic compared to that caused by compression. Limited bandwidth contributed to poor performance in the expansion conditions (Baskent & Shannon, 2004; Baskent & Shannon, 2007). The results of the present study are in agreement with the literature on vocoder studies which have generally shown a reduction in speech recognition for frequency shift and spectral mismatch conditions (Baskent & Shannon, 2003; Baskent & Shannon, 2004; Dorman et al., 1997; Shannon et al., 1998). For simulation of a 16 mm electrode array, the shift only condition produced better performance compared to the shift and compression condition followed by the shift and expansion condition. The results of the current study thus suggest that a spectral shift and compression-expansion of 5 mm can affect speech recognition performance of young listeners.

The effects of frequency shifting on speech recognition have been documented in studies going back several decades to the 1960's and 70's. Results from three early (Daniloff, Shiner, & Zemlin, 1968; Nagafuchi, 1976; Tiffany & Bennett, 1961) indicated that the listeners could only tolerate up to a frequency shift of about 35% before their performance worsened significantly. When the speech frequencies were shifted by 60%, speech recognition was reduced to about 20% correct.

The common finding of studies on tonotopic shift and compression or expansion has been that speech recognition is best when the acoustic input range matches the simulated stimulation range, and decreases when spectral distortions in the form of a shift or compression-expansion are introduced (Baskent & Shannon, 2003; Baskent & Shannon, 2004; Dorman et al., 1997; Fu & Shannon, 1999; Fu, Shannon, & Wang, 1998). The present study aimed to explore the combined effects of two spectral distortions, shift and compression-expansion, on speech recognition. Similar to previous studies, the matched frequency-place map produced the best performance while the application of either shift or compression-expansion combined, resulted in lower scores. However, the results of the present study were not exactly similar to that of Baskent and Shannon study (2007). In the present study, the basal shift with compression produced better results compared to no shift with compression condition and even similar thing can be observed in which basal shift produced better results than matched baseline condition. Some of the reason for these different findings is that difference in the stimuli such as sentences in the present study and other factors may be due to basal shift, the intelligibility of the sentences is increased.

Also according to Fu and Shannon (1998) the effect of spectral shift was clearly dependent on talker gender, with male voices tolerating more basal shift and children's voices tolerating more apical shift. This may be the reason for the present findings on better performance on basal shift. Fu et al. (1998) observed no significant performance drop for tonotopic shifts equivalent to 3 mm along the basilar membrane, that is, for frequency shifts of 40%–60%. Dorman and colleagues measured speech recognition with a five-band sinusoid vocoder in conditions in which the carrier bands were shifted up to 5 mm basally from the analysis bands. Speech recognition performance dropped off dramatically (10% per mm) as the carrier bands were shifted basally, particularly for vowels (Dorman et al., 1997). In the present study, there was no significant drop in performance on both shift conditions as compared to the matched-frequency-place condition, like in previous studies.

The present results on compression and expansion are remarkably similar to previous studies of Baskent and Shannon (2003, 2004, 2007) in which the whole speech signal was frequency compressed or expanded. According to Baskent and Shannon (2003) manipulating the analysis and carrier bands by narrowing the overall bandwidth, results in decreased performance, which deteriorates further by adding expansion via widening the carrier band range When compression is added to the shift-only condition, more of the speech spectrum is included in the analysis band range. Due to this, the compression and shift condition results in reduced speech recognition scores only by a small amount compared to expansion. The present study shows that when there is substantial spectral distortion, matching the frequencies that are important for speech perception might be more beneficial than matching the overall acoustic input range.

In general, a large variance in speech performance is observed among CI recipients due to factors such as number of surviving spiral ganglion cells, electrode placement, and insertion depth. Determining the exact location of an implanted electrode needs a sophisticated radiological test such as CT scan, which is expensive in developing countries like India. So in normal's we can control these factors by making the measurements in normal-hearing listeners by simulating CI. The present results may help to know effect of electrode depth, number of electrodes, and frequency range on speech perception. A number of studies have shown a relation between electrode array length and location and speech perception in CI (Baumann & Nobbe, 2006; Blamey et al., 1996; Boex et al., 2006; Ketten et al., 1998; Pfingst, Franck, Bauer, & Zwolan, 2001; Skinner et al. 2002; Yukawa et al., 2004;)

Also studies have shown that spectral distortions such as apical or basal shift, nonlinear warping and compression or expansion of the applied frequency map decrease speech perception (Baskent & Shannon, 2003; Baskent & Shannon, 2004; Dorman, Loizou, & Rainey, 1997; Fu & Shannon, 1999; Shannon, Zeng, & Wygonski, 1995). Hence the literature as well as the present study suggest that improvement of CI performance may depend on optimizing the tonotopic mapping of individual channels relative to the actual position of stimulation sites in the cochlea.

The present study focused on studying the ability of normal hearing participants to correctly recognize and identify spectrally degraded stimuli without training or adaptation. Although studies have shown that speech perception with a frequency-shifted implant map can substantially improve with training or adaptation, but ultimately the proper mapping will provide maximum benefit because improved performance due to training with the shifted map at the end of the experiment was lower than the matched map (Fu, Nogaki & Galvin, 2005; Nogaki, Fu & Galvin , 2007; Rosen, Faulkner & Wilkinson, 1999)

The results of this simulation study should be considered with some prudence as it is almost impossible to establish in an unequivocal and definitive way, that a synthesized signal represents exactly, the perception through a cochlear implant or the perception of an individual who has had hearing loss.

Conclusions

The present study aimed to investigate the combined effects of the two spectral distortions, shift and compression-expansion, on speech recognition. Results from the present study reveal that speech recognition is affected when spectral distortions are introduced. Information from the present study on speech recognition in spectrally shifted and compressed-expanded conditions will help in understanding the sensitivity of speech recognition to spectral manipulations.

Acknowledgement: The authors express sincere gratitude to Deniz Başkent, PhD, Assistant Professor, University Medical Center Groningen, Department of Otorhinolaryngology, University of Groningen, School of Behavioral and Cognitive Neuroscience, The Netherlands for providing the MATLAB code to carry out the present study. The authors also wish to thank the anonymous reviewers for their valuable suggestions on the previous versions of this manuscript.

Address for correspondence: Ganesh A.C., Dr. M. V. Shetty College of Speech & Hearing, Maladi Court, Kavoor, Mangalore 575015, Email: gani0017@gmail.com.

References

Avinash, M. C., Meti R, & Kumar U., (2010). Development of Sentences for Quick Speech-in-Noise (QuickSin) Test in Kannada. *Journal of Indian Speech and Hearing Association, 24*(1), 59-65

Baskent, D., & Shannon, R. V. (2003). Speech recognition under conditions of frequency-place compression and expansion [Electronic version]. *The Journal of the Acoustical Society of America, 113,* 2064–2076.

Baskent, D., & Shannon, R. V. (2004). Frequency-place compression and expansion in cochlear implant listeners [Electronic version]. *The Journal of the Acoustical Society of America, 116,* 3130–3140.

Baskent, D., & Shannon, R. V. (2005). Interactions between cochlear implant electrode insertion depth and frequency-place mapping [Electronic version]. *The Journal of the Acoustical Society of America, 117,* 1405–1416.

Baskent, D., & Shannon, R. V. (2007). Combined Effects of Frequency Compression-Expansion and Shift on Speech Recognition. *Ear and Hearing, 27,* 277-289.

Baumann, U., & Nobbe A. (2006). The cochlear implant electrode-pitch function [Electronic version]. *Hearing Research, 213,* 34–42.

Blamey, P., Arndt, P., Bergeron, F., Bredberg, G., Brimacombe, J., Facer, G., & Larky J., et al. (1996). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiology & Neurotology, 1,* 293–306.

Boersma, P., & Weenink, D. (2008). Praat (Version5.0.17) [Computer Program]. Doing Phonetics by computer

Boex, C., Caud, L., Cosendai, G., Sigrist, A., Kos, M. I., & Pelizzone, M. (2006). Acoustic to electric pitch comparisons in cochlear implant participants with residual hearing [Electronic version]. *Journal of the Association for Research in Otolaryngology: JARO, 7,*110–124.

Daniloff, R. G., Shiner, T. H., & Zemlin, W. R. (1968). Intelligibility of vowels altered in duration and frequency. *The Journal of the Acoustical Society of America, 44,* 700– 707.

Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. *The Journal of the Acoustical Society of America, 102,* 2993–2996.

Fu, Q.-J., Shannon, R.V., & Wang, X. (1998). Effects of noise and number of channels on vowel and consonant recognition: Acoustic and electric hearing [Electronic version]. *The Journal of the Acoustical Society of America, 104,* 3586-3596.

Fu, Q.-J., & Shannon, R. V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing [Electronic version]. *The Journal of the Acoustical Society of America, 105,* 1889–1900.

Fu, Q.-J., Nogaki, G., & Galvin, J. J. (2005). Auditory training with spectrally shifted speech: implications for cochlear implant patient auditory rehabilitation. *Journal of the Association for Research in Otolaryngology: JARO, 6,* 180-189

Gaydecki, P. (2004). *Foundations of Digital Signal Processing: Theory, Algorithms, and Hardware Design.* United Kingdom: IET Publishing.

Greenwood, D. D. (1990). A cochlear frequency-position functions for several species—29 years later [Electronic Version]. *The Journal of the Acoustical Society of America, 87,* 2592–2605.

Ketten, D., Skinner, M. W., Wand, G., Vannier, M. W., Gates, G. A., & Neely, J. G. (1998). In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays. *The Annals of Otology, Rhinology, and Laryngology, 107,* 1–16. [Pub Med: 9826942]

Nagafuchi, M. (1976). Intelligibility of distorted speech sounds shifted in frequency and time in normal children. *Audiology, 15,* 326–337.

Nambi, A., & Anusha (2008). *Perception of spectrally degraded speech: implication for cochlear implants on 8 normal hearing subjects.* Paper presented at the 40th Annual conference of Indian Speech and Hearing Association, Mangalore, India.

Pfingst, B. E., Franck, K. H., Bauer, E. H., & Zwolan, T. A. (2001). Effects of electrode configuration and place of stimulation on speech perception with cochlear prostheses. *Journal of the Association for Research in Otolaryngology: JARO 2*(2), 87–103.

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

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

Shannon, R. V., Zeng, F. G., & Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. *The Journal of the Acoustical Society of America, 104,* 2467–2476.

Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., et al.(2002). CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients. *Journal of the Association for Research in Otolaryngology: JARO, 3,* 332–350.

Tiffany, W. R., & Bennett, D. A. (1961). Intelligibility of slow-played speech. *Journal of Speech and Hearing Research, 4,* 248-258.

Yukawa, K., Chen, L., Blamey, P., Pyman, B., Tungvachirakul, V., & O'Leary, S. (2004). Effects of insertion depth of cochlear implant electrodes upon speech perception. *Audiology & Neurotology, 9,* 163–172.