The problems of cochlear implants and music perception and what we can change: a review.

By: Julian Bakker

Supervisor: Deniz Başkent

Bachelor essay

The problems of cochlear implants and music perception and what we can change: a review.

Abstract:
Music is enjoyed by millions over the world. Unfortunately cochlear implant users are not able to enjoy music as normal hearing people can. Aside from not being able to enjoy it, they also do not benefit from the benefits music has. It has used to treat different disorders like OCD and depression, has proven useful in speech perception and is even able to “deceive” people about their physical status. Making them feel much better than they actually physically are. The CI has problems on many fronts. Fundamental frequencies are not picked up well by CI users resulting in problems with pitch perception. Timbre is not relegated by CIs which causes all instruments sound the same. The only thing CI users seem to be able to pick up when listening to music are clear lyrics and rhythm. A solid solution is not easy to find. There are problems with the quality of the sound itself, increasing fidelity could lead to clearer sharper sound. This is however dependent on the advances in the electronics of CIs and a actively developed but years away from reaching normal hearing fidelity. Another possibility is adding presets that can be selected by the user. CIs boost certain frequencies which are useful for speech but can hinder the perception of music. Giving presets might remedy this. Furthermore the so called musician effect might be useful. If CI users were to receive musical training they could learn to value music as well a gain better pitch perception which will increase the perception of music as well as help speech perception.

Index:

Introduction:

Music, everyone listens to it and for good reason. Music is seen as beneficial to quality of life and has many other benefits, for example: treating depression and OCD. There are however people who cannot enjoy music to its fullest. Cochlear implant (CI) users and other hearing impaired people do not experience music the way normal hearing (NH) people do. This script is about why this could be the case and what we can do to improve their perception of music. To look at ways to improve the way music is experienced by CI users, we will first visit how natural hearing works and determine the differences with electrical hearing (i.e. hearing with a CI) and determine in what aspects music differs from speech. The benefits of music will be explored and finally what can be improved for CI users.

Natural Hearing:

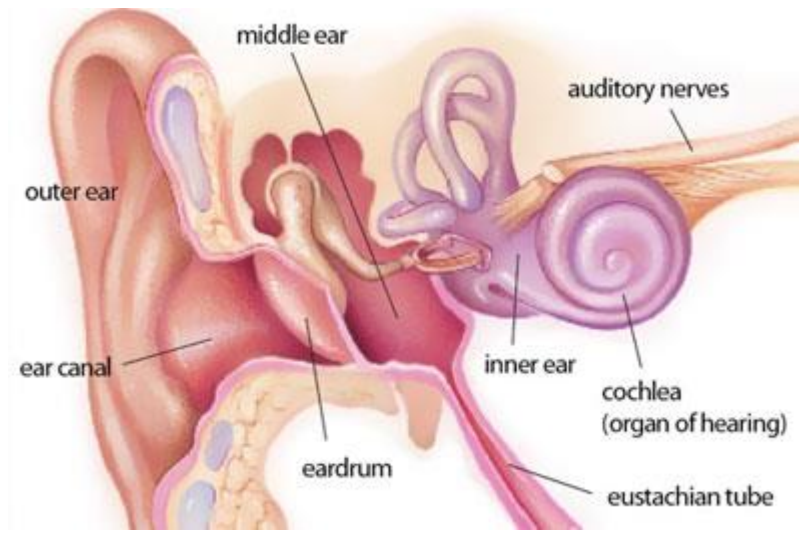

Figure 1: The human ear.
Source: http://www.findfunfacts.com/human_body/ear.html

The ear can be divided into three parts: the outer, middle and inner ear. The outer ear clusters the pressure changes from airborne auditory signals and directs them to the middle ear, furthermore it allows us to determine where a sound is coming from. By means of head-related transfer functions, one can use two ears to listen in three dimensions: distance, above or below and side, front or rear. By comparing arrival times and intensity differences between both ears one can determine the location of a sound source in the environment[1].

The middle ear starts at the eardrum, the tympanic membrane, and ends at the cochlea. The middle ear contains three tiny bones: the malleus (hammer), the incus (anvil) and the stapes (stirrup) which comprise the ossicular chain. The ossicular chain conducts the energy of the pressure changes from the eardrum to the cochlea using mechanical advantage. Connected to the tympanic membrane

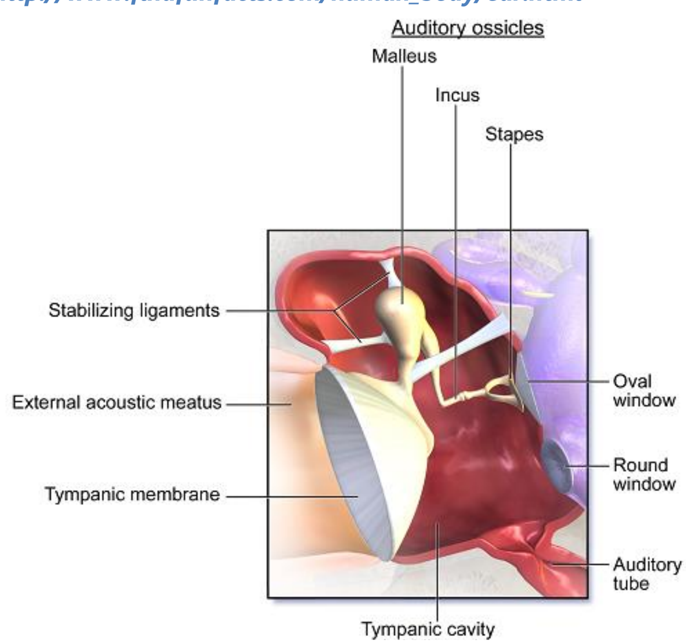

Figure 2: The ossicular chain
Source: Blaussen Gallery 2014 Wikiversity Journal of Medicine

is the malleus. When the tympanic membrane vibrates the malleus starts vibrating as well. The malleus is connected to the incus and forms a lever-like construction. The incus then relays the energy to the stapes. The footplate of the stapes is loosely connected to the oval window and transfers the energy to the oval window. Due to the lever like connection of the malleus and incus and the size difference of the eardrum and the smaller footplate of the stapes, the forces on the oval window can be 20 times larger. By using mechanical advantage the middle ear can enhance the intensity of the vibrations so they remain powerful enough to be able to travel through the pressurized fluid in the inner ear[1].

The cochlea, a snail shaped organ, contains three spiraling chambers known as the scalae: the scala tympani, scala vestibuli and scala media. The oval window vibrates the perilymph in the scala vestibuli, which ends at the round window. The round window bulges out when the oval window bulges in. The scala vestibuli is separated from the scala media by thin membrane known as the Reissner's membrane. The scala media and vestibuli however act a single duct, so the movements of the perilymph in the vestibuli cause the endolymph in the media to move as well. The scala media contains basilar membrane and on top of the basilar membrane lies the organ of Corti. The basilar membrane and the organ of Corti extend all the way to the end cochlea.

The organ of Corti functions similar to a Fourier analyzer, registering high frequencies at the base and low frequencies at the apex. The basilar membrane is stiffest at the beginning and gradually loses this stiffness towards the end. Because the membrane is stiff at its beginning, only the highest frequencies can make it vibrate. The lower the frequency of the sound wave, the further it travels along the length of the membrane with the least stiff part at the end of the basilar membrane reacting to the lowest frequencies one can hear. In addition, the coiling of the cochlea enhances the low frequency vibrations to allow them to reach the end of the cochlea. This layout of the cochlea is known as the tonotopical layout[1].

The organ of Corti harbors the hair cells. As the basilar membrane vibrates, the outer hair cells amplify the vibrations and this displaces the inner hair cells. The inner hair cells then depolarize via tip-link channels which opened due to the vibrations. The inner hair cells are different in length, going from short to long. The cells are linked to each other via tip links which, when under enough tension, cause the inner hair cells to depolarize.

When the hair cells are displaced in the direction of the longer cells, the tip links are extended and the cell depolarizes. Neurotransmitters are then released to the spiral ganglion, which relays the now electrical information to the brain.

Music and Speech:

Speech and music are both complex sound waves with frequency, temporal, intensity and timbral components which are presented in an organized manner. Spectral and temporal envelopes vary in time in music as well as speech, but they also differ in many ways[2,5]. In speech and music the fundamental frequency (F0) plays an important part, however the range of frequencies is much greater in music than in speech. The accuracy with which the F0 is recognized is important when listening to music to recognize individual notes, but also to listen to the voice of your friend in a crowded stadium. The fundamental frequency is the lowest frequency of a sound. Formants, spectral shape and the offset or onset of transients define timbre, but are not essential for melody recognition[5].

Music has four main attributes: pitch, duration, loudness and timbre[2]. Loudness is dependent on intensity and pitch is primarily dependent on frequency, but also loudness can change the perception of pitch. Timbre is dependent on: rise time, spectral centroid and the spectral flux. Recognizing sound as music is primarily attributed to the perception of patterns in the sound one hears[3], where the sequencing of pitches correlates to melody and harmony and the sequencing of durations or temporal patterns correlate with rhythm[2]. The interactions of these attributes with each other make music the way we experience it. Aside from the aforementioned main attributes, a listener's musical experience, training, listening preferences, age, culture or demographics may affect the way the listener listens to music and the way it is experienced (e.g. what it means to one or which emotions are in play)[4].

Pitch in psychoacoustic terms is the ordering of sounds from low to high on a melodic scale, with systematic variations in pitch providing a sense of melody[6]. The Oxford dictionary defines pitch as: "The quality of a sound governed by the rate of vibrations producing it; the degree of highness or lowness of a tone". As mentioned before the F0 is the most important part of pitch perception, but intensity and other factors can have a small effect as well[7]. The harmonic content, especially the low harmonics, are the most important in determining the strength of the pitch[8].

Pure tone perception is associated with the location on the basilar membrane that frequency has, however there is no consensus yet about how complex tones are processed. A host of models has been suggested with the predominant theories divided into three classes: Pattern recognition, temporal based models or spectro-temporal models.

All these models have in common that the basilar membrane works like an array of bandpass filters which analyze the input signal and divide it into different frequencies. The information can be determined by resolving the lower harmonics of the complex tone[9-11]. Another way to gain pitch information is by resolving the temporal cues from unresolved harmonics. This results in a complex pattern with a repetition rate equal to the F0 which can be used as a clue to gain the information[12].

Timbre is the attribute defines the differences in sound between two different instruments playing the same tone. There are three dimensions to timbre: rise time, spectral centroid and spectral flux. The rise time is the onset or attack time, the spectral centroid is the center of the spectrum in respect to perceived brightness or dullness and the spectral flux is the number of components the spectrum has and the spread of these components[13]. Instruments all have their own specific timbre and this allows someone to recognize the instrument (i.e. a trumpet is mostly described as bright whereas a clarinet is usually more associated with a dull sound).

Music and its Benefits:

Music can be an important factor in quality of life (QOL), but music in the form of therapy has also proven to be a very effective method for treating all kinds of disorders. A study from 2003 looked at the quality of life and length of life of people with terminal cancer. The Participants their QOL was determined before and after musical therapy using the Hospice Quality of Life Index-Revised (HQLI-R), a self-report questionnaire designed for cancer patients. It looks at different aspects of life: functional (i.e. daily activities, social life, concentration ability, enjoyable activity), psychophysiological (i.e. anger, pain, nausea, sex life, worry, anxiety) and social/spiritual (i.e. meaning in life, physical contact, family support, relationship with god). Higher scores indicate a higher perceived quality with 280 being the maximum. To determine if their physical status deteriorated the Palliative Performance Scale (PPS) was used. The PPS determines treatment effects at various stages of physical decline. A nurse determines the conscious level, intake, self-care, ambulation, activity and evidence of disease of a participant.

The lowest score 0% is equal to death and 100% indicates a healthy person. The control group showed a decrease between QOL test 1 and 2. Participants that underwent musical therapy had an increasing QOL and more interesting their PPS scores went down. They became more ill, but their QOL scores increased. Showing the effect music therapy can have[14].

Music therapy has also been used to treat OCD and co-morbid anxiety and depression. 30 participants diagnosed with OCD were randomly assigned in either "standard treatment" or "standard treatment + musical therapy". Both groups underwent Cognitive Behavioral Therapy. The musical therapy consisted of sessions of receptive musical therapy three days a week for four weeks. The participants were asked to do several self-administered tests: the Maudsley Obsessive-Compulsive Inventory (MOCI), the Beck Anxiety Inventory (BAI) and the Beck Depression Inventory – Short Form (BDI-SF) twice, once before and once after the therapy. It showed that the musical therapy had a good effect on the group that underwent the therapy. They showed less obsessive behavior, were less depressed and relieved anxiety. For anxiety and depression the musical therapy + standard treatment was approximately 47% more effective than standard treatment alone[15].

Aside from treating disorders, music can also increase speech perception skills in the normal hearing, especially for challenging listening tasks[16]. Musicians appear to be better at the decoding of vocal sound, perception of voice cues and better at pitch cues in speech as well as in music[17]. Aside from training the ear, this might also improve quality of sound for hearing impaired in and of itself. A study performed by Fuller et al. looked at the so called musician effect when musicians and non-musicians were presented with CI simulated sound. By using temporal and spectral degradation sound was modulated to appear as if the listener had a CI. Although all performance was poorer when simulated speech was presented, the musician effect persisted and musicians had higher word recognition scores than the non-musicians[18].

Cochlear Implants:

An integral part of this study are the CI users. A CI is an implant that helps people with loss of hearing of various degrees to regain their hearing. CI's are implanted in people with a functional auditory nerve system, but have severe hearing loss which is 70+ Db on average. The implant takes over the function of the hair cells in the inner ear, transforming sound into an electrical signal the brain can process.

By placing an array of electrodes in the cochlea and sending electrical pulses directly to the cochlear nerves, hearing can be restored. The external microphone picks up sound and the sound is divided into frequency bands by the speech processor connected to the microphone. The bits of sound are send to the receiver, which is a computing unit as well and is placed next to the skull subcutaneous, which transforms the sound information into an electrical signal. The receiver conveys the signal to the array of electrodes in the cochlea which relay the electrical signal to the auditory nerves, high frequencies at the base of the cochlea and low frequencies at the apex corresponding to the tonotopical layout of the cochlea.

Having a cochlear implant restores your hearing, but many studies indicate that the perception of sound is different[19, 20]. A decrease in pitch discrimination ability, for pure as well as complex tones, has been observed[21, 22]. Looi et al. set up a pitch ranking test and a melody recognition test. The pitch ranking test used recordings of the vowels /a/ and /i/ from trained singers (male and female) at a number of fundamental frequencies (e.g. 98, 370, 740 Hz). Each vowel was presented eight times using three different intervals in the whole octave, three intervals in half the octave and four intervals in a quarter octave. The melody recognition test used ten excerpts of familiar melodies on an electronic keyboard, set to produce clarinet and oboe sounds. 15 CI users participated in the study. The control group scored almost 100% on all three subsets and almost 100% on the melody test. CI users however scored 70% correct on the whole octave subset, lowering to 50% in the quarter octave subset and a 50% correct on the melody test[23]. Looi et al. demonstrated that tones that are reasonably far apart can be distinguished by CI user, but when the range of the pitch of the tones decreased CI users clearly had trouble distinguishing the different tones from each other.

The cochlea in NH people is non-linear and level-dependent, whereas a CI's sound processor has a limited amount of filterbands. Wide filters may interfere with the lower harmonics of complex sounds making it hard to resolve them. This in turn makes it difficult for the user to determine the harmonic frequencies and make a pitch judgement. Even if the components were resolved more precisely, it would be impossible to determine the exact frequency, as the CI user can only determine to which filter the component was sent to and the filters themselves are not precise enough. If the components are spread over adjacent filters which in turn activate adjacent electrodes it becomes difficult to determine pitch information with great precision.

Distinguishing one electrode from another is dependent on the degree of spatial specificity of the individual electrodes and the current spread in the cochlea[4]. Furthermore, the depth to which the electrodes reach is most likely insufficient. The lower frequencies of the electrode stimulate parts of the cochlea that are actually associated with a higher frequency (e.g. 150 Hz stimulation on the electrode stimulates the part of the cochlea associated with 300 Hz).

A study in which modified speech samples were used that simulates speech as heard by CI users, demonstrated the problems CI users can have with understanding speech[24]. The samples were recordings of basic Dutch sentences (e.g. "Buiten is het donker en koud." English: "Outside it is dark and cold") that were filled with steady-state noise and were then further modified using a noise-band vocoder. The bandwidth was limited to 150-7.000 Hz and a spectral degradation of 4, 8, 16 and 32 channels was applied using a Butterworth filter. Participants (normal hearing) were instructed to listen to the different samples and were asked to repeat the sentence that was spoken. The more noise (4 channels being the most, 32 the least amount of noise) the less intelligible the samples were. 16 Channel vocoded samples attained a 50% correctly heard sentence rate. 8 channel vocoded samples were most in line with speech as heard by CI users. Indicating that improvements can still be made to the audio quality of CIs.

Music and CIs:

Leek et al. asked 68 hearing aid wearers about their enjoyment of music and its importance. Two third of the questioned indicated that music was important to them. 28% reported that having a hearing loss negatively impacted their enjoyment of music[25].

The perception of timbre, the attribute that sets apart instruments, is generally not great in CI users[26]. Most timbre tests are based on instrument recognition. A comparison between CI and NH participants was made in respect to their ability to differentiate eight different instruments. The instruments played the same seven notes long melody. 91% of the NH participants were able to distinguish the instruments. Yet only 47% of the CI users was able to distinguish the instruments from one another[27].

Melodies and CIs don't go well together either as seen in the aforementioned study by Looi et al.. Gfeller et al. set up a similar study with 36 familiar melodies, which were divided into three genres.

The 59 CI users had a mean recognition rate of 16%, considerably lower than the 55% of the normal hearing group. The classical genre had the worst score as it was purely instrumental, whereas pop and country did better due to the fact that the CI users were able to recognize the lyrics[28].

The ability to determine the rhythm of the music however is not significantly worse in CI users. The rhythm is mostly in the range of 0.2Hz to 20 Hz. Gfeller and Lansing developed the primary measures of music audiation (PMMA) test, a test to determine rhythmic skill in children and other musical skills. The subtest for rhythm consists of pairs of short sequences with no variance in pitch or timbre. The sequences contain either identical or different rhythms and the listener must determine if the sequences have the same rhythm or not. 18 adult CI users were asked to do this test and their scores ranged from 80-95% correct[29]. In a similar study the ability to discriminate different tempos was investigated. Five CI users were charged with determining which of two sequences had the faster tempo, with a maximal increase or decrease of 20 BPM. The five CI users performed just as well as the control group of four NH. A change of 4-6 BPM was approximately the minimal discernible change in tempo[30].

What can we improve for CI users:

The most simple yet perhaps hardest way to improve the way CI users experience music is to improve the CIs themselves. If the CI were to be more precise, by using more electrodes and/or altering the location of the electrodes to allow for more precise delivery of the signal, all sound could increase in quality. These are however material and production obstacles and thus impossible to control.

More fine tuning of the CI might help increase the fidelity of music as well. Many CI’s have boosted low frequencies, around 1000 Hz, and boosted high frequencies, around 3000 Hz. These boosts are to help with speech recognition as vowels usually have frequencies at around a 1000 Hz and consonants have frequencies at around 3000 Hz. If the CI’s would be able to use presets, a user could switch from speech settings (boosts at 1000 Hz and 3000 Hz) to a music setting with no boosts or boosts at frequencies between 100-3000 Hz. The ranges indicated here are estimates and a separate study should be performed to determine the most efficient ranges to be boosted for music, as music covers a great range of frequencies practically simultaneously (e.g. base guitar, drum, female singer, a screaming guitar).

The so called musician effect is interesting to look into as well. Pitch perception is often worse in CI users. Musical therapy could help increase pitch perception, which in turn enhance the music experience and also speech intelligibility in general. Several training regimes for: timbre, pitch perception, melody recognition etc., could all help CI users enjoy music more bringing with it the benefits of listening music. Ranging from more joy to being able to understand someone in a crowded room.

References:

1: Geisler CD; From sound to synapse; physiology of the mammalian ear.; NY: Oxford University Press 1998

2: Krumhans CL, Iverson P; Perceptual interactions between musical pitch and timbre.; Journal Exp Psychol Hum Percept Perform 1992 18(3) 739-751

3: Gfeller K, Woodworth G, Robin DA, Witt S, Knutson JF; Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users.; Ear Hear 1997 18(3) 252-260

4: Looi V, Gfeller K, Driscoll V; Music appreciation and training for cochlear implant recipients: a review.; Semin Hear 2012 33(4) 307-334

5: Wolfe J; Speech and music, acoustics and coding and what music might be 'for'.; 7th international conference on music perception and cognition 2002 10-13

6: Moore BCJ; An introduction to the psychology of hearing 5th edition; Academic Press 2003

7: Morgan CT, Garner WR, Galambos R; Pitch and intensity.; Journal Acoust Soc Am 1951 23 658-663

8: Rasch RA, Plomp R; The perception of musical tones.; The psychology of music NY: Academic Press 1982 1-22

9: Moore BC; Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants.; Otol Neurol 2003 24(2) 243-254

10: Plomp R; Pitch of complex tones.; Journal Acoust Soc Am 1967 41 1523-1533

11: Ritsma RJ; Frequencies dominant in the perception of the pitch of complex sounds.; Journal Acoust Soc Am 1967 42 191-198

12: Moore BC, Rosen SM; Tune recognition with reduces pitch and interval information.; Q J Exp Psychol 1979 31(2) 229-240

13: Grey JM; Multidimensional perceptual scaling of musical timbres.; J Acoust Soc Am 1977 61(5) 1270-1277

14: Hilliard RE; The effects of music therapy on the quality of life and length of life of people diagnosed with terminal cancer.; J Music Ther 2003 40(2) 113-137

15: Shirani Bidabadi S, Mehryar A; Music therapy as an adjunct to standard treatment for obsessive compulsive disorder and co=morbid anxiety and depression: a randomized clinical trial.; J Affect Disord 2015 184 13-17

16: Wong PC, Skoe E, Russo NM, Dees T, Kraus N; Musical experience shapes human brainstem encoding of linguistic pitch patterns.; Nat Neurosci 2007 10 420-422

17: Schon D, Magne C, Besson M; The music of speech: music training facilitates pitch processing in both music and language.; Psychophysiology 2004 41 341-349

18: Fuller CD, Galvin III JJ, Maat B, Free RH, Başkent D; The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations.; Front Neurosci 2014 8 179

19: Moore BCJ, Carlyon RP; Perception of pitch by pwople with cochlear hearing loss and by cochlear implant users.; Pitch: Neural Coding and Perception NY Springer-Verlag 2005 234-27

20: Moore BCJ, Skrodzka E; Detection of frequency modulation by hearing-impaired listeners: effect of carrier frequency, modulation rate, and added amplitude modulation.; J Acoust Soc Am 2002 111 327-33;

21: Moore BCJ; Perceptual consequences of cochlear damage.; Oxford University Press 1995

22: Moore BCJ, Glasberg BR; Effects of the relative phase of the components on the pitch discrimination of complex tones by subjects with unilateral cochlear impairments.; Basic issues in hearing proceedings of the 8th international symposium on hearing 1988 421-430

23: Looi V, McDermott H, McKay C, Hickson L; Pitch discrimination and melody recognition by cochlear implant users.; Cochlear Implants. Proceedings of the VIII International Cochlear Implant Conference 2004 1273 197-200

24: Başkent D; 2012; Effect of speech on top-down repair: phonemic restoration with simulations of cochlear implants and combined electric-acoustic stimulation.; Journal of the association for research in otolaryngology 13(5) 683-692

25: Leek MR, Molis MR, Kubli LR, Tufts JB; Enjoyment of music by elderly hearing impaired listeners.; J Am Acad Audiol 2008 19 519-526

26: McDermott HJ; Music perception with cochlear implants: a review.; Trends Amplif 2004 8 49-82

27: Gfeller K, Witt S, Woodworth G, Mehr MA, Knutson J; Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal.; Ann Ottol Rhinol Laryngol 2002 111 349-356

28:Gfeller K, Olszewski C, Rychener M, Sena K, Knuston JF, Witt S; Recognition of "real-world" musical excerpts by cochlear implant recipients and normal-hearing adults.; Ear Hear 2005 26 237-250

29: Gfeller K, Lansing CR; Melodic, rhythmic and timbral perception of adult cochlear implant users.; J Speech Hear Res 1991 34(4) 916-920

30: Kong YY, Cruz R, Jones JA, Zeng FG; Music perception with temporal cues in acoustic and electric hearing.; Ear Hear 2004 25(2) 173-185