Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Customized selection of frequency maps in an acoustic simulation of a cochlear implant

Matthew B. Fitzgerald, Tasnim A. Morbiwala, and Mario A. Svirsky, *Senior Member, IEEE*

Abstract— **Cochlear implants can restore hearing to deaf individuals by electrically stimulating the auditory nerve. They do so by assigning different frequencies to different stimulating electrodes via a frequency map. We have developed a device that enables us to change the frequency map in real time. Here, in normal-hearing adults listening to an acoustic simulation of a cochlear implant, we investigate what frequency maps are initially preferred, and how the ability to understand speech with that preferred map compares with two other maps. We show that naïve listeners prefer a map that balances the need for low-frequency information with the desire for a naturally-sounding stimulus, and that initial performance with this listener-selected map is better than that with a map that distorts the signal to provide low-frequency information.**

I. Introduction

Among the most compelling advances in medicine over the past 30 years is the development of a prosthetic device known as a cochlear implant (CI). This device consists of an internal receiver and a series of electrodes that are surgically implanted into the inner ear, and an external part that includes a microphone, speech processor, and transmitter antenna. Sound is picked up by the microphone, processed, encoded, and sent to the receiver via a radiofrequency link. The receiver decodes the information and electrically stimulates auditory neurons, causing a hearing sensation. CI's restore hearing even in cases of profound deafness, and most CI users can understand speech even without lipreading [1].

One of the difficulties faced in CI design is how to mimic the function of the normally-functioning cochlea. A primary purpose of the cochlea is to serve as a frequency analyzer. Neurons at the base of the cochlea respond preferentially to frequencies as high as 20 kHz and neurons located progressively closer to the apex of the cochlea respond to progressively lower frequencies, down to about 20 Hz. This property is called tonotopicity [2].

Cochlear implants attempt to mimic this tonotopic organization by assigning different electrodes to different frequency bands. This frequency-to-place mapping is determined by the speech processor, and is generally called a *frequency map*. Complicating the assignment of the frequency map is that the average length of the cochlea is about 35 mm [3], while the typical insertion depth of a CI is only 20 mm [4]. Thus, when programming the frequency map of a postlingually deaf adult, (someone who lost hearing after learning an oral language), we are faced with a dilemma.

Does one remove the low-frequency information in an attempt to keep the frequency-to-electrode map as closely matched as possible to the frequency map the listener had prior to hearing loss? This processing strategy has been advocated for two reasons. First, speech-perception ability deteriorates both in implant users, and in normal-hearing individuals listening to an acoustic simulation of an implant, when the signal is frequency shifted beyond a certain amount [5], [6]. Second, it is difficult for postlingually deafened adults to adapt to substantial frequency-to-electrode shifts [7]. Unfortunately, the removal of low-frequency information may be suboptimal for long term use [8].

Alternatively, is it better to map low-frequency information to an electrode that will stimulate neurons normally tuned to high-frequency neural fibers? Such a processing strategy provides low-frequency information that is necessary for optimal speech understanding [9], but does so by shifting low-frequency information to neurons that are tuned to higher frequencies. It may take a postlingually deafened adult implant user years to fully adapt to this frequency shift [10]. However, once adult listeners have sufficiently adapted, their speech-perception scores may be significantly better than those obtained with frequency maps that are matched as closely as possible to the tonotopic array but lack some low frequency information [8].

We are interested in developing clinical tools that can help CI recipients adapt more quickly to their implant. Toward this goal, we have developed a real-time processor that enables an individual to adjust the frequency map in real time while they listen to running speech [11], [12]. One advantage of a real-time processor is that the listener can rapidly adjust the frequency map until he or she reaches a preferred map; such a process is not possible using currently available clinical software. Here, we used an acoustic simulation of a CI where noise bands of different frequencies were used to simulate the percepts that might be elicited by the stimulation of electrodes inside the cochlea. This is a relatively common procedure for pilot studies involving signal processing in CI's. We required naïve,

Manuscript submitted April 3, 2006. This work was supported by NIH/NIDCD grant DC03937.

Matthew B. Fitzgerald and Mario A. Svirsky are at the New York University School of Medicine, Department of Otolaryngology, 550 First Ave., NBV-5E5, New York, NY 10016. Phone number: 212-263-6346; fax number: 212-263-8256; email: matthew.fitzgerald@med.nyu.edu; Mario.svirsky@med.nyu.edu.

Tasnim A. Morbiwala was at Indiana University-Purdue University, Indianapolis, IN 46202 She is now with Advanced Bionics Corp., Sylmar, CA 91342 (e-mail: tmorbiwala@advancedbionics.com).

normal-hearing adults to listen to an acoustic simulation with fixed noise bands, and to select a preferred frequency map. After that we required them to perform a word-recognition task with this map, and with two other frequency maps. In doing so, we sought to answer two questions.

Our first question is how their preferred frequency map differs from 1) a "*right-place*" map: a map that matches the electrodes (or in the case of an acoustic model, the noise bands) as closely as possible to the analysis filter bank, but at the expense of failing to provide some low-frequency information, and 2) a map that provides both low-and high-frequency information, but must use a frequency shift to do so because the noise bands are higher in frequency than the range that is important for speech perception. This map is known here as a "*right-information*" map. This question provides us with some indication of what map a postlingually deafened adult implant user might prefer on his or her initial fitting after cochlear implantation. Our second question is whether word- and phoneme-recognition ability with their preferred frequency map differs from that obtained with the "*right-place*" and "*right-information*" frequency maps. This question provides some insight as to what sort of maps would provide the best performance for a newly-fit CI recipient.

II. Methods

A. Cochlear-implant simulation using an acoustic model

The signal processing for the acoustic model of the CI was adapted from the methods employed by Kaiser and Svirsky [11]. The signal was low-pass filtered with a cutoff frequency of 48,000 Hz. A bank of sixth order bandpass Butterworth filters was then used to divide the auditory signal into eight frequency channels. The cutoff frequencies of these filters could be changed in real time using the computer keyboard. For each filter, the temporal envelope was extracted and was used to modulate eight noise bands that spanned the range from 851 to 16,982 Hz.

B. Listeners

Fifteen listeners (ten of them female) between the ages of 23 and 44 (mean 28 years) participated in this experiment. All listeners were native speakers of English and had normal hearing. None of the listeners reported any history of hearing or learning disorder, nor did they have any prior experience listening to speech through an acoustic simulation of a CI.

C. Selection of listener maps

Listeners were seated in a sound-treated room, and were presented with running speech (taken from a book on tape) that was processed through our acoustic simulation of a CI. The listeners were instructed that the speech would sound somewhat distorted, and that they were to adjust it until it sounded best to them. By pressing keys on a keyboard, they could adjust two parameters of the filter bank: the average location, and the total frequency range. When the average location was changed, the logarithm of the frequency range remained constant and vice versa. This process was repeated three times with different starting points for the frequency map. One starting point was the frequency map traditionally used by CI recipients, while the second and third starting points were the lowest and highest possible frequency maps, respectively. The starting points were randomized across listeners. For an individual listener, we defined the *listener-selected* frequency map as the average of the three frequency maps that they selected.

D. Speech-recognition testing

After selecting their preferred frequency map, listeners were tested on their ability to understand speech processed through an acoustic simulation of a CI with three different frequency maps. The first map employed the *listener-selected* frequency map described above. The second frequency map ranged from 851 to 16982 Hz, the same exact frequency range used by the output noise bands. This map corresponds to a rough estimate of the preferred frequencies for the neurons that are presumably stimulated by the average insertion depth of a CI [4]. As such, it eliminates the low-frequency information below 851 Hz. This map is defined as the "*right-place*" frequency map. The final frequency map encompassed frequencies from 250 to 6800 Hz; this map corresponds to that commonly used with CI recipients because it provides the range that is believed to be most important for speech perception. It may, however, result in considerable basalward (low- to high-frequency) shift. This map is defined as the "*right-information*" map.

For each frequency map, listeners were tested on two 50-word CNC word lists. Listeners were instructed to repeat the target word, and their response was recorded by the experimenter prior to beginning the next trial. The number of words and the number of phonemes correctly identified using each frequency map were counted. The testing procedure was as follows: Listeners were presented with one CNC list for each frequency map; the presentation order of the different frequency maps was determined by a pseudo-randomized latin-square design. After completing testing with all three maps, the listeners were given a short break. Following the break, the order of the frequency maps was reversed, and listeners were re-tested on a second CNC list for each map. All word lists were presented at 70 dB SPL via speakers in a sound-treated booth.

III. Results

A. Listener-selected frequency maps

Our first question was how the *listener-selected* frequency map compares to a "*right-information*" frequency map, and to a "*right-place*" frequency map. Figure 1 shows the frequency ranges for the "*right-information*" map (far left column), for the "*right-place*" map (far right column), and for each of the fifteen *listener-selected* frequency maps (white columns in the middle). For all fifteen listeners, the lowest frequency in the *listener-selected* map (ranging from

481 to 817 Hz) always fell between the low-frequency edges of the "*right-information*" map (250 Hz) and the "*right-place*" map (851 Hz), although closer to the latter. These data indicate that naïve listeners who select their own frequency map shift the low-frequency edge somewhat basally relative to the "*right-place*" map, but that this shift is incomplete when compared to the clinically-used "*right-information*" map. Consistent with the idea of an incomplete basal shift, the low-frequency edges of both the "*right-information*" and the "*right-place*" maps (250 and 851 Hz, respectively) fell outside of the 95% confidence interval computed from the *listener-selected* frequency maps (575 to 682 Hz).

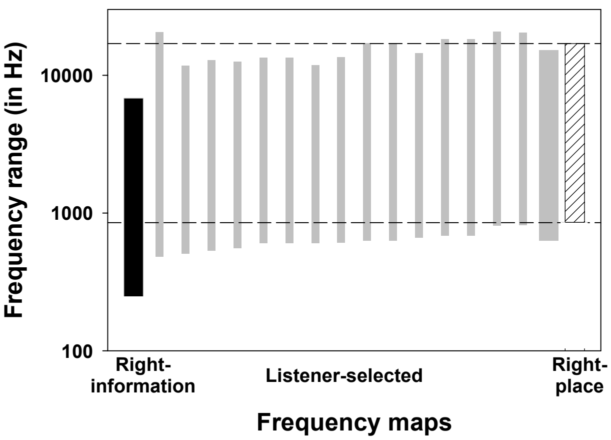

Fig. 1 Frequency maps are shown for the "*right-information*" (far left column), "*right-place*" (far right column) and the 15 individual *listener-selected* maps (white columns in the middle). Note that the *listener-selected* frequency maps fell between the "*right-information*" and "*right-place*" frequency maps, but closer to the latter..

The highest frequency in the *listener-selected* maps (white columns) ranged from 11684 to 20000 Hz. Ten of these values fell between the high-frequency edges of the "*right-information*" (left column, 6800 Hz) and "*right-place*" (right column, 16982) maps, while the remaining five (ranging from 17486 to 20000 Hz) were higher in frequency than the "*right-place*" map. In summary, the average location of the *listener-selected* maps fell within the "*right-information* and the "*right-place*" maps, but closer to the latter.

B. *Word-recognition performance with different frequency maps*

Our second question was whether word- and phoneme-recognition scores obtained with the *listener-selected* frequency map were different from those obtained with 1) the "*right-information*" map, and 2) the "*right-place*" map. To compare performance across these three frequency maps, we performed two one-way analyses of variance (ANOVA) with repeated measures across the listener-selected, "*right-information*," and "*right-place*" frequency maps. One ANOVA was done on the word-recognition scores, and the other on the phoneme-recognition scores. Each ANOVA was performed on the combined percent correct identifications from the two CNC word lists. The word-recognition percent-correct scores were determined by the number of words correctly identified / the total number of words presented. The phoneme-percent correct scores were determined in the same way except that they were computed on the number of phonemes presented in each word list (three per word).

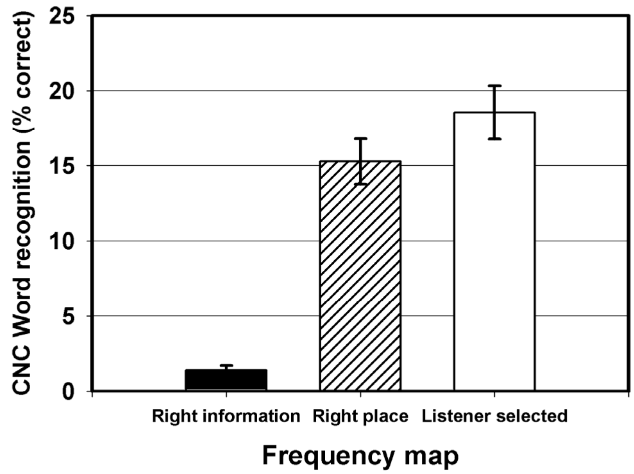

Fig. 2 Mean word-recognition scores are shown for the 15 listeners for each of the three frequency maps. Error bars indicate +/- one standard error of the mean across listeners. Performance with the *listener-selected* and "*right-place*" frequency maps was significantly better than with the "*right-information*" map.

Both ANOVAs revealed a significant main effect of frequency map (word-recognition scores: $F_{(2,14)} = 51.45$; $p<0.001$; phoneme-recognition scores: $F_{(2,14)} = 101.64$; $p<0.001$). While the word-recognition scores for the *listener-selected* map (Fig. 2, right column) were not significantly different from those obtained with the "*right-place*" map (middle column; Holm-Sidek test; $t=1.82$, $p=0.08$), they were significantly higher than those seen with the "*right-information*" frequency map (left column; $t=9.55$, $p<0.001$).

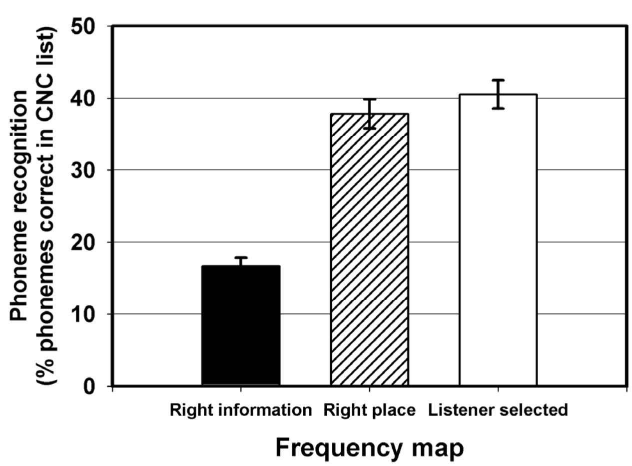

Fig. 3 Same as Fig. 2 except that the phoneme-recognition scores are plotted. Again, performance is better for the *listener-selected* and "*right-place*" frequency maps than for the "*right-information*" map.

A similar result was observed with the phoneme-recognition scores (see Fig. 3). Again, performance was equivalent for the *listener-selected* map and the "*right-place*" map (Holm-Sidek test; t=1.47, p=0.15). In contrast, the phoneme-recognition scores were significantly higher with the *listener-selected* map than with the "*right-information*" map (t=13.02, p<0.001). Taken together, these measures of speech-recognition ability indicate that, for naïve, normal-hearing adults who are listening to speech through an acoustic model of a CI, performance with a *listener-selected* map is equivalent to that seen with a "*right-place*" map, and is significantly better than that observed with a "*right-information*" map.

IV. Discussion

A. Characteristics of listener-selected frequency maps

The present data suggest that when naïve listeners listen to an acoustic simulation of a CI, they prefer to do so using a map that balances the need for the "*right information*" with their desire for the "*right place*." On average, both the low- and high-frequency edges of the listener-selected frequency map fell between those of the "*right-information*" and the "*right-place*" maps. This result suggests that listeners are trying to reconcile the need for low-frequency information with their own internal template for how speech should sound. Both post-lingually deafened adult implant users and normal-hearing adults who are listening to an acoustic model of a CI learned to understand speech by listening to a signal in which the place of excitation in the cochlea matches the neurons that respond preferentially to a given frequency. Therefore, a major shift of this frequency-to-place match may make speech less intelligible [5], [6]. However, these individuals are also used to having low-frequency information contained in the speech signal. For typical electrode insertion depths of a CI, low-frequencies may have to be represented via a basalward frequency-to-place shift. Finally, it is likely that this tradeoff is most pronounced for the low-frequency edge of the frequency map. Here, all 15 listeners had low-frequency edge values that fell between the *"right-information"* and *"right-place"* maps, while only 10 of 15 listeners displayed this pattern for the high-frequency edge.

B. Word-recognition performance with different frequency maps

Both the word- and phoneme-recognition results indicate that, for naïve listeners who listen to an acoustic model of an implant, optimal speech-perception performance is achieved with maps that have little to no frequency-to-place shifts. These data are consistent with previous results of cochlear implant simulation studies using naïve normal-hearing listeners, indicating that that substantial frequency-to-place shifts result in poor performance [5]. It is worth noting though that equivalent performance was observed for the listener selected and the "right-place" maps, even though the listener-selected map had a small amount of basalward frequency shift.

C. Implications for fitting frequency maps in cochlear implant users

These data suggest that cochlear-implant recipients could initially be fit with their individual preferred frequency map rather than the fitting software's default map [13]. Then, over time, the frequency map could be gradually shifted even further basally to provide more low-frequency information. We hypothesize that the initial use of the self selected map will facilitate speech perception shortly after initial fitting, and the gradual shift towards the right-information map will optimize long term speech perception performance. Finally, it is important to remember that the present results were obtained using an acoustic model of a cochlear implant. An important next step in this research will be to test the study's findings in a clinical population of cochlear implant users.

References

[1] Hochmair-Desoyer, I.J., Hochmair, E.S., Fischer, R.E., and Burian, K. "Cochlear prostheses in use: recent speech comprehension results," *Arch. Otorhinolaryngol*. vol 229, pp. 81-98, 1980

[2] von Bekesy, G. "*Experiments in Hearing*" E.G. Wever (ed. and trans.) New York: McGraw-Hill, 1960

[3] Hinojosa, R. and Marion, M. "Histopathology of profound sensorineural deafness," *Ann. NY Acad. Sci*., vol 405, pp. 459-484, 1983

[4] Ketten, D.R., Skinner, M.W., Wang, G., Vannier, M.W., Gates, G.A., and Neely, J.G. "In vivo measures of cochlear length and insertion depth of nucleus cochlear implant arrays," *Ann. Otol. Rhinol. Laryngol. Suppl*. Vol 175, pp. 1-16, 1998

[5] Dorman, M.F., Loizou, P.C., and Rainey, D. "Simulating the effect of cochlear-implant electrode insertion depth on speech understanding," *J. Acoust. Soc. Am*. vol 102, pp. 2993-2996, 1997

[6] Fu, Q.J., and Shannon, R.V. "Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant," *Ear Hear*. vol 20, pp. 332-344, 1999

[7] Fu, Q.J., Shannon, R.V., and Galvin, J.J. 3rd, "Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant," *J. Acoust. Soc. Am*. vol 112, pp. 1664-1674, 2002

[8] Faulkner, A., Rosen, S., and Norman, C. "The right information may matter more than frequency-place alignment: simulations of frequency-aligned and upward shifting cochlear implant processors for a shallow electrode array insertion," *Ear Hear*. vol 27, pp. 139-152, 2006

[9] French, N.R. and Steinberg, J.C. "Factors governing the intelligibility of speech," *J. Acoust. Soc. Am*. vol 19, pp. 90-119, 1947

[10] Tyler, R.S., Parkinson, A.J., Woodworth, G.G., Lowder, M.W., Gantz, B.J. "Performance over time of adult patients using the Ineraid or Nucleus cochlear implant," *J. Acoust. Soc. Am*. vol 102, pp. 508-522, 1997

[11] Kaiser A.R., and Svirsky, M.A. "Using a personal computer to perform real-time signal processing in cochlear implant research," *Proceedings of the IXth IEEE-DSP Workshop*, October 15-18, 2000

[12] Morbiwala, T.A., Svirsky, M.A., El-Sharkway, M., and Rizkalla, M. "A PC-based speech processor for cochlear implant fitting that can be adjusted in real-time," *Proceedings of the 48th IEEE Midwest Symposium on Circuits and Systems*, August 7-10, 2005

[13] Svirsky, M.A. "Gradual adaptation to a modified peripheral frequency map by post–lingually deaf cochlear implant users," *Proceedings of the 2005 Conference on Implantable Auditory Prostheses,* July 30-August 4, 2005.