VALIDATION OF ACOUSTIC MODELS OF AUDITORY NEURAL PROSTHESES

Mario A. Svirsky, Nai Ding, Elad Sagi, Chin-Tuan Tan, Matthew Fitzgerald, E. Katelyn Glassman, Keena Seward and Arlene C. Neuman.

Department of Otolaryngology, NYU School of Medicine, New York, USA

ABSTRACT

Acoustic models have been used in numerous studies over the past thirty years to simulate the percepts elicited by auditory neural prostheses. In these acoustic models, incoming signals are processed the same way as in a cochlear implant speech processor. The percepts that would be caused by electrical stimulation in a real cochlear implant are simulated by modulating the amplitude of either noise bands or sinusoids. Despite their practical usefulness these acoustic models have never been convincingly validated. This study presents a tool to conduct such validation using subjects who have a cochlear implant in one ear and have near perfect hearing in the other ear, allowing for the first time a direct perceptual comparison of the output of acoustic models to the stimulation provided by a cochlear implant.

Index Terms— Cochlear implants, Acoustic models, Sensory aids, Speech processing, Single sided deafness

1. INTRODUCTION

Cochlear implants (CIs) are sensory aids usually intended for patients whose hearing loss is at least severe. Unlike hearing aids, which stimulate the ear acoustically, CIs stimulate the auditory nerve directly using electrical pulses delivered to an electrode array placed in the inner ear. These devices represent the first successful attempt to replace a human sense with an electronic prosthesis.

Over the past thirty years [1], acoustic models have been used in numerous studies (e.g., [2]-[18]) to simulate the percepts elicited by auditory neural prostheses such as CIs. These models are extremely useful for experiments that cannot be done or would be impractical with actual CI users, and are also useful as informative control conditions in CI experiments. Acoustic models use the same front end signal processing as a CI, which includes a frequency analysis step that is carried out using a set of bandpass filters or an FFT algorithm. However, the percepts that would be caused by electrical stimulation in a real CI are simulated by modulating the amplitudes of either noise bands or sinusoids. These noise bands or sinusoids span a range of frequencies in an attempt to simulate percepts with low pitch (corresponding to electrodes that are more deeply inserted into the cochlea) or with higher pitch (electrodes that are more shallowly inserted). With few exceptions (e.g., [4]) acoustic models have been implemented using a set of noise bands that are an exact match (in both number of bands and the cutoffs of each band) to the analysis filters used in the front end, or using a set of sinusoids that are identical to the center frequencies of the analysis filters. Here we refer to this type of model as the “standard” acoustic model.

1.1. Relation to Prior Work

Models in general and acoustic models in particular can be extremely useful, but only to the extent that they are valid. Validating acoustic models of CIs is complicated, however, because it is impossible to know what the percepts are that result from stimulation of a given intracochlear electrode in a given CI user. In fact, due to individual variations in neural survival, cochlear size, and electrode insertion depth, it seems very likely that stimulation of a given electrode can sound very different to different CI users. Fortunately, there is a new group of CI patients (those with single-sided deafness or SSD, [19]-[21]) who can greatly facilitate the validation of CI acoustic models. These are patients who have a CI in one ear and have very good hearing in the other ear, allowing the direct comparison of two signals: a speech signal presented to the implant through a direct connection to the speech processor, followed by an acoustic model version of the same speech signal presented to the acoustic hearing ear.

In previous work [22], [23], [24] we developed a tool that allowed us to process stimuli using a variety of analysis filter banks, with the output consisting of either electrical stimulation to intracochlear electrodes, or an acoustic model using fixed noise bands. The purpose of that tool was to facilitate the selection of the analysis filter bank that would result in maximum speech intelligibility for a CI user, or for a normal hearing subject listening to an acoustic model of a CI. In the present study we used a new tool with a similar graphic interface and similar signal processing but also one important difference: instead of using variable analysis filters and fixed noise bands, here we used a fixed set of analysis filters and a variable set of noise bands (or a variable set of sinewaves). The fixed set of analysis filters is the same one used by a given SSD subject in his/her clinical processor, thus guaranteeing that the front end processing is as similar as possible between the CI and the acoustic model. The variable set of noise bands (or sinewaves) is used by the subject to select the acoustic model that provides the best match to the CI.

2. METHODS

Two test sentences were processed using CI acoustic models which had a number of similarities. In all cases, stimuli were digitized at 48,000 samples per second and lowpass filtered at 20,000 Hz. A bank of sixth order bandpass Butterworth filters was then used to divide the acoustic signal into twenty-two frequency channels. The cutoff frequencies were identical to those in the standard frequency table used by the vast majority of Nucleus Freedom™ CI users (188 Hz to 7,938 Hz), and also used by the individual with SSD who participated in this study. For each analysis filter, the temporal envelope was extracted by half wave rectification and third order Butterworth low pass filtering at 100 Hz. The temporal envelopes were then used to modulate either a set of twenty-two noise bands or twenty-two sinewaves. The acoustic models differed in the specific noise bands (or sinewaves) that were used. The noise bands were adjacent in frequency, and they spanned a frequency range whose minimum could assume one of thirteen possible values (ranging from 63 Hz to 1,813 Hz) and whose maximum could assume one of nine possible values (ranging from 3,372 Hz to 18,938 Hz). Thus, 117 acoustic models were implemented whose noise bands spanned different frequency ranges. Likewise, 117 acoustic models were implemented using sinewaves instead of noise bands. The frequencies of the sinewaves were equal to the center frequencies of the corresponding noise bands. The frequency ranges were chosen to ensure that they would cover the highest and lowest pitch percepts normally heard by CI users.

A MATLAB program was developed to handle stimulus presentation and record subject responses. When the program starts, the experimenter selects the sentence to be used, the type of model to be used (noise bands or sinewaves), and the orientation of the x- and y-axes for the grid shown in Figure 1. That grid is the graphical user interface that is employed by the subject to assess different acoustic models. In the 9 by 13 grid each square represents one of the 117 acoustic models. The low frequency edge of the set of noise bands (or sinewaves) is determined by the vertical position in the grid, and the high frequency edge is determined by the horizontal position. "Standard" acoustic models are those where the center frequencies of the analysis filters are equal to those of the noise bands (or sinewaves), and they are indicated by an asterisk located at the intersection of the second row and the fifth column in Figure 1. When the subject clicks on a given square, he hears a sequence of two stimuli. The first one is a signal that is presented to his/her CI processor over a direct connection (with the microphone deactivated), and the second one is the same signal processed with the acoustic model that corresponds to the selected square, and presented through a loudspeaker. Because subjects are totally deaf in the implanted ear and the speech processor microphone is deactivated, they only hear the acoustic signal (processed by the acoustic model) in the acoustically stimulated ear, and the electrical signal (processed by the CI's speech processor) in the implanted ear. The subject's task is to assess the similarity between the two stimuli and assign a rating that is shown using a color code in the corresponding square. As indicated in Figure 1, red represents the worst match (i.e., the acoustic model sounds very different from the CI) and blue represents the best match.

Once the subject has selected a noise-band acoustic model and a sinewave acoustic model that sound most similar to the CI, he fills out a questionnaire comparing the CI and four acoustic models: the standard/noise-band acoustic model, the self-selected/noise-band model, the standard/tone model, and the self-selected/tone model. The question to be answered is: "The sound I hear through my unimplanted ear is [***] to the sound I hear through my implant" where [***} can assume values ranging from 1-not at all similar (completely different) to 9-identical.

The third and last part of the evaluation consists on speech testing. Two 50-word lists of the CNC word identification test were presented under each condition. Presentation to the CI ear was done with the direct connection described above and the input gain adjusted to a comfortable level. The acoustic stimuli were presented via loudspeaker at 70 dB SPL-C, with the speech processor of the CI turned off.

One SSD subject was recruited to pilot the use of the tools described above. He was a unilaterally, postlingually deaf 37 year old male with normal hearing in his right ear and a profound hearing loss in the left (implanted) ear due to a severe ear infection one year prior to implantation. He received a Freedom™ CI that was programmed with the ACE strategy and the standard (188-7938 Hz) frequency table. Selection of the best acoustic models and questionnaire completion were done 6.5 months after initial activation, and speech testing was done two weeks later.

3. RESULTS

Figure 1 shows the models that were selected by the listener. The top panel shows results for the noise-band models and the bottom panel shows results for the sinewave (or "tone") models. In both panels, the black asterisk at the intersection of the second row and the fifth column indicates the standard model where the frequency range of the noise bands or sinewaves is identical to the frequency range of the map in the patient's speech processor. As can be observed in the top panel, the acoustic model that was the best match to the CI had noise bands that ranged from 188 Hz to 14,924 Hz (intersection of the second row and the eighth column). When using the sinewave acoustic models (bottom panel) the standard model was the best match to the CI, but the similarity rating was not as high as that achieved when using the self-selected noise-band acoustic model.

Table 1 shows the similarity ratings between the CI and each one of the acoustic models. All responses fell between 5 (somewhat similar) and 7 (very similar). In the case of the sinewave (tone) acoustic models, the self-selected model coincided with the standard one so the ratings are identical. Consistent with the ratings obtained during the acoustic model selection process (Figure 1), the best results were obtained with the self-selected noise-band model.

Lastly, the word recognition scores (shown in Figure 2) also point to the self-selected noise-band acoustic model as being the one that provided the best match to the CI, at least for the individual who was tested in this pilot study. In terms of speech perception scores all acoustic models resulted in overestimates of the real scores obtained with the CI. These overestimates were very large with the sinewave model (95% vs 42% correct). The overestimate obtained with the self-selected/noise-band model was the smallest one but it was still significant (67% vs 42%). This is consistent with informal comments made by subjects indicating that the sinewave models and the standard noise-band model actually sounded "better" and "more intelligible" than the implant itself.

Ratings	NOISE		TONE	
	Std	SS	Std	SS
Sentence 1	5	7	5	5
Sentence 2	6	7	7	7
Average	5.5	7	6	6

Table 1: Ratings indicating perceived similarity between sounds heard through the CI ear and sounds heard using each one of the acoustic models. The self-selected/noise-band model was the only one that was rated 7 (i.e, "very similar" to the CI) for both of the sentences that were used.

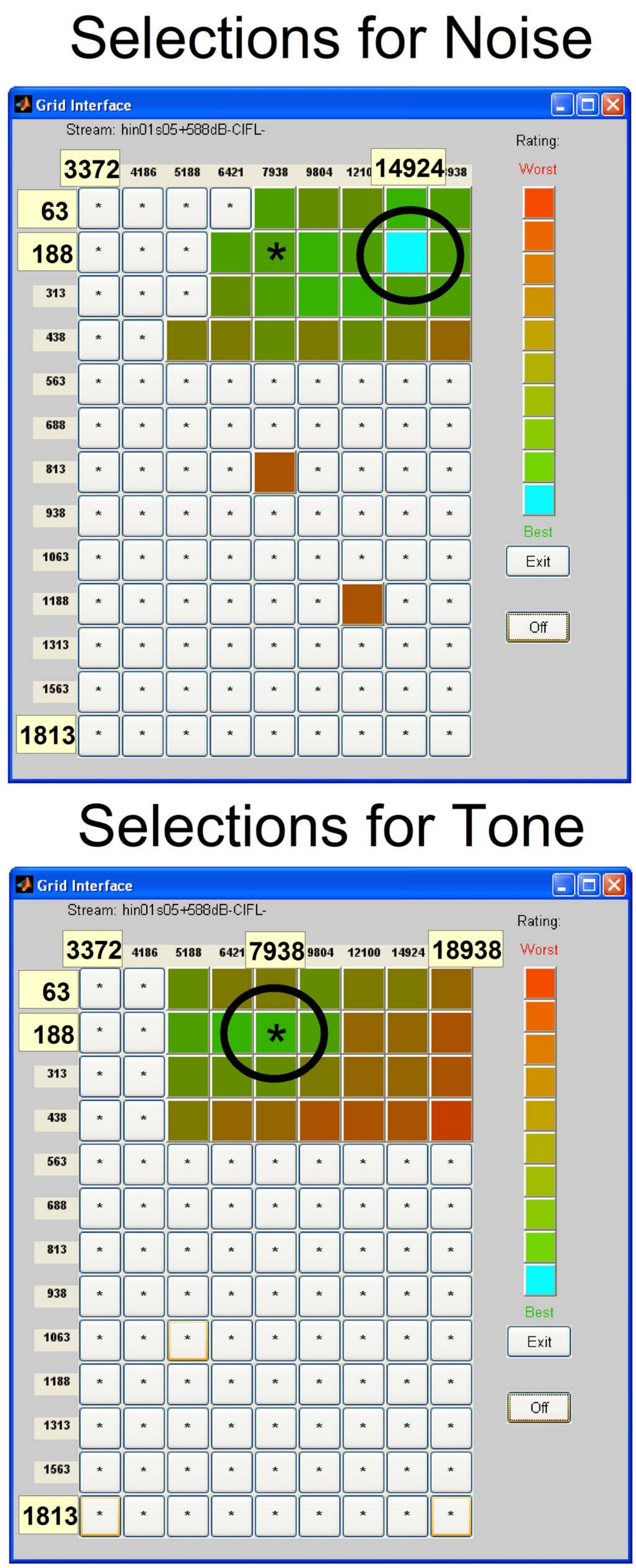

Figure 1:GUI used to select the acoustic models that sounded most similar to the CI. Color indicates similarity between CI and acoustic model (blue-most similar).

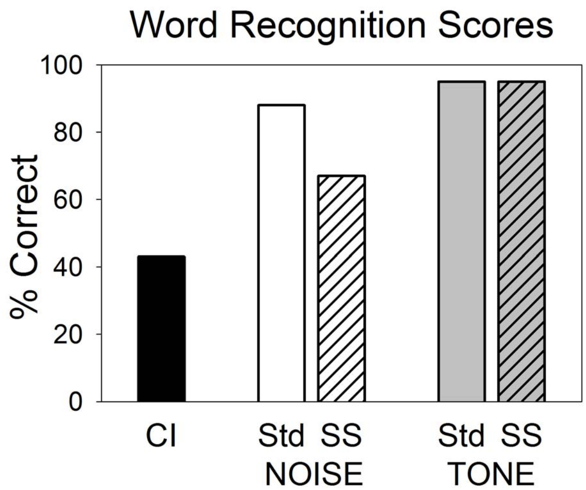

Figure 2: CNC word identification scores obtained by a CI user with single sided deafness using his CI or one of four acoustic models (standard/noise-band, self-selected/noise-band, standard/tone and self-selected/tone). All acoustic models resulted in overestimates of word identification scores. The self-selected/noise-band model provided the best fit (smallest overestimate).

4. CONCLUSIONS

We have developed a tool that can be successfully used with single-sided deafness CI users to validate acoustic models of neural auditory prostheses. For the individual whose data are shown here as a case study the best acoustic model was the self-selected/noise-band model. The noise bands in this model differed from those used in the majority of published studies. All models in the present study overestimated the speech perception score actually obtained with the CI ear, and the smallest overestimate was obtained with the self-selected/noise-band model. This model was also the only one that was rated as "very similar" to the sound provided by the CI. The other models were less similar because they sounded "better" and "more intelligible" than the CI actually does.

In this study we optimized the noise bands used in acoustic models, but there are other parameters whose manipulation may result in even more accurate models. In particular, the bandwidth of the noise bands was used by Nogaki and Fu [18] to simulate different degrees of spread of excitation in the cochlea.

Results from this case study suggest that great caution must be used when extrapolating results of acoustic model studies to actual CI users. Depending on the specific acoustic model that is used, such extrapolations may not be warranted. The tool that is being introduced by the present study may help develop a new, more accurate generation of CI acoustic models.

5. ACKNOWLEDGEMENTS

The work described in these pages was supported by the following NIH-NIDCD grants, R01-DC03937 (PI: Svirsky), R01-DC011329 (PIs: Svirsky and Neuman), K99-DC009459 (PI: Fitzgerald), K25DC010834 (PI: Tan), and by Cochlear Americas (PI: J. Thomas Roland, Jr). We appreciate the support of Susan Waltzman and J. Thomas Roland, Jr., co-directors of the NYU Cochlear Implant Center, and Jacquelyn Schmit-Hoffman for her help with subject recruitment.

6. REFERENCES

[1] P. J. Blamey, R. C. Dowell, and Y.C. Tong, "Speech Processing Studies Using an Acoustic Model of a Multiple-Channel Cochlear Implant," Journal of the Acoustical Society of America, vol. 76 (1), pp. 104-110, 1984.

[2] R. V. Shannon, F. G. Zeng, and V. Kamath, "Speech Recognition with Primarily Temporal Cues," Science, vol. 270 (5234), pp. 303-304, 1995.

[3] M. F. Dorman, P. C. Loizou, and D. Rainey, "Speech Intelligibility as a Function of the Number of Channels of Stimulation for Signal Processors using Sine-Wave and Noise-Band Outputs," Journal of the Acoustical Society of America, vol. 102 (3), pp. 2403-2411, 1997.

[4] M. F. Dorman, P. C. Loizou, and D. Rainey, "Simulating the Effect of Cochlear-Implant Electrode Insertion Depth on Speech Understanding," Journal of the Acoustical Society of America, vol. 102 (5), pp. 2993-2996, 1997.

[5] M. F. Dorman, and P. C. Loizou, "Speech Intelligibility as a Function of the Number of Channels of Stimulation for Normal-Hearing Listeners and Patients with Cochlear Implants," American Journal of Otology, vol. 18 (6, S), pp S113-S114, 1997.

[6] M. F. Dorman, and P. C. Loizou, "The Identification of Consonants and Vowels by Cochlear Implant Patients using a 6-Channel Continuous Interleaved Sampling Processor and by Normal-Hearing Subjects using Simulations of Processors with Two to Nine Channels," Ear and Hearing, vol. 19 (2), pp 162-166, 1998.

[7] P. C. Loizou, M. Dorman, and Z. M. Tu, "On the Number of Channels Needed to Understand Speech," Journal of the Acoustical Society of America, vol. 106 (4), pp. 2097-2103, 1999.

[8] P. C. Loizou, M. Dorman, and J. Fitzke, "The Effect of Reduced Dynamic Range of Speech Understanding: Implication for Patients with Cochlear Implants," Ear and Hearing, vol. 21 (1), pp. 25-31, 2000.

[9] P. C. Loizou, M. Dorman, and O. Poroy, "Speech Recognition by Normal-Hearing and Cochlear Implant Listeners as a Function of Intensity Resolution," Journal of the Acoustical Society of America, vol. 108 (5), pp. 2377-2387, 2000.

[10] M.F. Dorman, P.C. Loizou, L.L. Kemp, and K.I. Kirk, "Word Recognition by Children Listening to Speech Processed into a Small Number of Channels: Data from Normal-Hearing Children and Children with Cochlear Implants," Ear and Hearing, vol. 21 (6), pp. 590-596, 2000.

[11] K. Kasturi, P.C. Loizou, M. Dorman, and T. Spahr, "The Intelligibility of Speech with "Holes" in the Spectrum," Journal of the Acoustical Society of America, vol. 112 (3), pp. 1102-1111, 2002.

[12] M.F. Dorman, P.C. Loizou, A. Spahr, and C.J. Dana, "Simulations of Combined Acoustic/Electric Hearing," Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1999-2001, 2003.

[13] M.F. Dorman, A.J., Spahr, P.C. Loizou, C.J., Dana, and J.S. Schmidt, Acoustic Simulations of Combined Electric and Acoustic Hearing (EAS)," Ear and Hearing, vol. 26 (4), pp. 371-380, 2005.

[14] K. Kasturi and P.C. Loizou, "Effect of Filter Spacing on Melody Recognition: Acoustic and Electric Hearing," Journal of the Acoustical Society of America, vol. 122 (2), pp. EL29-EL34, 2007.

[15] N. Li and P.C., Loizou, "A Glimpsing Account for the Benefit of Simulated Combined Acoustic and Electric Hearing," Journal of the Acoustical Society of America, vol. 123 (4), pp. 2287-2294, 2008.

[16] W.B. Cooper, E. Tobey, and P.C. Loizou, "Music Perception by Cochlear Implant and Normal Hearing Listeners as Measured by the Montreal Battery for Evaluation of Amusia," Ear and Hearing, vol. 29 (4), pp. 618-626, 2008.

[17] N. Li and P.C. Loizou, "Factors Affecting Masking Release in Cochlear-Implant Vocoded Speech," Journal of the Acoustical Society of America, vol. 126 (1), pp. 338-346, 2009.

[18] Fu, Q.J. and Nogaki, G. "Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing," J Assoc Res Otolaryngol, vol. 6(1), pp. 19-27, 2005.

[19] S. Arndt, R. Laszig, A. Aschendorff, R. Beck, C. Schild, F. Hassepass, G. Ihorst, S. Kroeger, P. Kirchem, and T. Wesarg, "Unilateral Deafness and Cochlear Implantation: Audiological Diagnostic Evaluation and Outcomes," HNO, vol. 59 (5), pp. 437-446, 2011.

[20] A. K. Punter, K. Vermeire, A. Hofkens, M. De Bodt, D. De Ridder, and P. Van De Heyning, "Cochlear Implantation as a Durable Tinnitus Treatment in Single-sided Deafness," Cochlear Implants International, 12 (S 1), pp. S26-29, 2011.

[21] R. A. Arts, E. L. George, R. J. Stokroos, K. Vermeire, "Review: Cochlear Implants as a Treatment of Tinnitus in Single-Sided Deafness," Current Opinion in Otolaryngology and Head and Neck Surgery, vol. 20 (5), pp. 398-403, 2012.

[22] D. Jethanamest, C. T. Tan, M. B. Fitzgerald, and M. A. Svirsky, "A new software tool to optimize frequency table selection for cochlear implants," Otol Neurotol, vol. 31 (8), pp. 1242-1247, 2010.

[23] M. B. Fitzgerald, T. A. Morbiwala, and M. A. Svirsky, "Customized Selection of Frequency Maps in an Acoustic Simulation of a Cochlear Implant," Proceedings of the IEEE 2006 International Conference of the Engineering in Medicine and Biology Society, pp. 3596-3599, 2006.

[24] M. B. Fitzgerald, E. Sagi, T. A. Morbiwala, C. T. Tan, and M. A. Svirsky, "Feasibility of Real-Time Selection of Frequency Tables in an Acoustic Simulation of a Cochlear Implant," Ear and Hearing, submitted.