

Eighth Quarterly Progress Report

January 1 through March 31, 2004
NIH Project N01-DC-2-1002

Speech Processors for Auditory Prostheses

Prepared by

Blake S. Wilson, Xiaoan Sun, Reinhold Schatzer, and Robert D. Wolford

Center for Auditory Prosthesis Research
RTI International
Research Triangle Park, NC

CONTENTS

I. Introduction and review of principal activities for the quarter.....	3
II. Plans for the next quarter.....	5
III. Acknowledgments	6
Appendix 1: Summary of reporting activity for this quarter	7
Appendix 2: Representation of fine structure or fine frequency information with cochlear implants	8

I. Introduction

The main objective of this project is to design, develop, and evaluate speech processors for implantable auditory prostheses. Ideally, such processors will represent the information content of speech in a way that can be perceived and utilized by implant patients. An additional objective is to record responses of the auditory nerve to a variety of electrical stimuli in studies with patients. Results from such recordings can provide important information on the physiological function of the nerve, on an electrode-by-electrode basis, and can be used to evaluate the ability of speech processing strategies to produce desired spatial or temporal patterns of neural activity.

Work and activities in this quarter included:

- A two-week visit by new subject ME-23, using a Med-El cochlear implant on one side and having substantial residual acoustic hearing on both sides, January 5-16.
- A visit by investigator Artur Lorens from Warsaw, Poland, to collaborate in studies with ME-23.
- Studies with new local subject ME-24, who has bilateral Med-El implants, January 19-20.
- Initial studies with subject NP-8, the second in a series of four patients implanted with an experimental version of the Nucleus device, with a Contour electrode array and a percutaneous connector, February 2-3.
- An invited presentation by Blake Wilson at the *II Meeting Consensus on Auditory Implants*, in Valencia, Spain, February 22.
- Further studies with subject NP-8 the week of March 1-5, to evaluate new processing strategies using dual-resonance, non-linear filters, designed to reproduce aspects of non linear responses and tuning at the basilar membrane and associated structures in normal hearing (see the prior Quarterly Progress Reports 6 and 7 for this project).
- A visit by consultant Enrique Lopez-Poveda to collaborate in studies with NP-8.
- Studies with Nucleus Contour Electrode percutaneous subject NP-7, March 10-12.

The studies with subject ME-23 in particular increased our knowledge about the efficacy of combined electric and acoustic stimulation (EAS) of the auditory system and also afforded the possibility to compare pitches evoked by electric stimuli with those evoked by acoustic stimuli. In general, the pitches evoked by electric stimuli are matched with frequencies of acoustic sinusoids that are far lower than would be predicted by Greenwood's map.

Manuscripts are in preparation to describe our findings from studies with six EAS subjects, and from the combined findings from our studies at RTI and studies conducted at the J. W. Goethe Universität in Frankfurt, Germany, at the University Hospital Clinic in Vienna, Austria, and at the International Center of Hearing and Speech in Kajetany (near Warsaw), Poland. In addition, a separate manuscript is in preparation to describe findings from pitch scaling and matching studies conducted with three subjects in our laboratories and with additional subjects at collaborating institutions (the University of

Iowa and the International Center of Hearing and Speech). These manuscripts should be ready for submission soon, and, when ready, will be included as appendices to future reports.

In addition to these activities, work is underway in our laboratories and elsewhere to represent "fine structure" or "fine frequency" information with implants in a way that the information can be perceived and utilized by patients. The proof pages for a paper in press describing some of this work are presented in Appendix 2 to this report. Results from the other studies mentioned in the list above will be presented in future reports.

II. Plans for the next quarter

Among the activities planned for the next quarter are:

- A one week return visit by subject ME-21, May 3-7.
- Special Guest Address by Blake Wilson at the *VIIIth International Cochlear Implant Conference*, to be held in Indianapolis, Indiana, May 11-14.
- Honorary Speaker Presentation by Blake Wilson at the Med-El Satellite Meeting, *VIIIth International Cochlear Implant Conference*, Indianapolis, IN, May 11–14.
- A visit by Peter Nopp following the Indianapolis conference.
- A return visit for continued studies with Nucleus Contour Electrode percutaneous subject NP-8, May 24-25.
- A return visit for continued studies with Nucleus Contour Electrode percutaneous subject NP-6, June 7-8.
- Initial studies with Nucleus Contour Electrode Percutaneous subjects NP-7 and NP-9.
- Additional studies with local subjects.

III. Acknowledgments

We thank volunteer research subjects ME-23, ME-24, NP-7, and NP-8, who participated in the studies conducted during this quarter.

Appendix 1: Summary of reporting activity for this quarter

Reporting activity for this quarter, covering the period of January 1 through March 31, 2004, included the following:

Invited presentation

Wilson BS, Cartee LA, Cox JH, Lawson DT, Lopez-Poveda E, Schatzer R, Sun X, Wolford RD: Future developments of CI. *II Meeting Consensus on Auditory Implants*, Valencia, Spain, February 19-21, 2004.

Publication

Loeb GE, Wilson BS: Cochlear prosthesis. In Adelman G, Smith BH (eds.), *Encyclopedia of Neuroscience*, 3rd edition. Amsterdam: Elsevier, 2004. (The *Encyclopedia* is available both in book and CD-ROM formats.)

Appendix 2: Representation of fine structure or fine frequency information with cochlear implants

This appendix presents the page proofs for a paper that will be published in the Elsevier International Congress Series, volume 1273.



ELSEVIER

International Congress Series xx (2004) xxx–xxx



www.ics-elsevier.com

1

Representation of fine structure or fine frequency information with cochlear implants

B.S. Wilson*, X. Sun, R. Schatzer, R.D. Wolford

*Center for Auditory Prosthesis Research, RTI International, 3040 Cornwallis Road, Research Triangle Park,
NC 27709, United States*

Abstract. One possible approach for improving cochlear implants is to represent “fine structure” or “fine frequency” information in a way that it can be perceived and utilized by patients. The importance of the information and various ways of representing it are described. © 2004 Published by Elsevier B.V.

Keywords: Cochlear implants; Speech perception; Music perception; Speech processing

1. Introduction

This is an exciting time in the history of cochlear implants. Significant improvements in performance have been demonstrated recently with bilateral implants and with combined electric and acoustic stimulation (EAS) of the auditory system, the latter for patients with some residual (low frequency) hearing. Additional approaches are being developed. Work is underway in our laboratories and elsewhere to (a) provide a closer mimicking than was heretofore possible of the signal processing that occurs in the normal cochlea, and (b) represent frequency variations within channels in a way that they can be perceived. These new approaches may support further gains in performance, either alone or in combination with other approaches.

The principal purpose of this short paper is to describe approach (b) above. A detailed description of approach (a) is presented in Ref. [1].

2. Importance of “fine structure” or “fine frequency” information

The mathematician David Hilbert showed that signals can be decomposed into a slowly varying envelope that modulates a high-frequency carrier. The instantaneous phase, or

* Corresponding author. Tel.: +1 919 541 6974; fax: +1 919 990 8385.

E-mail address: bsw@rti.org (B.S. Wilson).

frequency (first derivative of the phase signal), of the carrier varies continuously. Hilbert described the carrier as the “fine structure” (FS) portion of the original signal.

Smith [2] have investigated the relative importance of envelope and FS information for speech reception, melody recognition, and sound localization. They created “auditory chimeras” by first processing two separate inputs with identical banks of band-pass filters and then multiplying the FS carriers derived from one bank of filters with the envelope signals derived from the other bank of filters. The modulated carriers were then summed to form the output. Thus, the chimeras presented conflicting cues—the envelope variations in a given number of bands for one sound, versus the FS variations in the same bands for another sound. Pairings of inputs included sentences versus noise, sentences versus different sentences, melodies versus different melodies, and sentences with an interaural time delay (ITD) corresponding to a sound image at the left versus the same or different sentences with an ITD corresponding to a sound image at the right.

The sound heard or correctly identified by subjects with normal hearing depended on the sounds in each pairing and on the number of processing channels (bands). Speech was identified by its envelope information for eight or more channels, whereas the FS information was more important for one or two channels. Both envelope and FS information contributed to sentence recognition for intermediate numbers of channels. Melodies were recognized almost exclusively by their FS information up to 32 channels. Envelope cues became dominant at 48 and 64 channels. Lateralization of sentences was difficult with a small number of channels, but improved with increasing numbers up to the tested limit of 32. Lateralization was cued by the FS information in all cases.

These findings indicate the importance of the FS information for speech reception using fewer than about eight processing channels, and for music reception using fewer than about 40 channels.

Present-day electrode arrays for cochlear implants appear to support no more than 4–8 independent channels, as indicated by a lack of increases in speech reception scores when the number of processing channels and associated sites of stimulation is increased beyond that number (e.g., Ref. [3]). In this 4–8 range, both envelope and FS information contribute strongly to speech recognition. Music is conveyed almost solely by FS cues.

The importance of the FS information seems indisputable, given the maximum number of effective channels with current implant devices. The question is: how can this information be presented in a way that it can be perceived by the patient?

3. Present processing strategies for implants

The CIS, ACE, SPEAK, *n-of-m*, and other processing strategies now in use for cochlear implants [4] extract envelope signals from contiguous band-pass filters, that span the overall frequency range of speech and other inputs. The envelope signals are used to determine patterns of stimulation in the implant. Thus, only the processed envelope information is presented to the user, and most or all FS information is discarded at the envelope-extraction stage.

In CIS processors, for instance, envelope signals are derived at the outputs of the band-pass filters and those derived signals are compressed into the narrow dynamic range of electrically evoked hearing. The compressed signals modulate trains of pulses with a

constant pulse rate. Frequency variations of signals within each band-pass are not represented in the stimuli unless (a) the cutoff frequency of the low-pass filter in the envelope detector is comparable to or higher than frequencies in the band-pass, (b) a half-wave rectifier is used instead of a full-wave rectifier in the envelope detector to avoid frequency-doubling effects, and (c) the pulse rate for the channel is sufficiently high to represent the relatively rapid variations in the modulation without significant distortions. Typical implementations of the CIS and other processors use cutoff frequencies for the envelope detectors in the range of 200–400 Hz. This allows representations of the fundamental frequency for voiced speech sounds, voiced versus unvoiced distinctions, and rapid transient events in speech such as those associated with stop consonants. However, little or no information about frequencies within channels, even the channel with the lowest center frequency, is represented.

A variation of CIS, called “HiRes,” uses high pulse rates and an effective cutoff for the envelope detectors at half the rates [4,5]. In addition, the detectors use a half-wave rectifier. The rates may be as high as 2800 pulses/s for each channel and associated electrode, for a 16-channel implementation that uses non-simultaneous stimulation across electrodes. The rate can be almost doubled using an implementation that presents pulses for two of the channels to their respective electrodes simultaneously. FS information may be presented for frequencies up to about 1400 Hz for the non-simultaneous mode and up to about 2800 Hz for the “paired pulses” mode.

Another way in which FS information might be presented is to represent directly the “analog” outputs of the band-pass filters at the electrodes, as in the *compressed analog* (CA) or *simultaneous analog stimulation* (SAS) strategies [4]. FS information is not discarded in the processing with these strategies.

Although FS information may be presented with the HiRes, CA, or SAS strategies, implant patients may not be able to perceive much if any of it. In particular, most patients do not perceive differences in the frequency of stimulation at individual electrodes as differences in pitch above a “pitch saturation limit” of about 300 Hz (e.g., Ref. [6]). Thus, frequency variations may be presented by these strategies but they cannot be utilized by the patients for any but the lowest frequencies. For loudness-balanced stimuli, a sinusoid at 500 Hz does not sound any different to most patients than a 300-Hz sinusoid. Similarly, modulation of a pulse train at 500 Hz is not discriminable from modulation at 300 Hz.

A further concern with strategies that use simultaneous stimulation across electrodes (CA, SAS, and the “paired pulses” version of HiRes) is that such stimulation may exacerbate interactions among electrodes and thereby reduce the salience of channel-related cues [4,7]. Thus, any gain produced through presentation of FS information may be counteracted, or more than counteracted, by an increase in electrode interactions.

This possible tradeoff between a representation of FS information on the one hand, and increased electrode interactions on the other hand, may vary across patients. Some patients have relatively high pitch saturation limits and thus may have greater access than others to the presented FS information. In addition, some patients have relatively low electrode interactions, perhaps due to excellent survival of neural elements in the implanted cochlea or close placements of electrodes next to excitable tissue or both. Patients with both attributes may achieve especially good results with the CA, SAS, or “paired pulses” HiRes strategies. Patients with high pitch saturation limits may receive the greatest benefit from

the non-simultaneous pulses version of HiRes or other implementations of CIS that use high cutoff frequencies for the envelope detectors.

In general, though, representation of FS information as variations in frequencies or rates of stimulation at the electrodes seems limited. Most patients will not have any access to the information above about 300 Hz, and no patient will have access to it above about 1000 Hz. In addition, the difference limens for frequencies below the limit are very much poorer for implant patients than for listeners with normal hearing, usually more than 10 times worse. Thus, even for frequencies below the pitch saturation limit, patients may be able to perceive only gross features in the presented FS information.

The typical pitch saturation limit of 300 Hz just reaches or barely exceeds the lower cutoff frequency for the band-pass filter with lowest center frequency in the various processing strategies mentioned above. With the possible exception of this lowest band, no information about frequency components within the bands of the filterbank will be available to most implant users with “temporal” representations of FS information.

4. Some alternatives and future directions

Possibilities for better representations of FS information include (a) the acoustic stimulation part of combined EAS and (b) fine adjustments in the sites of stimulation along the electrode array of an implant, as instructed by an instantaneous frequency signal for each channel. The acoustic stimulation part of combined EAS may be perceived in a way similar to that of low-frequency sounds in normal hearing. In such a case, FS information would be fully or largely available to the user up to the frequency limit of the residual hearing. Fine resolution of frequencies in the low-frequency range can support a high level of music reception. In addition, the FS information in the low-frequency band, when combined with electric stimuli for a coarse representation of higher frequencies, may support especially high levels of speech reception in noise.

Fine adjustments in sites of stimulation might be made using virtual channels [8], or through selection of a particular electrode among many. Coding by place of stimulation might be far more effective than coding by frequency or rate of stimulation, as described above. Coding by place may allow representation of frequencies within bands for all channels of the implant processor, not just the lowest band (at best). Studies are now underway to evaluate these possibilities for place coding of FS information.

References

- [1] B.S. Wilson, et al., Cochlear implants: some likely next steps, *Annu. Rev. Biomed. Eng.* 5 (2003) 207–249.
- [2] Z.M. Smith, et al., Chimaeric sounds reveal dichotomies in auditory perception, *Nature* 416 (2002) 87–90.
- [3] K.E. Fishman, R.V. Shannon, W.H. Slattery, Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor, *J. Speech Lang. Hear. Res.* 40 (1997) 1201–1215.
- [4] B.S. Wilson, Engineering design of cochlear implants, in: F.-G. Zeng, A.N. Popper, R.R. Fay (Eds.), *Auditory Prostheses*, Springer-Verlag, New York, 2004, pp. 14–52.
- [5] D.B. Koch, et al., HiResolution and conventional sound processing in the HiResolution bionic ear: using appropriate outcome measures to assess speech recognition ability, *Audiol. Neuro-otol.* 9 (2004) 214–223.
- [6] F.G. Zeng, Temporal pitch in electric hearing, *Hear. Res.* 174 (2002) 101–106.
- [7] E. Favre, M. Pelizzone, Channel interactions in patients using the Ineraid multichannel cochlear implant, *Hear. Res.* 66 (1993) 150–156.
- [8] B.S. Wilson, et al., Recent developments with the CIS strategies, in: I.J. Hochmair-Desoyer, E.S. Hochmair (Eds.), *Advances in Cochlear Implants*, Manz, Vienna, 1994, pp. 103–112.