Hearing Research 333 (2016) 185–193

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Information theoretic evaluation of a noiseband-based cochlear implant simulator

Daniel E. Aguiar [a], N. Ellen Taylor [b], Jing Li [a], Daniel K. Gazanfari [a], Thomas M. Talavage [a, b, *], J. Brandon Laflen [a], Heidi Neuberger [c], Mario A. Svirsky [c, d]

[a] School of Electrical & Computer Engineering, Purdue University, West Lafayette, IN, USA
[b] Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA
[c] DeVault Otologic Research Laboratory, Department of Otolaryngology/Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, IN, USA
[d] Department of Otolaryngology-Head & Neck Surgery, New York University School of Medicine, New York, NY, USA

ARTICLE INFO

Article history:
Received 2 January 2015
Received in revised form
25 August 2015
Accepted 20 September 2015
Available online 25 September 2015

Keywords:
Cochlear implant
Simulation
Noiseband vocoder
Psychophysics
Vowel perception

ABSTRACT

Noise-band vocoders are often used to simulate the signal processing algorithms used in cochlear implants (CIs), producing acoustic stimuli that may be presented to normal hearing (NH) subjects. Such evaluations may obviate the heterogeneity of CI user populations, achieving greater experimental control than when testing on CI subjects. However, it remains an open question whether advancements in algorithms developed on NH subjects using a simulator will necessarily improve performance in CI users. This study assessed the similarity in vowel identification of CI subjects and NH subjects using an 8-channel noise-band vocoder simulator configured to match input and output frequencies or to mimic output after a basalward shift of input frequencies. Under each stimulus condition, NH subjects performed the task both with and without feedback/training. Similarity of NH subjects to CI users was evaluated using correct identification rates and information theoretic approaches. Feedback/training produced higher rates of correct identification, as expected, but also resulted in error patterns that were closer to those of the CI users. Further evaluation remains necessary to determine how patterns of confusion at the token level are affected by the various parameters in CI simulators, providing insight into how a true CI simulation may be developed to facilitate more rapid prototyping and testing of novel CI signal processing and electrical stimulation strategies.

1. Introduction

Cochlear implants (CIs) represent the first example in history where a human sense has been successfully replicated through introduction of electrical stimulation of sensory nerve fibers. These neuroprostheses enable restoration of hearing to profoundly deaf patients, and have had a significant impact on a worldwide basis with approximately 324,000 implantees as of December 2012 (NIDCD, 2014).

CI performance is recognized as being sub-optimal, as evidenced by degraded levels of speech perception. In post-lingually deafened individuals, CIs are intended to restore speech perception by replicating excitation patterns in the auditory nerve as produced under normal acoustic hearing conditions. Under this ideal, perception of external acoustic stimuli would be unchanged relative to performance evidenced prior to the onset of deafness. However, even with extensive post-surgical rehabilitation to improve speech perception, most CI users only achieve open set consonant-nucleus-consonant word identification rates, in quiet, of around 60% (e.g., Bassim et al., 2005; Alkaf and Firszt, 2007), with significantly worse performance in common social conditions of background noise.

Improvements in speech perception performance in CI users have been achieved via both hardware and software changes. Primary examples of these include increases in the number of electrodes and alterations in signal processing strategies. Evaluation of changes using either approach is, at best, challenging (e.g., modifying the number of electrodes requires re-implantation, which is impractical for testing purposes).

While readily manipulated, the evaluation of novel signal processing strategies in a patient population is time-consuming and

* Corresponding author. 465 Northwestern Ave, West Lafayette, IN 47907-2035, USA.
E-mail address: tmt@purdue.edu (T.M. Talavage).

http://dx.doi.org/10.1016/j.heares.2015.09.008
0378-5955/

subsequent generalization of findings may be limited. Evaluation typically requires repeated extended periods in which the patient's implant is alternately programmed with a known or a test strategy (e.g., A-B-A-B paradigm), sometimes necessitating days, weeks, or even months for consistent performance levels to be achieved (Bassim et al., 2005). In addition, several individual-specific factors have been identified in CI users as affecting auditory-only performance, including age of onset and duration of deafness, age of implantation (Holt et al., 2004; Svirsky et al., 2004, 2007; Habib et al., 2010; Tajudeen et al., 2010; Lazard et al., 2012; Blamey et al., 2013) and duration of CI use (Meyer et al., 1998), and experience with a given signal-processing strategy, electrode location and insertion depth, and electrical dynamic range (Dorman and Loizou, 1997b; Loizou et al., 2000). The inherent variability of these parameters across a random selection of likely-to-be-heterogeneous CI users complicates absolute quantification of the performance of a novel signal processing strategy to be applied to a larger population of CI users.

To control for population heterogeneity in CI experimentation, researchers often use a normal-hearing (NH) test population that is presented stimuli altered by a simulation of a CI, commonly a noiseband vocoder (Shannon et al., 1995). Use of a CI simulator increases the control over experiments and increases the size of the potential subject pool. Some factors, such as “patient” age and duration of usage are readily controlled during the selection process, while others, including signal processing strategy, number of electrodes, and (effective) electrode location are configurable within the simulator (Shannon et al., 1995; Dorman et al., 1997a, b; Kaiser and Svirsky, 2000).

Critically, however, it is unclear how readily improvements observed in acoustic simulator users should be expected to translate to the clinical population. Little validation of simulator findings has been performed on CI users. Perhaps as a consequence, the clinical benefit from simulation of CIs is ambiguous, and some studies have noted a disconnect between the behavior of simulator users and CI users (e.g., Friesen et al., 2001; Fu and Nogaki, 2005; Laneau et al., 2006; Litvak et al., 2007; Svirsky et al., 2013).

While comparisons have been made between simulator users and CI users, the focus has not generally been on whether comparable amounts or types of information are being presented by these two modalities. Typical analyses have relied on percent correct identification rates, and even those efforts that have addressed concepts of information transfer have generally not focused on issues of providing subjects with feedback (i.e., training) or making direct comparison of confusion matrices (Strydom and Hanekom, 2011)—e.g., via the Kullback-Leibler Distance or comparable measures. Past comparisons therefore cannot assess whether simulator users receive the same kinds of information as CI users—a key factor in determining if advances in one modality should be expected to translate to the other.

The behavioral study presented here was undertaken to assess the similarity and differences (and thus translational relationship) between a group of CI users and NH subjects using an eight-channel CI simulator, with both groups performing a common task (vowel identification). Evaluation of error patterns and rates at the token level can provide insight regarding the potential for the simulator to serve as a proxy for actual CI users. If the two populations exhibit comparable error patterns and are found to use similar cues to identify presented tokens, the CI simulator can likely serve as a useful testbed for developing strategies that should be assessed in the clinic. If the two populations exhibit differences, it indicates that the particular simulator has limited predictive value for novel strategies as applied to CIs, and raises concerns regarding the unconstrained use of the larger class of comparable simulators—i.e., specific benefits identified through alteration of typical CI simulators may not be consistent with results obtained in subsequent testing on a CI population.

2. Materials and methods

2.1. Subjects

One hundred and four (104) adult native speakers of American English participated as paid volunteers in this experiment. *Cochlear Implant Users:* Twenty-eight (28) subjects were post-lingually deafened CI users (Nucleus 22: 9; Nucleus 24: 8; Clarion: 7; Med-El: 4) using a mix of stimulation strategies (CIS: 11; SPEAK: 10; MPEAK: 3; SPEAK/ACE: 3; ACE: 1). CI users had an average age of 59 ± 14 years (range: 31–79) with an average at implantation of 53 ± 13 years (range: 23–75), for an average period of CI use of 6 ± 3 years (range: 1–13). *Normal Hearing Subjects:* Seventy-six (76) subjects were college-age (range: 18–31) volunteers with no known hearing problems or prior experience with a CI simulator or the test procedure.

2.2. Acoustic stimuli

Source acoustic stimuli consisted of nine/h/-vowel-/d/words (“Had”, “Hawed”, “Head”, “Heard”, “Heed”, “Hid”, “Hood”, “Hud”, and “Who'd”) spoken by an adult male, native speaker of American English, and recorded as WAV files at a sampling frequency of 44.1 kHz.

The simulator used herein is that of Kaiser and Svirsky (2000) as further modified by Morbiwala et al. (2005) and Fitzgerald et al. (2013). Briefly, the acoustic signal was digitized, low pass filtered, and divided into adjacent frequency bands by a bank of analysis filters. For each analysis filter, the temporal envelope was extracted by half wave rectification and low-pass filtering and the temporal envelopes were then used to modulate noise bands that were created using a set of synthesis filters. This was an implementation of the continuous interleaved sampling (CIS) strategy commonly used in CIs, and the update rate of the temporal envelopes was 1250 times per second.

The simulator was configured in one of two manners for the presentation of speech stimuli. In one case (“Unshifted”) a common set of analysis and synthesis filters was selected to deliver relevant frequencies with no mismatch between the frequency allocation table and the (simulated) electrode location. In the second case (“Fullshifted”), a set of synthesis filters representing incomplete insertion of the cochlea was used in conjunction with a set of analysis filters that provided a mismatch of approximately 6 mm (roughly 1.4 octaves) between the analysis filters and the (simulated) electrode location. Note that incomplete insertion means that neurons stimulated by a given electrode have a characteristic frequency that may be significantly higher than the input stimulus.

Table 1
Filter cutoff frequencies (Hz) used in acoustic simulations.

Channel	Channel cutoff frequencies (Hz)			
	Unshifted stimulus		Fullshifted stimulus	
	Analysis	Synthesis	Analysis	Synthesis
1	251–498	251–498	280–545	854–1468
2	502–728	502–728	547–794	1466–2032
3	730–1015	730–1015	794–1099	2032–2732
4	1015–1450	1015–1450	1099–1565	2732–3800
5	1450–2000	1450–2000	1565–2154	3800–5150
6	2000–2600	2000–2600	2154–2798	5150–6622
7	2600–3800	2600–3800	2798–4084	6622–9568
8	3800–6800	3800–6800	4084–7294	9568–10,400