HHS Public Access
Author manuscript
Hear Res. Author manuscript; available in PMC 2017 March 01.

Published in final edited form as:
Hear Res. 2016 March ; 333: 185–193. doi:10.1016/j.heares.2015.09.008.

Information Theoretic Evaluation of a Noiseband-Based Cochlear Implant Simulator

Daniel E. Aguiar[1], **N. Ellen Taylor**[2], **Jing Li**[1], **Daniel K. Gazanfari**[1], **Thomas M. Talavage**[1,2], **J. Brandon Laflen**[1], **Heidi Neuberger**[3], and **Mario A. Svirsky**[3,4]

[1]School of Electrical & Computer Engineering, Purdue University, West Lafayette, IN, USA

[2]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

[3]DeVault Otologic Research Laboratory, Department of Otolaryngology/Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, IN, USA

[4]Department of Otolaryngology-Head & Neck Surgery, New York University School of Medicine, New York, NY, USA

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.

Keywords

cochlear implant; simulation; noiseband vocoder; psychophysics; vowel perception

Corresponding Author: Thomas M. Talavage, 465 Northwestern Ave, West Lafayette, IN 47907-2035, Tel: (765) 494-5475, Fax: (765) 494-3358, tmt@purdue.edu.

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 reimplantation, which is impractical for testing purposes).

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

Materials and Methods

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.

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 6mm (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. Filter center frequencies for the two shift conditions spanned the ranges indicated in Table 1, approximating typical frequency ranges used in cochlear implant speech processors—e.g., the frequency allocation table in some Nucleus devices spans the range 188-7,938 Hz (Jethanamest et al., 2010), and some Advanced Bionics devices use 250-8,700 Hz (Svirsky et al., 2015). In all cases, the analysis and synthesis filters comprised a bank of eight Butterworth bandpass filters. CI simulators are commonly used with filters ranging from 3rd-order through 8th-order (Shannon et al., 1998; Fu and Shannon, 1999; Baskent and Shannon, 2003, 2006). Here, the filters were chosen to be 5th-order for 60 of the assessments (related to the effects of feedback/training), with 16 assessments conducted using 7th-order filters for a preliminary evaluation of the effect of filter sharpness.

Test Protocol

Cochlear Implant Users—Subjects were tested at an average post-implant time of 6 ± 3years. Stimuli were presented in a sound proof booth, passed through the subjects’ own clinical speech processors using a loudspeaker, with stimuli calibrated to 70 dB SPL, C-scale (e.g., Alkaf and Firszt, 2007). After each presentation, subjects were asked to indicate the perceived stimulus in a 9-alternative forced-choice (9AFC) test. Each stimulus was presented 15 times. Subsequent to each presentation, subjects were provided visual feedback (and, in effect, training) as to the correct identity of the preceding stimulus.

Normal Hearing Subjects—The same WAV files as used for the CI users were input to the simulator, with the outputs from the synthesis filters saved as new WAV files. These output WAV files served as the stimuli for testing, being presented to the subjects in a sound proof booth via electrostatic headphones attached to the output of a PC sound card. Each subject selected a comfortable intensity level for presentation of the stimuli. Two forms of experimentation (see below) were conducted: (1) *without feedback/training*, and (2) *with feedback/training*.

Experiments without Feedback/Training (-FT)—After presentation of each stimulus, subjects reported the perceived vowel via a MATLAB-based interface, on a PC. Following a brief pause, the next stimulus was presented. Each stimulus was presented 15 times. Within each of two testing blocks, subjects were presented only a single type of shift (Unshifted or Fullshifted), always using 5th-order filters—abbreviated below as *Unshifted/5th/-FT* and *Fullshifted/5th/-FT*.

Experiments with Feedback/Training (+FT)—The presentation of stimuli and response recording were conducted using the same MATLAB-based interface, as above, but following each response, the interface would provide visual feedback regarding the actual presented stimulus. To avoid inter-subject variability in learning/adaptation, each of the vowel stimuli was presented 45 times: the first 30 presentations were considered training, followed by the data collection experiment involving 15 presentations of each stimulus. For the Unshifted condition, these experiments were conducted twice, once with 5^{th}-order filters, and once with 7^{th}-order filters—i.e., *Unshifted/5th/+FT* and *Unshifted/7th/+FT*. Experiments using the Fullshifted stimuli were only conducted using 5^{th}-order filters—ie., *Fullshifted/5th/+FT*. Therefore, subjects performed three testing blocks, in each of which only a single type of stimulus (e.g., same filter order and shift) was presented.

Data Analysis

Presented and perceived data from the tests described above were organized into six confusion matrices based on user group, shift and filter order—CI users (with feedback/training); simulator users identifying Unshifted stimuli (5th-order filters), both with and without feedback/training; simulator users identifying Unshifted stimuli (7th-order filters) with feedback/training; and simulator users identifying Fullshifted stimuli (5th-order filters), both with and without feedback/training. Confusion rates were normalized (converted to a probability such that all possible responses to presentation of a given vowel sound sum to one) and averaged across subjects within each testing condition.

CI User Analysis—A statistical analysis was performed to see if CI users having different speech-processing strategies may be grouped together for purposes of comparing the CI aggregate confusion matrix with those of the simulator users. Confusion matrices were converted to 9 × 9 rank order matrices, showing the rank order (e.g., most likely = 1, 2, …, least likely = 9) in which CI users perceived each presented stimulus. Rank order matrices were compared within the CI population to determine if the error patterns were consistent across subjects. A Monte Carlo cross-validation test was used wherein, for 10^5 random permutations, subjects were split into two equal-sized groups and the corresponding confusion matrices and rank order matrices computed. For each permutation, the average squared absolute difference between the two group-level rank order matrices was calculated as follows:

$$E_{squared}=\frac{1}{9^2}\sum_{i=1}^{9}\sum_{j=1}^{9}\left(R_1\left[i,j\right]-R_2\left[i,j\right]\right)^2 \quad (1)$$

$E_{squared}$ is the 2-norm error between the rank order matrices R_1 and R_2, obtained from groups 1 and 2, respectively; and wherein $R_n[i,j]$ is the rank order associated with perception of stimulus *j* when stimulus *i* was presented to the members of group *n*. The probability that the observed error norm could arise by chance was then computed via a permutation test of all possible orderings of the ranks one through nine (Weerahandi, 2003).

Comparison of CI Users to Simulator Users—Features of the signals that are being delivered to the CI users and simulator users will inherently not be equivalent, but we desire to understand how much of the acoustic signal delivered to the simulator user is equivalent in value and meaning to that delivered to the CI user. Rather than making *a priori* assumptions about the features of the acoustic signal that are being delivered, we have chosen initially to examine percent correct (over the whole matrix) and subsequently employ information theoretic tools to examine the information received by subjects.

Confusion Matrix Analysis—Percent correct measures, as represented in confusion matrices, were evaluated in both an absolute and rank order sense, for each configuration of CI simulator used (i.e., by shift, filter order and feedback/testing condition). The root mean square error in the percent correct rate was computed for each of the simulator configurations, permitting a coarse assessment of the “similarity” of each to the CI user population. The 2-norm (Eq. 1; $E_{squared}$) was again used to compare the performance of the CI user and simulator user groups, in the rank order sense, with the probability that the observed 2-norm measures represented chance (i.e., the distribution of rank orders is equivalent between two groups) assessed via a permutation test, as above.

Information Theoretic Analysis—While commonly cited as evidence of effective simulation of CI user processing, comparison of aggregate correct identification rates (Dorman and Loizou, 1997a; Dorman et al., 1997b) between simulator users and CI users is not sufficient to determine if the normally hearing group is accurately simulating CI user perception. Ultimately, this metric provides no potential for comparison of amounts of information used, nor does it permit comparison (direct or indirect) of the presence/absence of underlying cues on which a listener may make a categorical decision. The use of rank orders—as above—provides an opportunity to assess the pattern of responses, but remains insufficient with regard to concluding whether two groups are similar.

Given the limitations noted above, two information theoretic approaches were applied to evaluation the similarity of the CI user and the CI simulator user groups: (1) Mutual Information (MI), and (2) the Kullback-Leibler (KL) Distance (Cover and Thomas, 2012). MI was computed for each confusion matrix (Miller and Nicely, 1955) to quantify how

Author Manuscript
Author Manuscript
Author Manuscript

much presented stimulus information (in bits) was transmitted by the system (CI or CI simulator). Note that isolated application of MI results in limitations comparable to those based on percent correct alone. Knowing the quantity of information that is conveyed to one group does not provide insight into the nature of that conveyed information— i.e., it is possible to have two systems convey similar quantities of information while not conveying *equivalent* information (e.g., in one system the subject identifies “Had” as “Hid”, while in the other “Had” is identified as “Heed”). It is even possible for two groups who receive identical quantities of information to have received completely disjoint information. Therefore, to evaluate equivalence of information transfer, the KL Distance was computed for each pair of confusion matrices. The KL Distance, *D(p//q)*, measures the likelihood of observing data with one distribution (q) given the real distribution (p). Note that because the KL Distance is not symmetric, it was computed in both directions (i.e., CI-vs.-simulator and simulator-vs.-CI).

To effect computation of MI and the KL Distance, each confusion matrix was converted from a conditional probability matrix to a joint probability matrix. This provided the joint probability of a subject responding that they heard stimulus y and that the stimulus was actually x, rather than the conditional probability of y given x (as in the original confusion matrix). Since the identification experiment design used equally likely stimuli, the distribution of the input (x) was uniform and the joint distribution was found by dividing all elements of the confusion matrix by the number of stimuli (here, nine), as specified by Bayes’ Theorem (Eq. 2).

$$p\left(x,y\right)=p\left(y|x\right)p\left(x\right)=\frac{p\left(y\vee x\right)}{9} \quad (2)$$

Results

All Monte Carlo trials conducted to assess similarity among the CI users were found to yield less than 5% probability of arising by chance. Therefore, all CI users are hereafter considered as coming from one group as far as the rank orders of their confusions are concerned.

Qualitative differences were observed between simulator users and CI users in terms of aggregate correct identification rates (see Table 2). Unshifted (5th-order filters) stimulation with feedback/training resulted in the highest level of percent correct performance. In fact, both conditions (5th- and 7th-order filters) involving Unshifted stimulation with feedback/ training resulted in higher average rates of correct identification than were exhibited by the CI users, but these rates were not statistically significantly elevated. The underlying spread of correct identification rates on a per-subject basis is shown in Figure 1.

While some aggregate rates of correct identification for the simulator users were comparable to those exhibited by CI users, rates of specific token correct identification were not generally consistent with CI users (Table 2). Particularly note that the conditions without feedback/training—independent of shift and filter order—often resulted in the most common perceived stimulus being something other than the presented stimulus (i.e., the rank order for the correct perception was not 1). The variation in token-level response may be seen by comparing the aggregated CI user percent correct and rank order confusion matrix (Table 3) with those obtained from the simulator users (Tables 4-6).

Comparisons of CI users and simulator user rank order matrices are shown in Table 7. Only those conditions involving feedback/training (+FT) produced rank orders significantly ($p < 0.01$) consistent with those of CI users. Note that the aforementioned similarity between aggregate rates of correct identifications for CI users and the Unshifted stimulation with feedback/training is further observed in these data, but now extends to the Fullshifted stimulation with feedback/training case, as well.

Within the context of the ensemble of correct identification rates, Unshifted (7th-order filters) stimulation with feedback/training was found to provide the closest approximation of the CI user population. For this case, the correct perceived stimulus was always the most common perception, and the specific token correct identifications exhibited the lowest root mean square error relative to the CI users (16.7%; see Table 2). Note that the relatively minor change (i.e., within the range typically used in CI simulator experiments) made to the order of the bandpass filters (5th- vs. 7th-order) did produce noticeable changes in the confusion patterns of subjects and correct performance (see Tables 4A and 6).

Results of the analysis of information theoretic approaches—MI and KL Distance—are presented in Table 7 and Figure 2. From the MI perspective, CI users receive 1.42 bits of information per stimulus, which is observed to be below the information transfer rates observed for both (5th- and 7th-order filters) cases of Unshifted stimulation with feedback/training, but greater than the rates observed for any of the Fullshifted stimulation conditions or cases of Unshifted stimulation without feedback/training. Note that the smallest KL Distance observed between the CI users and any simulator group was approximately 0.28 bits, for Unshifted (7th-order filters) stimulation with feedback/training (Figure 2; *Unshifted/7th/+FT* vs. *CI User*). This distance remains larger than 98.38% of all *CI user* vs. *CI user* comparisons— i.e., this distance indicates that results obtained with the simulator remain statistically significantly different from those obtained from actual CI users ($p < 0.017$). Comparably, Unshifted (5th-order filters) stimulation with feedback/training was associated with a KL Distance of 0.45 bits, with respect to the CI users (Figure 2; *CI user* vs. *Unshifted/5th/+FT*). Given that this distance is greater than that exhibited by 99.89% of the permutations comparing CI users amongst themselves, it remains a highly significant difference (i.e., $p < 0.0012$).

Discussion

This study investigated the low-level consistency of speech perception performance of CI users and normally hearing subjects using a noiseband vocoder as a simulation of the processing and stimulation associated with the electric hearing of a CI. While similar aggregate rates of token identification (and comparable ordering of “preferred” errors) were achieved across these two cohorts, in no case was token-level performance comparable, particularly under information theoretic analysis. Further, the limited predictive capability of the normally hearing cohort performance to the CI users' performance exhibited marked sensitivity to non-standardized parameters of the vocoder configuration. These findings raise concern that the use of noiseband vocoder (or other) simulations of CI signal processing may not allow novel CI processing approaches or stimulation strategies to be meaningfully evaluated with normally hearing subjects. General use of such vocoders to simulate the effects or benefits of CIs is cautioned.

CI Simulator Performance and Sensitivity to Implementation

The outcomes obtained in this study suggest that a noiseband vocoder based simulation represents, at best, a rough approximation of speech perception performance by CI users. When feedback/training was provided (i.e., +FT conditions) to the normally hearing listeners after making perceptual judgments of acoustic simulations of CI inputs, these subjects generated aggregate correct identification rates (Figure 1) and rank orders of errors (Table 7) similar to those of CI users (Table 3). However, subtle characteristics of CI user performance at the individual token level were not reproduced in the responses of those listening to the simulation (Tables 4A, 5A, 6).

The amount of information received from the simulator, as assessed using MI, was generally markedly different from that obtained using actual CIs, even though correct identification rates were not significantly different (Tables 2 and 7). Additionally, KL Distances (Table 7) between simulator users and CI users were notably higher than the KL Distances obtained when comparing across individuals within the CI user group (Figure 2). In other words, no tested simulator condition—common to all such subjects—was able to achieve speech perception performance that was significantly consistent with the CI user population.

An important consideration to take into account, however, is that the specific implementation of a CI simulator is likely to have a significant effect on most measures of speech perception performance. For example, in this investigation the variation of bandpass filter order from 5th (30 dB/oct) to 7th (42 dB/oct) led to appreciable—and arguably surprising—reductions in both percent correct identification rate (80% down to 73%; Table 2) and MI (2.08 bits per stimulus down to 1.72 bits per stimulus; Table 7). If such small changes in the simulator configuration can produce changes in the cues presented to subjects, conclusions regarding application to CIs of new algorithms developed from the use of that simulator are likely to be configuration-specific.

It remains to future investigations to identify the set of simulator parameters that will best provide a match to the speech perception performance of CI users (Svirsky et al., 2013; Butts et al., 2015).

Effects of Feedback/Training on Task Performance

When using a noiseband vocoder to simulate CI user perception, feedback/training should be provided to increase the similarity of the simulator users' response to those of CI subjects. The rank orders of confusions for the two groups were found to be statistically similar only for conditions with feedback/training (i.e., +FT; Table 7).

Metrics for Evaluation of Speech Perception Performance

Reflection on the results obtained in this investigation strongly argues for use of information theoretic approaches when evaluating or developing simulator configurations. Looking at Tables 2 and 7, it may be seen that the Fullshifted (5th-order filters) stimulation with feedback/training provided a lesser amount of information than, for example, the Unshifted (5th-order filters) stimulation without feedback/training—0.85 vs. 0.98 bits—even though it resulted in a higher percent correct—44.3% vs. 31.9%. Such disagreements between measures have more to do with the “background” of the performance. For example, even though it had the lower percent correct identification rate, the distribution of errors for the Unshifted (5th-order filters) stimulation without feedback/training (Table 5A) is notably more consistent with the CI users (Table 4) than is the corresponding distribution for the Fullshifted (5th-order filters) stimulation with feedback/training (Table 6B). Given that meaningful, predictive simulation of CI users must also include error patterns, the information theoretic approach is here more effective.

The rank order analysis in Table 7 highlights that providing feedback/training was important to better approximating CI users, but did not provide meaningful insight as to which simulator version was most useful. Conversely, the subtle indication from the rank order analysis that the Unshifted (7th-order filters) stimulation with feedback/training was closest in performance to the CI users is strongly supported by information theoretic approaches. This testing condition yielded the MI value closest in absolute value to that of CI users (1.77 bits vs. 1.44 bits) and had the smallest KL Distances from the CI users (0.28 and 0.43).

Therefore, it is advocated that evaluation of simulated CI speech perception performance be conducted, relative to actual CI user data, using analytical techniques comparable to the KL Distance approach implemented herein. This metric provides a direct comparison of the probability distributions of experimental and reference groups, providing insight whether the information provided to users of the simulator could potentially be the same as that provided to CI users. In the absence of such probabilistic analysis, there is no way to determine if simulator subjects are performing the same task as CI users or if they are performing a fundamentally different task that simply has comparable complexity.

Possible Considerations for Simulating CIs

As noted above, it is likely possible to refine current CI simulators or create an altogether new algorithm to better model CI user speech perception performance results. Such work must be done from the perspective of minimizing measures such as the KL Distance between CI and simulator users. Critically, metrics such as percent correct identification rates or even rank ordering of errors are, in and of themselves, insufficient as objective functions for such simulators. Use of such functions could yet result in superficial matches to performance of CI users. However, the simulator will actually represent a task that has equivalent difficulty but likely results in different central processing than in the implant case, making application to CI users of algorithms or stimulation patterns developed on simulator users dubious.

Conclusions

Normally hearing subjects listening to the output of acoustic simulations of CI electrical stimulation may not accurately model the underlying central processing effected in CI users during speech perception performance. Documented here for an eight-channel noiseband vocoder, used during vowel identification experiments, gross similarities in performance are not necessarily reflected at a more granular (e.g., token) level. Testing or development of new CI signal processing algorithms in such settings, therefore, will not necessarily provide a strong indication of the potential value of such an algorithm to improve CI user speech perception. In this study, the most promising configuration tested was that of an unshifted simulator with feedback provided after each token presentation and the associated training. However, this configuration provided more information to the subject than is received by CI users, resulting in an expected higher correct identification rate. 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.

Acknowledgments

This work was partially supported by grants from the Purdue University/Indiana University School of Medicine Collaboration in Biomedical Research Committee (Svirsky/Talavage), NIH/NIDCD (Svirsky: R01 DC003937), the Purdue Biomedical Engineering NSF REU program (Grant Number: 0353901-EEC) and the Purdue Summer Undergraduate Research Fellowship (SURF) program. The authors wish to thank Kyle Aitken and Krithika Iyer for their assistance with data collection.

References

Alkaf FM, Firszt JB. Speech recognition in quiet and noise in borderline cochlear implant candidates. Journal of the American Academy of Audiology. 2007; 18:872–882. [PubMed: 18496996]

Baskent D, Shannon RV. Speech recognition under conditions of frequency-place compression and expansion. J Acoust Soc Am. 2003; 113:2064–2076. [PubMed: 12703717]

Baskent D, Shannon RV. Frequency transposition around dead regions simulated with a noiseband vocoder. J Acoust Soc Am. 2006; 119:1156–1163. [PubMed: 16521776]

Bassim MK, Buss E, Clark MS, Kolln KA, Pillsbury CH, Pillsbury HC 3rd, Buchman CA. MED-EL Combi40+ cochlear implantation in adults. Laryngoscope. 2005; 115:1568–1573. [PubMed: 16148696]

Blamey P, et al. Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. Audiol Neurootol. 2013; 18:36–47. [PubMed: 23095305]

Butts, AM.; Cook, SJ.; Berisha, V.; Dorman, MF. A modular signal processing program to simulate and test the sound quality of a cochlear implant. Conference on Implantable Auditory Prostheses; Lake Tahoe, CA. 2015.

Cover, TM.; Thomas, JA. Elements of information theory. John Wiley & Sons; 2012.

Dorman MF, Loizou PC. Speech intelligibility as a function of the number of channels of stimulation for normal-hearing listeners and patients with cochlear implants. The American journal of otology. 1997a; 18:S113–114. [PubMed: 9391623]

Dorman MF, Loizou PC. Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with continuous interleaved sampling (CIS) processors. Ear Hear. 1997b; 18:147–155. [PubMed: 9099564]

Dorman MF, Loizou PC, Rainey D. Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. J Acoust Soc Am. 1997a; 102:2993–2996. [PubMed: 9373986]

Dorman MF, Loizou PC, Rainey D. Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. J Acoust Soc Am. 1997b; 102:2403–2411. [PubMed: 9348698]

Fitzgerald MB, Sagi E, Morbiwala TA, Tan CT, Svirsky MA. Feasibility of real-time selection of frequency tables in an acoustic simulation of a cochlear implant. Ear Hear. 2013; 34:763–772. [PubMed: 23807089]

Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J Acoust Soc Am. 2001; 110:1150–1163. [PubMed: 11519582]

Fu QJ, Shannon RV. Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. J Acoust Soc Am. 1999; 105:1889–1900. [PubMed: 10089611]

Fu QJ, Nogaki G. Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. J Assoc Res Otolaryngol. 2005; 6:19–27. [PubMed: 15735937]

Habib MG, Waltzman SB, Tajudeen B, Svirsky MA. Speech production intelligibility of early implanted pediatric cochlear implant users. Int J Pediatr Otorhinolaryngol. 2010; 74:855–859. [PubMed: 20472308]

Holt, RF.; Svirsky, MA.; Neuburger, H.; Miyamoto, RT. International congress series. Elsevier; 2004. Age at implantation and communicative outcome in pediatric cochlear implant users: Is younger always better?; p. 368-371.

Jethanamest D, Tan CT, Fitzgerald MB, Svirsky MA. A new software tool to optimize frequency table selection for cochlear implants. Otol Neurotol. 2010; 31:1242–1247. [PubMed: 20729774]

Kaiser, AR.; Svirsky, MA. Using a personal computer to perform real-time signal processing in cochlear implant research. 9th Digital Signal Processing Workshop; Hunt, Texas. IEEE Signal Processing Society; 2000.

Laneau J, Moonen M, Wouters J. Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants. J Acoust Soc Am. 2006; 119:491–506. [PubMed: 16454303]

Lazard DS, et al. Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. PLoS One. 2012; 7:e48739. [PubMed: 23152797]

Litvak LM, Spahr AJ, Saoji AA, Fridman GY. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. J Acoust Soc Am. 2007; 122:982–991. [PubMed: 17672646]

Loizou PC, Poroy O, Dorman M. The effect of parametric variations of cochlear implant processors on speech understanding. J Acoust Soc Am. 2000; 108:790–802. [PubMed: 10955646]

Meyer TA, Svirsky MA, Kirk KI, Miyamoto RT. Improvements in speech perception by children with profound prelingual hearing loss: effects of device, communication mode, and chronological age. J Speech Lang Hear Res. 1998; 41:846–858. [PubMed: 9712131]

Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonants. J Acoust Soc Am. 1955; 27:338–352.

Morbiwala, T.; Svirsky, M.; El-Sharkway, M.; Rizkalla, M. A PC-based speech processor for cochlear implant fitting that can be adjusted in real-time. Circuits and Systems, 2005; 48th Midwest Symposium on; IEEE; 2005. p. 1310-1313.

NIDCD. Cochlear Implants. 2014. http://www.nidcd.nih.gov/health/hearing/coch.asp

Shannon RV, Zeng FG, Wygonski J. Speech recognition with altered spectral distribution of envelope cues. J Acoust Soc Am. 1998; 104:2467–2476. [PubMed: 10491708]

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Science. 1995; 270:303–304. [PubMed: 7569981]

Strydom T, Hanekom JJ. The performance of different synthesis signals in acoustic models of cochlear implants. J Acoust Soc Am. 2011; 129:920–933. [PubMed: 21361449]

Svirsky MA, Teoh SW, Neuburger H. Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. Audiol Neurootol. 2004; 9:224–233. [PubMed: 15205550]

Svirsky MA, Chin SB, Jester A. The effects of age at implantation on speech intelligibility in pediatric cochlear implant users: Clinical outcomes and sensitive periods. Audiological Medicine. 2007; 5:293–306.

Svirsky MA, Fitzgerald MB, Sagi E, Glassman EK. Bilateral cochlear implants with large asymmetries in electrode insertion depth: implications for the study of auditory plasticity. Acta oto-laryngologica. 2015; 135:354–363. [PubMed: 25719506]

Svirsky, MA.; Ding, N.; Sagi, E.; Tan, C-T.; Fitzgerald, M.; Glassman, EK.; Seward, K.; Neuman, AC. Validation of acoustic models of auditory neural prostheses. Acoustics, Speech and Signal Processing (ICASSP), 2013; IEEE International Conference on; IEEE; 2013. p. 8629-8633.

Tajudeen BA, Waltzman SB, Jethanamest D, Svirsky MA. Speech perception in congenitally deaf children receiving cochlear implants in the first year of life. Otol Neurotol. 2010; 31:1254–1260. [PubMed: 20814343]

Weerahandi, S. Exact statistical methods for data analysis. Springer; 2003.

Highlights

- Speech perception was tested for a noiseband vocoder cochlear implant simulator.
- Information theoretic metrics are recommended for evaluation of novel strategies.
- The simulator replicated aggregate, but not token-level, behavior of implantees.
- The simulator performance is sensitive to subtle changes in filter parameters.
- Simulators should be used with caution in cochlear implant strategy development.

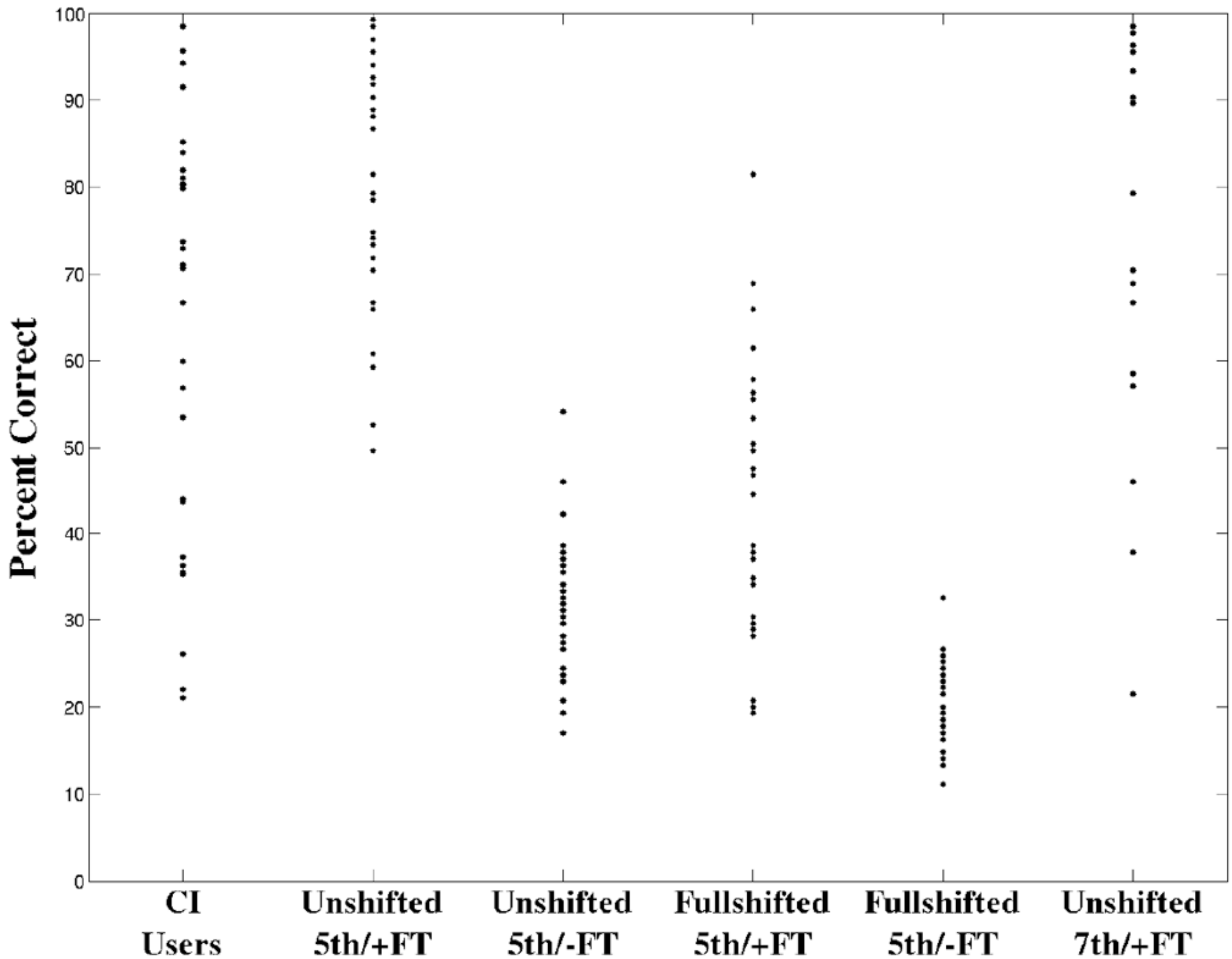

Figure 1.
Scatter plot of distribution of accuracy by subject group (9-alternative forced choice vowel test).

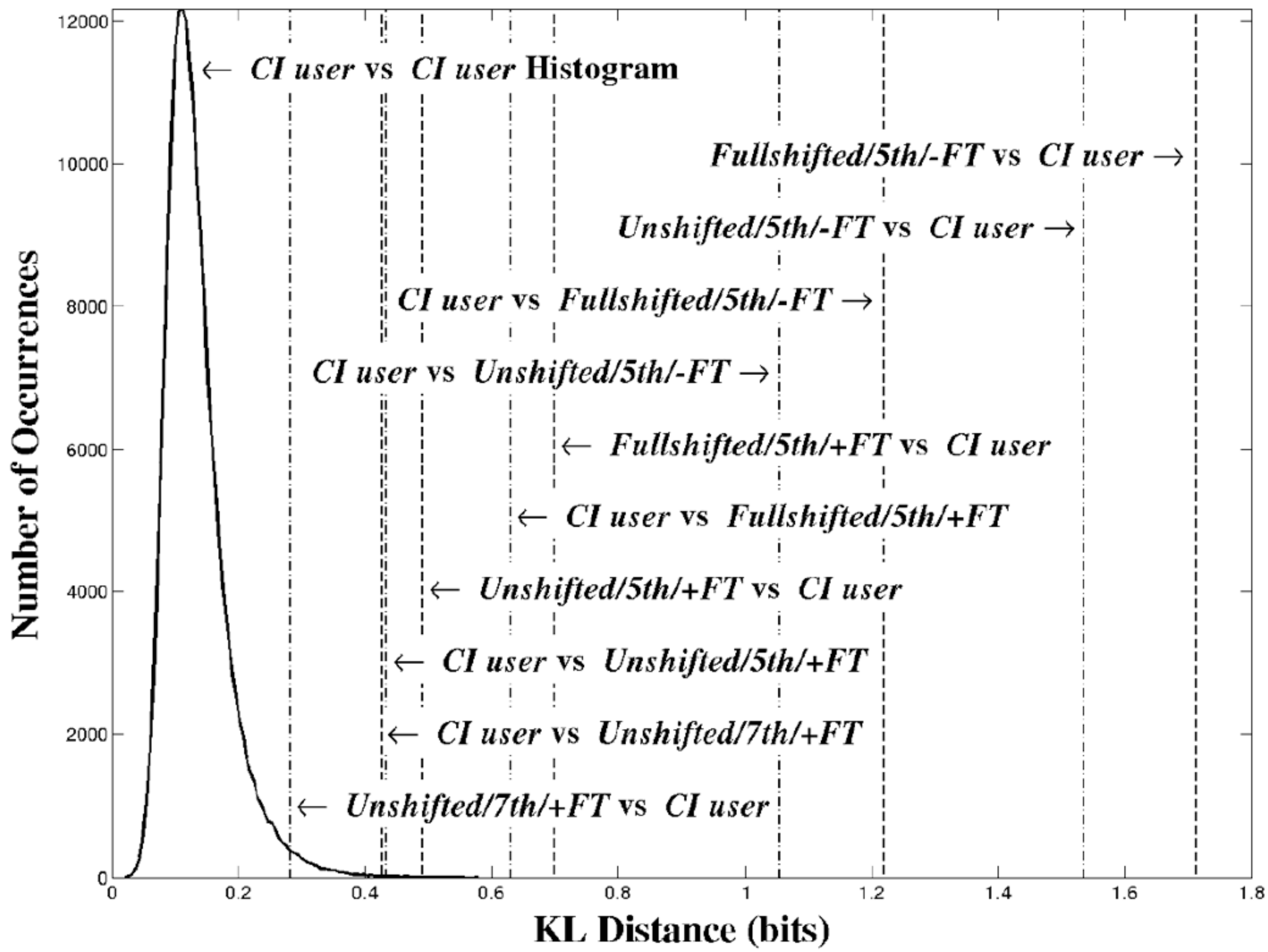

Figure 2.

Comparison of KL Distances between CI user groups, and between CI users and each tested condition for simulator users. The histogram of KL Distances within the CI user group was constructed through 100,000 permutations of random assignment of CI users into two equal-sized groups, with computation of both KL Distance (given the KL Distance is not symmetric) measures conducted for each permutation. This histogram provides a perspective to the KL Distances calculated between the confusion matrices of CI users (Table 3) and simulator users (Tables 4-7). Note that all comparisons of CI users to simulator users resulted in distances greater than 98.38% of all (randomly generated) subdivisions of the CI users, indicating that the simulator conditions resulted in statistically significant KL Distances at the $p < 0.017$ level.

Table 1

Filter cutoff frequencies (Hz) used in acoustic simulations.

	Channel Cutoff Frequencies (Hz)			
	Unshifted Stimulus		**Fullshifted Stimulus**	
Channel	**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	9568-10400

Table 2

Vowel-specific and aggregate speech perception performance for CI and simulator users. Both Percent Correct (9AFC; chance rate = 11%) and rank order of the correct selections (ideal = 1) are indicated.

Testing Group/Condition		**Mean % Correct (Rank)**									**Aggregate % Correct *(Mean ±StdDev)***	**RMSE vs. CI *(%)***
	Feedback/Training	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**		
CI Users	*+FT*	79.8 (1)	68.6 (1)	72.5 (1)	65.6 (1)	77.8 (1)	67.2 (1)	46.1 (1)	56.3 (1)	51.5 (1)	65.0 ± 32.1	*N/A*
Unshifted (5th-order filters)	*+FT*	60.7 (1)	83.8 (1)	70.2 (1)	66.7 (1)	96.7 (1)	89.3 (1)	86.2 (1)	76.4 (1)	91.8 (1)	80.2 ± 26.1	23.8
	-FT	22.2 (2)	28.2 (2)	26.9 (2)	13.6 (3)	42.7 (1)	58.7 (1)	48.7 (1)	31.3 (2)	15.1 (3)	31.9 ± 29.0	38.0
Fullshifted (5th-order filters)	*+FT*	33.3 (1)	39.6 (1)	30.7 (1)	37.8 (1)	57.8 (1)	80.7 (1)	29.8 (1)	32.2 (1)	56.9 (1)	44.3 ± 31.1	27.9
	-FT	33.8 (1)	26.9 (2)	16.0 (2)	6.9 (7)	31.3 (2)	34.0 (1)	9.6 (4)	10.9 (4)	8.2 (6)	19.7 ± 20.9	46.0
Unshifted (7th-order filters)	*+FT*	70.0 (1)	82.9 (1)	71.7 (1)	65.4 (1)	83.8 (1)	56.2 (1)	66.2 (1)	71.2 (1)	89.2 (1)	73.0 ± 32.2	16.7

Table 3

Vowel identification task (9AFC) confusion matrix for CI users (with feedback/training). Both the Percent Correct (chance rate = 11%) and rank order of the correct selection (ideal = 1) are indicated.

Cochlear Implant Users with Feedback/Training [*CI Users*]

	% Perceived (Rank)								
Played	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HAD	79.8 (1)	1.1 (7)	1.3 (6)	0.9 (8.5)	6.3 (2)	3.4 (4)	0.9 (8.5)	2.5 (5)	3.8 (3)
HAWED	0.0 (9)	68.6 (1)	9.7 (2)	5.0 (4)	2.7 (6)	0.5 (8)	3.8 (5)	8.8 (3)	0.9 (7)
HEAD	0.7 (8.5)	6.8 (3)	72.5 (1)	2.7 (6)	3.9 (4.5)	0.7 (8.5)	3.9 (4.5)	8.0 (2)	1.0 (7)
HEARD	0.9 (9)	4.9 (4)	9.9 (2)	65.6 (1)	3.1 (6)	4.3 (5)	6.3 (3)	2.5 (7.5)	2.5 (7.5)
HEED	0.9 (8.5)	5.2 (3)	3.4 (4)	6.5 (2)	77.8 (1)	1.6 (6)	2.7 (5)	0.9 (8.5)	1.1 (7)
HID	4.3 (4)	0.2 (9)	0.9 (8)	9.4 (3)	1.3 (7)	67.2 (1)	2.2 (6)	3.8 (5)	10.6 (2)
HOOD	0.4 (9)	6.3 (4)	13.3 (3)	18.7 (2)	2.5 (8)	3.1 (7)	46.1 (1)	4.3 (6)	5.4 (5)
HUD	0.9 (9)	8.6 (3)	10.6 (2)	3.8 (6)	2.5 (7)	2.3 (8)	8.3 (4)	56.3 (1)	6.8 (5)
WHOD	1.1 (7.5)	0.4 (9)	2.3 (6)	4.8 (4)	1.1 (7.5)	26.9 (2)	4.2 (5)	7.8 (3)	51.5 (1)

Table 4

Vowel identification task (9AFC) confusion matrix for simulator users experiencing Unshifted (5th-order filters) stimulation (A) with feedback/training (*Unshifted/5th/+FT*) and (B) without feedback/training (*Unshifted/5th/-FT*). Both Percent Correct (chance rate = 11%) and rank order of the correct selection (ideal = 1) are indicated.

(A) Unshifted (5th-order filters) with Feedback/Training [*Unshifted/5th/+FT*]

	% Perceived (Rank)								
Played	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HAD	60.7 (1)	14.9 (2)	13.3 (3)	6.4 (4)	1.1 (6)	0.9 (7)	0.7 (8)	1.8 (5)	0.2 (9)
HAWED	4.4 (2)	83.8 (1)	0.2 (7.5)	4.2 (3)	0.0 (9)	0.2 (7.5)	2.7 (4.5)	1.8 (6)	2.7 (4.5)
HEAD	7.6 (3)	1.6 (6)	70.2 (1)	2.7 (5)	0.4 (8)	9.6 (2)	0.7 (7)	7.3 (4)	0.0 (9)
HEARD	10.7 (2)	4.4 (5)	4.4 (5)	66.7 (1)	0.9 (8)	0.4 (9)	4.4 (5)	1.6 (7)	6.4 (3)
HEED	0.0 (7)	0.0 (7)	0.7 (3)	0.0 (7)	96.7 (1)	2.4 (2)	0.2 (4)	0.0 (7)	0.0 (7)
HID	0.2 (6.5)	0.0 (8.5)	3.3 (2.5)	0.9 (5)	0.0 (8.5)	89.3 (1)	2.7 (4)	3.3 (2.5)	0.2 (6.5)
HOOD	0.9 (6)	0.0 (9)	1.3 (4.5)	1.3 (4.5)	0.2 (7.5)	0.2 (7.5)	86.2 (1)	7.1 (2)	2.7 (3)
HUD	4.7 (4)	0.9 (7)	5.8 (3)	1.6 (6)	0.0 (9)	4.0 (5)	6.4 (2)	76.4 (1)	0.2 (8)
WHOD	0.0 (7)	0.0 (7)	0.0 (7)	2.2 (3)	0.0 (7)	0.0 (7)	5.6 (2)	0.4 (4)	91.8 (1)

(B) Unshifted (5th-order filters) without Feedback/Training [*Unshifted/5th/-FT*] % Perceived (Rank)

	% Perceived (Rank)								
Played	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HAD	22.2 (2)	6.2 (6)	25.3 (1)	10.7 (4)	6.0 (7)	15.6 (3)	6.9 (5)	3.1 (9)	4.0 (8)
HAWED	2.9 (6)	28.2 (2)	2.2 (7)	7.3 (4)	0.2 (8)	0.0 (9)	34.4 (1)	6.4 (5)	18.2 (3)
HEAD	10.7 (3)	1.6 (8)	26.9 (2)	1.1 (9)	8.2 (4)	42.9 (1)	3.1 (6)	3.6 (5)	2.0 (7)
HEARD	32.9 (1)	27.3 (2)	3.3 (7)	13.6 (3)	1.6 (9)	4.9 (6)	7.6 (4)	6.7 (5)	2.2 (8)

(A) Unshifted (5th-order filters) with Feedback/Training [*Unshifted/5th/+FT*]

	% Perceived (Rank)								
Played	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HEED	6.2 (4)	0.0 (8.5)	15.8 (3)	0.2 (7)	42.7 (1)	34.2 (2)	0.4 (5.5)	0.4 (5.5)	0.0 (8.5)
HID	7.8 (3)	0.2 (9)	20.7 (2)	0.4 (8)	6.7 (4)	58.7 (1)	1.3 (6)	1.1 (7)	3.1 (5)
HOOD	2.0 (7)	7.1 (4)	1.6 (8)	4.9 (5)	0.7 (9)	3.1 (6)	48.7 (1)	19.8 (2)	12.2 (3)
HUD	5.3 (4)	3.8 (5)	3.3 (6)	2.9 (7)	0.2 (9)	2.7 (8)	38.0 (1)	31.3 (2)	12.4 (3)
WHOD	5.1 (6)	18.9 (2)	2.4 (7)	7.3 (5)	0.7 (9)	1.6 (8)	36.7 (1)	12.2 (4)	15.1 (3)

Table 5

Vowel identification task (9AFC) confusion matrix for simulator users experiencing Fullshifted (5th-order filters) stimluation (A) with feedback/training (*Fullshifted/5th/+FT*) and (B) without feedback/training (*Fullshifted/5th/-FT*). Both Percent Correct (chance rate = 11%) and rank order of the correct selection (ideal = 1) are indicated.

(A) Fullshifted (5th-order filters) with Feedback/Training [*Fullshifted/5th/+FT*]

	% Perceived (Rank)								
Played	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HAD	33.3 (1)	22.4 (2)	13.8 (4)	14.4 (3)	6.7 (5)	2.4 (8)	2.7 (6.5)	1.6 (9)	2.7 (6.5)
HAWED	24.9 (2)	39.6 (1)	6.9 (4)	16.0 (3)	2.0 (8)	1.3 (9)	3.3 (6)	2.2 (7)	3.8 (5)
HEAD	14.9 (3)	2.9 (8)	30.7 (1)	6.9 (6)	7.6 (5)	13.1 (4)	6.7 (7)	15.1 (2)	2.2 (9)
HEARD	7.1 (6)	9.8 (4)	15.3 (2)	37.8 (1)	14.9 (3)	2.2 (8)	4.0 (7)	1.3 (9)	7.6 (5)
HEED	6.0 (4)	2.7 (7.5)	8.7 (3)	5.6 (5.5)	57.8 (1)	9.1 (2)	2.7 (7.5)	5.6 (5.5)	2.0 (9)
HID	2.9 (4)	0.2 (8.5)	6.4 (2.5)	0.9 (7)	0.2 (8.5)	80.7 (1)	1.1 (5.5)	6.4 (2.5)	1.1 (5.5)
HOOD	5.6 (5)	3.1 (8.5)	18.7 (2)	5.3 (6)	3.1 (8.5)	10.9 (4)	29.8 (1)	18.4 (3)	5.1 (7)
HUD	13.8 (4)	1.3 (9)	22.7 (2)	4.2 (6)	1.6 (7.5)	17.1 (3)	5.6 (5)	32.2 (1)	1.6 (7.5)
WHOD	4.0 (6)	12.2 (2)	2.7 (7)	8.7 (3)	1.6 (8)	1.1 (9)	7.3 (4)	5.6 (5)	56.9 (1)

(B) Fullshifted (5th-order filters) without Feedback/Training [*Fullshifted/5th/-FT*]

	% Perceived (Rank)								
Played	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HAD	33.8 (1)	18.4 (2)	10.4 (3)	10.0 (5)	4.9 (6)	4.4 (8)	10.2 (4)	4.7 (7)	3.1 (9)
HAWED	30.0 (1)	26.9 (2)	7.1 (5.5)	11.1 (3)	2.0 (9)	2.2 (8)	7.6 (4)	6.0 (7)	7.1 (5.5)
HEAD	29.6 (1)	6.7 (6)	16.0 (2)	3.8 (9)	4.4 (8)	15.6 (3)	9.6 (4)	8.2 (5)	6.2 (7)
HEARD	26.9 (1)	9.8 (4.5)	9.8 (4.5)	6.9 (7)	21.3 (2)	11.3 (3)	8.4 (6)	3.3 (8)	2.2 (9)

(A) Fullshifted (5th-order filters) with Feedback/Training [*Fullshifted/5th/+FT*]

Played	% Perceived (Rank)								
	HAD	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HEED	6.0 (4)	1.6 (7)	3.6 (6)	1.1 (9)	31.3 (2)	44.9 (1)	4.0 (5)	1.3 (8)	6.2 (3)
HID	22.2 (2)	3.6 (8)	11.6 (3)	1.8 (9)	5.8 (7)	34.0 (1)	7.3 (5)	7.8 (4)	6.0 (6)
HOOD	37.6 (1)	3.3 (8.5)	16.4 (2)	5.1 (6)	3.3 (8.5)	14.2 (3)	9.6 (4)	5.8 (5)	4.7 (7)
HUD	36.0 (1)	4.7 (6)	18.2 (2)	4.0 (7)	1.3 (9)	13.8 (3)	7.8 (5)	10.9 (4)	3.3 (8)
WHOD	33.1 (1)	11.6 (2)	9.1 (5)	6.2 (8)	3.6 (9)	10.7 (3)	7.8 (7)	9.8 (4)	8.2 (6)

Table 6

Vowel identification task (9AFC) confusion matrix for simulator users experiencing Unshifted (7th-order filters) stimulation with feedback/training (*Unshifted/7th/+FT*). Both Percent Correct (chance rate = 11%) and rank order of the correct selection (ideal = 1) are indicated.

Unshifted (7th-order filters) with Feedback/Training [*Unshifted/7th/+FT*]

	% Perceived (Rank)								
Played	**HAD**	**HAWED**	**HEAD**	**HEARD**	**HEED**	**HID**	**HOOD**	**HUD**	**WHOD**
HAD	70.0 (1)	2.1 (5.5)	14.2 (2)	6.2 (3)	2.1 (5.5)	1.2 (7.5)	1.2 (7.5)	2.9 (4)	0.0 (9)
HAWED	0.8 (7)	82.9 (1)	1.7 (6)	2.9 (3.5)	0.4 (8)	0.0 (9)	2.9 (3.5)	2.5 (5)	5.8 (2)
HEAD	8.3 (3)	1.2 (5.5)	71.7 (1)	4.2 (4)	1.2 (5.5)	12.5 (2)	0.0 (8.5)	0.8 (7)	0.0 (8.5)
HEARD	4.2 (5)	5.8 (4)	2.9 (6)	65.4 (1)	2.1 (8)	0.8 (9)	7.9 (3)	2.5 (7)	8.3 (2)
HEED	1.7 (5.5)	0.4 (7)	2.9 (3)	2.5 (4)	83.8 (1)	7.1 (2)	0.0 (8.5)	1.7 (5.5)	0.0 (8.5)
HID	3.3 (6)	1.7 (7.5)	15.0 (2)	4.6 (5)	0.8 (9)	56.2 (1)	6.2 (4)	10.4 (3)	1.7 (7.5)
HOOD	0.4 (9)	2.1 (6.5)	2.1 (6.5)	5.4 (3)	1.7 (8)	2.9 (5)	66.2 (1)	14.6 (2)	4.6 (4)
HUD	2.5 (7)	4.6 (3.5)	4.6 (3.5)	6.7 (2)	0.4 (9)	1.7 (8)	4.2 (5.5)	71.2 (1)	4.2 (5.5)
WHOD	0.0 (8.5)	2.1 (3)	0.8 (4)	0.4 (6)	0.4 (6)	0.0 (8.5)	6.7 (2)	0.4 (6)	89.2 (1)

Table 7

Rank order error and information theoretic analyses of confusion matrices. 2-norm rank order error metrics were assessed for significance of similarity between CI and simulator user confusion matrices. Information theoretic analyses included Mutual Information (MI) in bits/stimulus presented, percent of total information transferred (maximum = $\log_2 9$), and KL Distances between the CI user and simulator user distributions.

Testing Group/Condition				KL Distance	
	Feedback/Training?	$E_{squared}$	**MI *in Bits/Stimulus* (% total information)**	***CI vs. Simulator***	***Simulator vs. CI***
CI Users	*+FT*	*N/A*	1.42 (44.7%)	*N/A*	*N/A*
Unshifted (5th-order filters)	*+FT*	9.14**	2.08 (65.7%)	0.43	0.49
	-FT	11.04	0.98 (30.9%)	1.05	1.54
Fullshifted (5th-order filters)	*+FT*	9.53**	0.85 (27.0%)	0.63	0.70
	-FT	14.10	0.70 (22.0%)	1.22	1.71
Unshifted (7th-order filters)	*+FT*	9.09**	1.77 (54.3%)	0.43	0.28

** Confusion rank order matrix is statistically significantly similar to that of CI users, at the $p < 0.01$ level