fractory period. The dashed line denotes the stimulation function that would cause a nerve with an absolute refractory period of 0.5 ms (see Figure, A-C) or 0.7 ms (see Figure, D) to respond with the same PSTH.

In the Figure, A, the stimulation function is chosen to be a step function (remark 2). Here it can be seen that the response asymptotes to the constant value of \( s1(1 + s1) = s2(1 + bs2) \) as in equation 4.

In the Figure, B, D, the stimulation function \( s1 = \exp(\cos(2512t)) \), simulating a strongly phase-locked neural response to a 400-Hz signal. In the Figure, C, \( s1 = 2 \exp(\cos(2512t)) \), simulating a stronger signal, also at 400 Hz.

In the case of a nonstationary stimulation function, such as a strongly phase-locked nerve (see Figure, B-D), the shapes of the stimulation functions required to produce the same PSTH in nerves with different hazard functions are considerably different (especially the Figure, C). A larger stimulation (compare the Figure, B and C), or a larger difference between the hazard functions of the two types of stimulation (compare the Figure, D and B), results in greater differences between the shapes of stimulation functions that result in the same PSTH. This is why simply applying an electrical signal that is the same shape (even possibly resized) as an acoustic signal does not evoke the same PSTH as a nerve directly stimulated via the acoustic signal.

CONCLUSION

Improvements to the simulation of acoustically generated neural firing patterns by artificially evoked neural firing patterns should lead to improved speech perception of cochlear implant patients. To the knowledge of the authors, this paper is the first to present the idea of utilizing the point process models of auditory nerve response to design stimuli that cause this improvement. Work is under way to extend these ideas to a practical algorithm.

REFERENCES


CODING OF VOICE SOURCE INFORMATION IN THE NUCLEUS COCHLEAR IMPLANT SYSTEM

P. A. JONES, MCOGSC; H. J. MCDERMOTT, PhD; P. M. SELIGMAN, PhD; J. B. MILLAR, PhD

From the Human Communication Research Centre and the Department of Otolaryngology, University of Melbourne, Melbourne (Jones, McDermott), Cochlear Pty Limited, Sydney (Seligman), and Australian National University, Canberra (Millard), Australia.

Two studies are reported in which the effectiveness of explicitly coding voicing and fundamental frequency information for the Nucleus cochlear implant was investigated. In the first study, the voicing perception of a group of three experienced Multipeak users was evaluated when they were using Multipeak and a modified Multipeak in which the explicit fundamental frequency and voicing cues were eliminated and replaced with a 250-Hz constant rate of stimulation. The results of consonant and monosyllabic word tests showed that there was no significant difference in the subjects' ability to discriminate voicing.

In the second study, the ability of a group of five experienced users of the constant rate spectral maxima sound processor (SMSP) strategy to discriminate suprasegmental contrasts, the results of intonation, rowing stress, and question-statement tests showed that there was no significant difference between the scores recorded with these strategies. Since the temporal voicing cue is not a primary cue to voicing discrimination for Multipeak users, and the provision of an additional rate cue to the SMSP strategy does not improve SMSP users' ability to discriminate suprasegmental contrasts, the results of these studies indicate that in the cases investigated, the coding of voice source information by rate of stimulation does not significantly augment the cues present in the spatially distributed constant rate stimulation pattern.

INTRODUCTION

Voice source information is the acoustic information about the motion of the vocal folds during speech. In terms of speech perception, this voice source information contributes to the perception of the phonetic feature of voicing and to the perception of suprasegmental information that helps convey the higher-level rhythm and stress patterns of utterances and provides cues to word and syllable segmentation, speaker identity, question-statement contrasts, and the emotional state of the speaker. Since voice-source information is difficult to discern visually, the speech-coding strategies for the Nucleus implant have traditionally included an explicit encoding of this information. For example, within the Multipeak strategy, variation in stimulation rate is used to code both the fundamental frequency (F0), by the rate at which groups of spectral components are presented in voiced speech segments, and the phonetic feature of voicing, by the presence or absence of F0 periodicity in the stimulation pattern. The more recent spectral maxima sound processor (SMSP)2 does not include an explicit coding of this information — rather, the filter bank–derived estimates of the speech signal’s spectral components are presented at a constant rate of stimulation. Such a strategy successfully transmits voicing and suprasegmental information because the speech is highly redundant. In this paper the effectiveness of the explicit coding of voice source information by rate of stimulation is investigated by comparing coding strategies that do and do not maintain an explicit rate encoding of voice source information.
Two speech-coding evaluations were conducted. The first investigated the voicing perception of Multipeak users when the rate-encoded voice source information is removed and replaced with a constant stimulation rate. The second study, based on the finding that the standard 250-Hz carrier rate of the SMSP is insufficient to provide a full representation of F0 through the mechanism of amplitude modulation and the psychophysical adequacy of the rate domain to transmit F0 information, investigated the discrimination of suprasegmental contrasts by SMSP users when a full, rate-encoded representation of F0 is incorporated into this strategy.

**Method.** In the first study, the voicing perception of three experienced postlingually deaf adult Multipeak users was evaluated when they were using Multipeak and when they were using a constant rate strategy (CR) created by forcing Multipeak into its unvoiced mode and including a fixed delay recorded by a male talker and presented at a signal level of 65 dB A-weighted sound pressure level (dBA). Considered part of this study was to examine the effect of the standard voice-onset time and spectral weight, F1, F2, and F3, in which F4/F3 indicates that F4 is selected if F2 is in a high-frequency region and F3 otherwise, at a rate that varied by less than 10% around the average.

The data for this study were collected in four evaluation sessions for each subject, two with each strategy in an alternating sequence. Within each session the implantees responded to lists of /aCa/ consonants, each consisting of four repetitions of 12 consonants, and lists of monosyllabic words, each consisting of 50 unique CNC words. These lists were recorded by a male talker and presented at a signal level of 65 dB A-weighted sound pressure level (dBA). Considered across sessions and subjects, an unbalanced design was implemented, with a maximum amount of data collected in each fixed-length evaluation session. In total, the number of consonant lists presented in the Multipeak and CR conditions was 16 and 18, respectively, and 28 lists in each condition in the case of the word test.

**Results.** An analysis of variance of the percent correct scores for the medial consonant test (F[1.2] = 0.34, p = .567) and the various aspects of the word test (phonemes: F[1.8] = 1.66, p = .233; initial consonants: F[1.8] = 1.66, p = .756; and final consonants: F[1.8] = 2.96, p = .124) with subject and strategy as factors showed that there was no significant difference between the strategies and no significant interaction terms. The group mean information transmission analysis results for the medial consonant and the word tests are displayed in Figs 1 and 2, respectively. An analysis of variance of the individual information transmission scores for the medial consonant test (voicing: F[1.2] = 0.3, p = .872) and word test (initial consonant voicing: F[1.2] = 5.52, p = .143; final consonant voicing: F[1.2] = 0.65, p = .504) with the factors of strategy and subject showed that there were no significant differences between the strategies. Limiting the discussion to voicing, the results of both the consonant and word tests show that the CR stimulation pattern contains at least as much information about voicing as does the FO-based stimulation pattern. As a consequence of the multiplicity of cues to voicing that exist within the spatially distributed stimulation pattern — voice onset time and spectral weight, for example — the results suggest that explicitly coded temporal periodicity is not an essential cue to voicing discrimination for Multipeak users.

**FUNDAMENTAL FREQUENCY**

Method. In the second study the ability of a group of five experienced postlingually deaf adult users of the CR SMSP processor to discriminate suprasegmental contrasts was evaluated when they were using the SMSP and a modified SMSP strategy that included a rate-encoded representation of the F0 on the most apical stimulation channel. This dual-rate strategy, SMSP + F0, was implemented by using an MSP programmed with Multipeak as a source of interrupts to an SMSP such that the reception of an F0-based signal caused the immediate delivery of the most apical SMSP filter bank channel.

The data were collected in three evaluation sessions for each subject. Within each session, three speech tests were administered in each strategy condition: a novel randomization of the 50-item roving stress test (RST) and question-statement (Q/S) tests from the Speech Pattern and Contrasts
An analysis of variance of the percent correct trials were presented in both strategy conditions, but no extensions no feedback was given as to the correctness of the experimental environment and procedures. The speech material was spoken by a male speaker and presented via audiotape at a signal level that never exceeded 75 dBA. The first data collection session was preceded by an initial practice session used to familiarize the subjects with the experimental environment and procedures. The speech materials were presented in both strategy conditions, but no explicit training was undertaken. In the evaluation sessions, the subjects' responses.

Results. The group mean percent correct scores for the RST, Q/S, and INT tests are displayed in Fig 3. The chance-level scores for these tests are 33.3%, 50%, and 33.3%, respectively. An analysis of variance of the percent correct scores for the RST test (F[1,18] = 0.71, p = .410), Q/S test (F[1,18] = 0.01, p = .943), and INT test (F[1,73] = 0.12, p = .731) with subject and strategy as factors showed that there was no significant difference between the scores with the differing strategies and no significant interaction terms. There was no learning effect evident in any of the tests. It may be possible that more experience or a structured program of training might be necessary for the subjects to make use of this additional source of information or that the result represents a ceiling effect in F0 perception with the SMSP due to language knowledge or some more central psychophysical mechanism. However, the failure of the additional rate-encoded F0 information to increase the standard SMSP scores suggests that spectral amplitude and amplitude modulation variations are the primary carriers of stress and intonation information within the SMSP + F0 strategy and indicate that a single channel of rate-encoded F0 information is ineffective in enhancing SMSP performance in cueing stress and intonation.

CONCLUSION

Although these studies investigate voicing and F0 in different coding strategies, the common finding is that the rate domain is shown to be a secondary carrier of the information under investigation in each study. More formally, since the temporal voicing cue is not a primary cue to voicing discrimination for Multipeak users, and the provision of an additional rate cue to the SMSP strategy does not improve SMSP users' ability to discriminate suprasegmental distinctions, the results of these studies indicate that in the cases investigated, the coding of voice source information by rate of stimulation does not significantly augment the cues present within the spatially distributed CR stimulation pattern.

REFERENCES


RECOGNITION OF DUTCH PHONEMES BY COCHLEAR IMPLANT USERS WITH THE MULTipeAK STRATEGY

A. F. VAN OLPHEN, MD; J. E. VAN DIJK, PhD; M. C. LANGEREIS, MA; G. F. SMOORENBURG, PhD

From University Hospital Utrecht, Utrecht, the Netherlands.

INTRODUCTION

The purpose of this study was to investigate how the coding strategy of the Nucleus cochlear implant affects the recognition of Dutch phonemes. In a previous study a Kruskal analysis was used to create a multidimensional space in which the distance between two vowels was decreased with increasing...
Author/s:
Jones, P. A.; McDermott, H. J.; Seligman, P. M.; Millar, J. B.

Title:
Coding of voice source information in the Nucleus cochlear implant system

Date:
1995

Citation:

Persistent Link:
http://hdl.handle.net/11343/27435

File Description:
Coding of voice source information in the Nucleus cochlear implant system

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.