THE COCHLEAR IMPLANT: THE CODING OF AUDITORY INFORMATION

Mark W. White

Dept. of Otolaryngology, University of California, San Francisco

Investigators have compared psychophysical behavior in electrially stimulated patients with that in normal hearing subjects. Using current auditory coding models, investigators have found significant differences between model predictions and psychophysical measures. Such differences may help us develop useful extensions to current models of the auditory system.

Phase-locking to electrical stimuli: perceptual consequences

Electrophysiological recordings /1/ indicate that electrical stimulation of the auditory nerve elicits phase-locked responses to sinusoidal stimuli at frequencies up to 3 KHz. However, post-lingually deaf subjects implanted with similar intracochlear stimulating electrodes do not perceive changes in pitch for stimulus frequencies significantly above 300-600 Hz /2,3/. Also, frequency discrimination is extremely poor at stimulus frequencies above 300-600 Hz /2,3,4/. These results indicate that a simple periodicity-extraction mechanism is probably not operating for stimulus frequencies above about 300-600 Hz in these subjects. These results are consistent over a wide range of stimulus sites and over a wide range of stimulus intensities. By "simple periodicity-extraction mechanism", we refer to a neuronal frequency-extraction mechanism that measures the interspike intervals of a single neuron or group of neurons to estimate the stimulus spectrum /5,6/.

There are several possible explanations for the relative insensitivity of the implant recipient to differences in stimulus frequency above 300-600 Hz:

First, the normal nervous system may simply not be sensitive to such fine grain temporal information. If this is the case, the "place" theory would account for all pitch discrimination in both normal and prosthetic hearing. The classical "place" theory holds that the stimulus is completely coded in the discharge rate vs. cochlear place profile. However, this explanation is contrary to a considerable body of evidence supporting the perceptual significance of fine-grain temporal information (up to 3-4 KHz) in auditory nerve firing patterns /7-12/. For example, Viemeister's /11/ study indicates that behavioral performance in an intensity discrimination task (ie. tone in notch filtered noise) is relatively better than would be expected at the higher stimulus intensities if only rate-vs-cochlear-place information were used by the nervous system. Neurophysiological data /7,13/ indicate that a large percentage of eighth nerve fibers saturate in firing rate (or are nearly saturated) at the higher stimulus intensities, causing a broad plateau in the rate vs place profile. As a consequence, less information is available to the CNS at the higher stimulus levels <u>if</u> only rate-vs-place information is used. Such saturation occurs at the higher stimulus levels for both narrow-band and wide-band stimuli. Dye and Hafter /10/ have drawn a similar conclusion from their frequency discrimination experiments in which a wide-band masking noise was used. At the higher stimulus-masker intensities, they expected a poorer performance in the frequency discrimination task than at the moderate stimulus-masker intensities <u>if</u> only rate-vs-place information was being used. Only at stimulus frequencies of 3 KHz and above did the frequency discrimination performance decrease with stimulus-masker intensity. As a consequence, they concluded that a temporal mechanism is likely operating at frequencies up to 2-3 KHz, since such temporal information is not lost when overall firing rate saturates /7,8/. Sachs and Young /7,8/ studied eighth-nerve discharge patterns in cat as a function of stimulus intensity. They used a steady-state vowel stimulus and found that the vowel's spectrum was well represented in the temporal firing patterns distributed along the cochlea, even at the higher stimulus intensities. In contrast, the vowel's spectrum, when represented by the discharge-rate-vs-place profile, became severely distorted at the higher stimulus levels. The extreme distortion was primarily due to response rate saturation at the higher stimulus levels.

The second explanation for insensitivity to electrical stimulus frequency above 300-600 Hz is that these deaf subjects could have become insensitive to temporal features of eighth nerve excitation patterns as a result of prolonged auditory stimulus deprivation. Inotherwords, the CNS of these individuals may have changed as a result of such prolonged auditory stimulus deprivation.

Third, certain forms of auditory processing could be occurring which are not currently incorporated in classical "place" or "temporal" theories or in a more recent synchronization vs. place theory suggested by Sachs and Young /8/. The nervous system might be highly sensitive to the differences in discharge times between individual fibers. With electrical stimulation, it has not been possible to generate fine-grain interfiber discharge timing patterns corresponding to those generated by acoustic stimuli in the normal cochlea. In the normal cochlea, interfiber discharge timing is a function of the traveling-wave motion of the basilar membrane. Our data indicate that electrical excitation in the suprathreshold region around a scala tympani electrode is highly synchronous, with small latency increases in the liminal fringes. We expect that the "fringes of excitation" will extend more or less equally in all directions from the electrode site /14,15/. The implanted-subject's CNS might be able to make only limited use of the temporal-spatial information generated during electrical stimulation because it differs significantly from the temporal-spatial excitation patterns generated during acoustic stimulation. It may be that interfiber discharge timing sensitivities do exist in the nervous system and that they contribute significantly to spectral extraction mechanisms for both simple and complex stimuli /1/.

A more acceptable coding theory might be synthesized from Sach's and Young's periodicity-vs-place coding model and an inter-fiber timing model. Simple neural networks that perform a spatial-temporal cross-correlation across cochlear space might form a part of this extended model /1/.

An Improved Representation of Auditory Excitation: A Separate Excitation Pattern for each Spectral Band

After recording from a large population of eighth-nerve fibers, Young and Sachs used a "temporal synchronization index vs. cochlear place" function to generate a stable representation of the spectrum of their steady state vowel stimuli /16/. The authors refer to this synchronization index as the "average localized synchronization rate" or simply "ALSR". From each fiber's post-stimulus-time-histogram (PSTH), the magnitude of the fourier component at the fiber's characteristic frequency is calculated and displayed as a function of the fiber's cochlear origin. The fiber's characteristic frequency is that stimulus frequency at which the fiber is most sensitive. The "ALSR vs cochlear place" representation of auditory nerve activity has great appeal because the representation of the stimulus spectrum is relatively invariant across a wide range of stimulus intensities; whereas a "rate vs. cochlear place" representation of the spectrum severely degenerates at the higher stimulus intensities.

Although the ALSR representation of Sachs and Young is useful in preserving the stimulus spectrum, this representation fails to carry information pertaining to the absolute intensity of the stimulus (see Figure 12 of Young and Sachs /16/). In the ALSR representation, the synchronization index does not significantly change as the stimulus intensity of the steady-state vowel is increased. Perceptual measurements indicate that subjects are quite capable of utilizing absolute intensity information /17,18/. For example, a steady-state vowel presented at 40, 60, and 80 dB SPL will clearly be perceived as very different in loudness at each level. However, the ALSR representation of Sachs and Young would give little or no information about absolute intensity for those stimuli. With complex vowel-like stimuli, pulsation threshold measurements indicate that masking significantly increases with stimulus intensity /17/. This supports the view that the "internal representation" of the acoustic stimulus contains absolute intensity information.

The auditory system could obtain absolute intensity information and accurate spectral information from the more complete spatial-temporal representations displayed in figures 3-5 of Young and Sachs /16/. In these displays, the fourier magnitude for each spectral component of the stimulus is plotted as a function of cochlear place. The information contained in eighth nerve firing patterns is represented across three dimensions: spectral magnitude of the PSTH is plotted as a function of both frequency and cochlear place. In this representation, the formant peaks are clearly represented at all tested stimulus intensity levels - just as they are clearly represented in the two-dimensional ALSR representation. In the three-dimensional representation, there is information useful in estimating the absolute intensity of the stimulus. As the intensity of the vowel-like stimulus is increased, the synchronization to the first formant (F1) spreads to fibers originating from more more basal regions of the cochlea. To a somewhat lesser extent, this is also true for synchronization to the second formant (F2). In this 3-dimensional excitation pattern model, each spectral component or "band" has its own representation across the cochlea - as though each spectral band had a separate excitation pattern associated with it. However, each excitation pattern is not entirely independent of "interference" from the other spectral components of the stimulus. The excitation patterns representing the weaker spectral components are most affected by the stronger spectral components. Suppression of synchronization and suppression of average firing rate generate such cross-couplings between the spectral components. In the pulsation-threshold experiment, the observer adjusts the probe level so as to obtain the closest match to that specific excitation pattern associated with the masker component at or near the probe's frequency.

Florentine and Buus's analysis /19/ indicates that a "multi-band or multi-region" representation more accurately models "intensity discrimination vs. stimulus level" functions for a wide range of intensity discrimination experiments using masked tones. This finding further supports the concept that the auditory system is capable of decomposing the auditory nerve's fine-grain temporal patterns into components which represent the stimulus spectrum. And more importantly, this evidence supports the concept that this decomposition of multiple spectral components occurs over a large region of the cochlea. In the ALSR representation, each spectral component is represented by only one restricted region of the cochlea. Such a representation is less likely to yield models that accurately mimick measured intensity discrimination behavior. However, if spectral components are resolved at multiple sectors along the cochlea, better agreement with psychophysical intensity discrimination measures is likely.

References

/1/ Loeb GE, White MW, Merzenich MM. Spatial cross-correlation: A proposed model for acoustic pitch perception. Biol Cybernetics 1983, 47:149-164.

/2/ Simmons FB, White RL, Mathews RG, Walker MG. Pitch correlates of direct auditory nerve electrical stimulation. Ann Otol Rhinol Laryngol 1981; 90:suppl 82:15-18.

/3/ Merzenich MM, Michelson RP, Pettit CR, Schindler RA and Reid M. Neural encoding of sound sensation evoked by electrical stimulation of the acoustic nerve. Annals of Otolaryngology 1973; 82:486-503.

/4/ Herndon MK. Psychoacoustics and speech processing for a modiolar auditory prosthesis. 1981; Technical Report No. G906-5; Stanford Electronics Laboratories, Stanford University, Stanford, CA.

/5/ Green DM. An Introduction to Hearing, John Wiley and Sons, New York, 1976.

/6/ Goldstein JL, Srulovicz P. Auditory-nerve spike intervals as an adequate basis for aural frequency measurement. In Psychophysics and Physiology of Hearing. EF Evans and JP Wilson, eds., pp. 337-346, Academic Press.

/7/ Sachs MB, and Young ED. Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. J Acoust Soc Am 1979; 66:470-479.

/8/ Sachs MB, Young ED. Effects of non-linearities on speech encoding in the auditory nerve. J Acoust Soc Am 1980; 68:858-875.

/9/ Evans EF. Place and time coding of frequency in the peripheral auditory system: some physiological pros and cons. Audiology 1978; 17:369-420.

/10/ Dye RH, and Hafter ER. Just-noticeable differences of frequency for masked tones. J Acoust Soc Am 1980; 67:1746-1753.

/11/ Viemeister NF. Intensity discrimination of noise in the presence of band reject noise. J Acoust Soc Am 1974; 56:1594-1600.

/12/ Goldstein JL. The role of neural timing in auditory measurement of short-term fourier spectra. Presented at the midwinter meeting of the Association for Research in Otolaryngology 1979; St Petersburg Beach, FL (Abstract).

/13/ Lieberman MC. Auditory-nerve response from cats raised a low-noise chamber. J Acoust Soc Am 1978; 63:442-455.

/14/ Merzenich MM, White MW. Cochlear implant: the interface problem. In: Resnick J, Hambrecht T, eds. Functional electrical stimulation. New York: M Dekker, 1977:321-339.

/15/ White M. Design Considerations of a Prosthesis for the Profoundly Deaf. Doctoral Dissertation, University of California, Berkeley, 1978.

/16/ Young E, Sachs M, Representation of steady-state vowels in the temporal aspects of the discharge patterns of auditory-nerve fibers, J Acoust Soc Am 1979; 66:1381-1403.

/17/ Tyler R, Lindblom R, Preliminary study of simultaneous-masking and pulsation-threshold patterns of vowels, J Acoust Soc Am 1982; 71:220-224.

/18/ Stelmachowicz P, Small A, Abbas P, Suppression effects for complex stimuli, J Acoust Soc Am 1982; 71:410-420.

/19/ Florentine M, Buus S, An excitation-pattern model for intensity discrimination, J Acoust Soc Am 70:1646-1654.

Paper published in the Proceedings of the First Vienna International Workshop on Functional Electrostimulation Bioengineering Laboratory, University of Vienna, October 19-22, 1983.