On the Relationship Between Noise and Speech Recognition in Cochlear Implant Subjects: A Theoretical and Psychophysical Study

Leslie M. Collins[a,b], Roger L. Miller[b] and W. David Ferguson[a]
[a]Department of Electrical and Computer Engineering, Duke University
[b]Otolaryngology-Head and Neck Surgery, Department of Surgery, Duke University
Box 90291, Durham, NC, 27708-0291
Ph: 919-660-5260 FAX: 919-660-5293 Email: lcollins@ee.duke.edu

One factor that may impede speech recognition by cochlear implant subjects is that electrically stimulated nerves respond with a much higher level of synchrony than what is normally observed in acoustically stimulated nerves. Thus, the response patterns "seen" at higher processing centers are substantially different from those generated under normal acoustic stimulation. These differences may form the basis for a degradation of speech understanding since the patterns generated under electrical stimulation may be interpreted incorrectly by higher processing centers. Recent data from the implant research community has suggested that high rates of stimulation [1], addition of low levels of noise to the electrical stimulus [2-5], or elevated levels of "internal noise" [6] may provide some implanted individuals with improved speech recognition. In this paper, we present theoretical data that indicate that the addition of noise may provide improved neural responses to electrical stimulation along with psychophysical data that may indicate that individuals with higher levels of internal noise do in fact have higher levels of speech recognition.

INTRODUCTION

Impairment or disruption of the normal chain of mechanical to electrical transduction of information along the auditory pathway can result in sensorineural impairment or deafness. In the United States alone, there are 28 million individuals with hearing loss, and 1-2 million completely deaf individuals. As a result of increasing levels of noise in the modern environment, hearing loss occurs earlier in life and progresses more rapidly. In many instances of sensorineural deafness, the mechanical structures of the inner ear are irreparably damaged; however, the underlying neural population remains at least partially viable. In such cases amplification by hearing aids, which relies on the partial integrity of the mechanical structures in the ear, is an inadequate means of hearing assistance due to the severe amount of damage sustained by the inner ear. The insufficiencies of amplification are inherently addressed by the design of the cochlear implant. Cochlear implants stimulate the auditory system electrically, instead of acoustically, via an electrode array that is inserted into the inner ear, or cochlea. Thus, a cochlear implant replaces the mechanical to electrical conversion that occurs in a normal cochlea.

It is well established that cochlear implants restore some level of hearing to most deaf individuals [*cf.* 7]. Speech recognition abilities, however, vary widely across subjects and the mechanisms responsible for this variability are poorly understood. Factors such as electrode design, electrode placement, analog versus pulsatile stimulation, peripheral nerve survival, central auditory system integrity, speech-processing strategy, and complex and/or unexpected current paths within or around the target neural tissue may play a role in this variability. Unfortunately, these factors are complex and difficult to control or assess in implanted subjects. Research has indicated there are experimental, or psychophysical, measures that not only predict speech recognition performance, but also can be used to *improve* speech recognition by implanted subjects [8-12]. For example, reducing the interactions between electrodes has been shown to improve speech recognition performance [8,9,11]. These measures have, however, only addressed a subset of the possible factors that may influence speech recognition abilities.

Another factor that may impede speech recognition by cochlear implant subjects is the extreme degree of synchrony in the neural response to electrical stimulation [1-6, 13], a phenomenon not observed under acoustic stimulation [13,14]. Electrically stimulated nerves respond with a much higher level of synchrony than what is normally observed in acoustically stimulated nerves. Thus, the response patterns "seen" at higher auditory processing centers are substantially different from those generated under normal acoustic stimulation. These differences may form the basis for a degradation of speech understanding since the patterns generated under electrical stimulation may be interpreted incorrectly by higher processing centers. Two approaches have been proposed recently in the literature to address this issue. Rubinstein *et al.*, [1] have suggested that using a high rate conditioning stimulus may desynchronize the underlying neural population due to the refractory nature of nerve fibers.

Morse and Evans [2-5] have suggested that adding noise to a speech signal may decrease the synchrony of the neural response due to electrical stimulation, potentially leading to improved speech recognition for cochlear implant patients. In addition, Wilson et al., [6] have suggested that when neurons retain some level of stochastic activity (*i.e.* spontaneous activity remains) subjects may obtain higher speech recognition scores.

The phenomenon whereby additive noise, when presented at an optimal level, improves signal transmission is known as stochastic resonance (SR). Most SR work has focused on the addition of noise to a weak signal within the context of a nonlinear system in order to induce a "regularity" in the response pattern of the system which is related to the properties of the weak signal. SR has been applied to physical systems for many years, and evidence in support of SR has recently been found in biological systems [15]. Recent theoretical work investigated methods for optimizing the spectrum of the noise under normal (*i.e.* physiologically reasonable) stimulation of nerves [*e.g.* 16-18].

Currently there are no data concerning the optimal selection of the statistics of the noise process or its relationship to the speech signal for electrical stimulation of the auditory nerve. There are also no data that suggest how SR could be used to help alleviate factors known to deteriorate speech recognition performance, such as channel interactions. In this paper we present psychophysical data that indicate that *variability* in psychophysical threshold may be related to a form of internal stochastic resonance, as has been suggested previously [19-21], and that the data indicate that individuals with a higher level of naturally occurring SR have improved speech recognition performance. We also present theoretical results using a computational model to simulate neural response patterns to electrical pulse trains. Our preliminary data indicate that modulating the amplitude of the pulse train by low levels of white noise generates neural response patterns that are more similar to those elicited by acoustic stimulation than those elicited by un-modulated pulse trains. The patterns elicited also appear to de-correlate across fibers, an effect that has been hypothesized to be important for information coding under electrical stimulation [1-5].

PSYCHOPHYSICAL MEASURE OF INTERNAL NOISE: THRESHOLD VARIABILITY

Based on the SR hypothesis, and as suggested in [6] and [19-21], we hypothesized that some implanted subjects could have more "optimal" levels of internal noise than other subjects. In order to investigate this hypothesis, we measured *variability* in psychophysical threshold in a set of five implanted subjects. We chose a threshold task since it is simple to evaluate, and because internal noise levels will be at least partially captured by variability in threshold [*c.f.* 22].

Threshold data were obtained from five postlingually deafened adults. Four of the subjects were implanted with a Cochlear Corporation Nucleus 20+2L implant, and one was implanted with a Nucleus 20+2 implant. These devices are capable of bipolar as well as monopolar stimulation. Detailed information on these implants is given in [23]. A centrally located electrode, which was active in each subject's clinically programmed device and exhibited a normal dynamic range, was used to collect the threshold data. The extracochlear-indifferent electrode that was used for monopolar stimulation was common to both devices, and was the external receiver-stimulator plate electrode.

The stimuli presented consisted of single bi-phasic pulses and were delivered through an implanted receiver/stimulator driven by a Cochlear Corporation Spectra 22 processor connected to a Cochlear Corporation Dual Processor Interface. Threshold was measured at constant phase durations of 0.1, 0.2, and 0.4 ms. Phase duration ranges from 0.0192 – 0.4 ms during normal prosthesis operation [24]. The stimuli were presented via custom-designed software which interfaced to the implant through the DPI and Cochlear Corporation IF4 board. Threshold was measured using a 2-down 1-up adaptive 2IFC procedure [25] in which the stimulus amplitude was adjusted with a step size of one current level unit. A total of 14 reversals was required for each threshold determination, and the average of the last eight was used to estimate threshold. Four subjects completed 16 threshold measurements and one completed 15 measurements at each combination of phase duration and stimulation mode. The variability in threshold at each phase duration and stimulation mode was calculated from these data.

The standard deviations of the threshold data averaged across subjects are shown in Figure 1. The data exhibit an increase in standard deviation as phase duration increases. This trend occurs since the input/output functions that describe electrically stimulated neurons decrease in slope as phase duration increases [*e.g.*, 26]. In order to calculate correlation coefficients, the threshold variability for each subject was averaged across all phase durations for each stimulation mode. Variability in threshold was then correlated with various speech recognition measures that had been obtained clinically for these five subjects. Table I lists these correlation coefficients. Bold entries are significant at a level of 0.1 and italicized entries are significant at a level of 0.01.

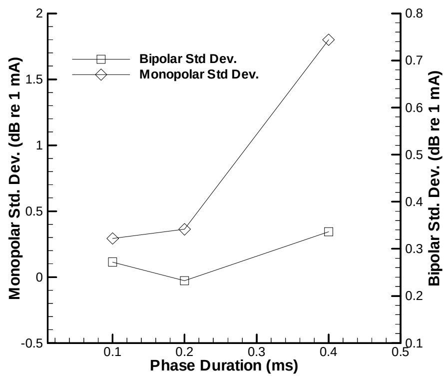

Figure 1. Standard deviation (dB) of threshold data averaged across subjects as a function of phase duration (ms).

Pearson	Monopolar	Bipolar
Average SR	*0.96*	**0.84**
BKB Sentences	**0.87**	0.63
Vowels	0.79	0.76
Consonants	0.53	0.76
NU6 – Words	0.54	0.42
NU6 – Consonants	0.75	0.33
Spearman	Monopolar	Bipolar
Average SR	0.70	**0.90**
BKB Sentences	**0.90**	0.70
Vowels	0.50	0.80
Consonants	0.30	0.50
NU6 – Words	0.67	0.56
NU6 – Consonants	0.40	0.30

Table I. Correlation coefficients between averaged threshold variability and average speech recognition (SR) and other speech recognition measures. Bold ($p < 0.1$) and italicized ($p < 0.01$) entries statistically significant.

All correlations were positive, and many were quite high. Significant Pearson correlations were observed between the average variance measured in monopolar stimulation mode and BKB sentences as well as the average of the speech recognition measures. The fact that the same pattern of results is observed with Spearman correlation data (Spearman correlations use relative rank) indicates that the correlation values observed are probably not an artifact of a small data set.

THEORETICAL EFFECTS OF NOISE ON THE RESPONSE TO ELECTRICAL STIMULATION

Analytic models can provide a useful framework for characterizing those parameters that are normally difficult to control in the physical system. The utility of a particular model depends on the extent to which predictions can be experimentally verified. White and colleagues [27-28] have proposed a stochastic neural-behavioral model that is computationally tractable. A primary feature of this model is the characterization of single-fiber input/output functions that are based on the phase duration of the eliciting stimulus with the slope of the function being inversely proportional to phase duration. At longer phase durations, the system behaves in a more stochastic fashion; thus the slope of the input/output function is more shallow. The modeled input/output relationship as a function of phase duration is consistent with that observed experimentally in electrically stimulated neurons [26,28].

In a previous theoretical study, we used signal detection theory to determine the optimal detector for threshold level detection of single pulse stimuli at the output of the model and then used this detector to predict the variability in threshold [19-21]. The results indicated that the model accurately predicted the trends in the experimental performance, and the predictions were consistent with human and feline psychophysical data (described above). In this study, we utilize the model to predict the changes in interval histograms for periodic electrical stimuli when low levels of Gaussian noise are used to modulate the amplitude of the stimulus. The model was modified slightly to include an exponential distribution for the actual spike times in response to an electrical pulse [29]. Interval histograms are a histogram of the inter-spike intervals elicited by a stimulus, and plot number of interval events versus the interval duration. Under acoustic stimulation at relatively low frequencies, neurons fire preferentially in one phase of the stimulus and may skip a random number of cycles between two successive firings. A typical interval histogram is shown in Figure 2. The characteristic multi-modal pattern associated with a phase-locked auditory nerve response is clearly evident [30].

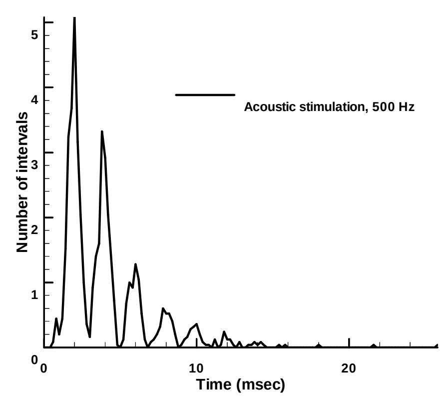

Figure 2. Interval histogram obtained by stimulating an auditory neuron acoustically with a 500 Hz sinusoid. Bin size is 0.2 msec and 1207 spikes were recorded.

Under electrical stimulation at moderate levels, neurons tend to respond on every cycle of the stimulus, as long as refractory effects are not in evidence. This corresponds to a uni-modal interval histogram where a response occurs in almost every interval [13,14]. In this study, we utilized the model described by Bruce *et al.* [28] to predict the response of a single neuron to a 500 Hz pulse train in both noise-free and low-level-noise conditions. The top portion of Figure 3 illustrates the interval histogram predicted by the model for the noise free condition. The bottom panel shows the response when a low level Gaussian noise has been used to adjust the amplitude of each pulse independently (SNR = 20dB). Superimposed in both panels (dashed lines) is the acoustic interval histogram, where all histograms have been normalized to a maximum of one for ease of comparison. Clearly, a more "natural", *i.e.* acoustically driven, response is obtained under the condition in which the stimulus is modified by a low level Gaussian noise.

The correlation between the interval histograms calculated from the responses predicted by the computational model and those measured under acoustic stimulation can be calculated over a range of SNRs. Figure 4 illustrates the relationship between noise level (which decreases as SNR increases) and the correlation between the theoretically predicted electrical interval histograms and the acoustic interval histogram. The noise level was not increased beyond an SNR of 10 dB as negative pulse amplitudes would have resulted. As was apparent in Figure 3, increasing the level of noise increases the similarity (as measured by a correlation) between the electrically induced theoretical interval histogram and the measured acoustic interval histogram. The data presented here indicates that a low level additive noise may indeed produce a more natural neural response.

As noted by Knight [31] and others [*e.g.* 1-5], cross-fiber de-correlation (in addition to increased correlation between electrical and acoustic interval histograms) is important for achieving optimal information transmission. In a modeling study by Morse et al. [3] using a population of neurons subject to analog (as opposed to pulsatile) electrical stimulation, results indicated that the spread of current in the cochlea can effectively provide a pseudo-independent noise source to each fiber, thus achieving *cross-fiber* de-correlation. The term pseudo-independent is used to indicate that the fiber responses were not *perfectly* uncorrelated. In fact, correlation between fiber responses decreased as the distance between the fibers increased when independent noise was applied to the analog stimulus presented to each electrode. While fibers in close proximity had correlation values, as measured by the Tanimoto measure, that approached unity, fibers spaced 2.5 mm apart had correlation values of approximately 0.3. (For reference, electrodes in the Nucleus device are spaced 0.75 mm apart). The Tanimoto measure of similarity can be used to determine the degree to which two binary sequences are similar [32]. A binary sequence is said to posses the i^{th} attribute if the i^{th} position equals one. The Tanimoto measure is then an indication of the common attributes shared by two binary sequences. This measure is well suited for calculations involving spike trains from nerve fibers and is used with the output of modeled auditory fibers to determine the similarity between the behavior of adjacent fibers. The Tanimoto measure of similarity is given by

$$s(x,z) = \frac{x'z}{x'x + z'z - x'z}$$

where *x* and *z* correspond to the two binary sequences.

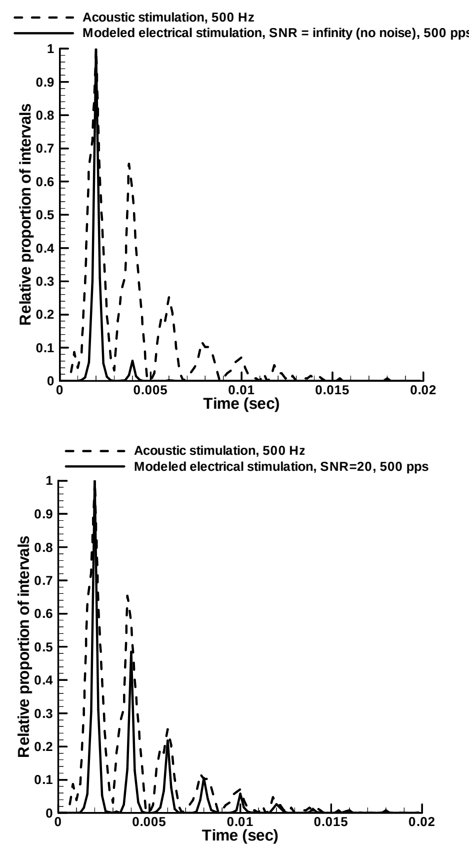

Figure 3. A comparison of the theoretically predicted interval histogram for an electrically stimulated neuron (500 pps stimulus) under conditions of no noise (top panel) and 20 dB SNR additive noise (bottom panel). Theoretical predictions are shown with solid lines. Histograms have been normalized to a maximum of 1. Normalized acoustic interval histogram has been superimposed for comparison and is shown with a dashed line.

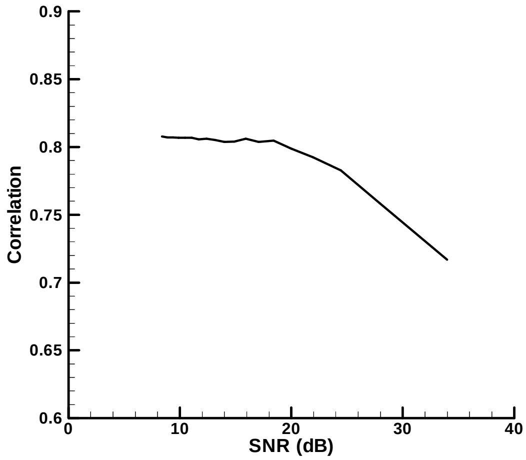

Figure 4. Correlation between the predicted electrical interval histograms and the acoustic interval histogram as a function of SNR

We used a similar approach with our computational model to determine if fibers subject to *pulsatile* electrical stimulation would show a similar pattern of de-correlation. In contrast to the approach described above, we applied a single stimulus to the center of the electrode array and computed the Tanimoto measure for 100 fibers spaced evenly across a 5 mm section of the cochlea that was centered at 15 mm. The simulations were run with the center fiber at an operating level (probability of firing) of 95%. The level of adjacent fibers was determined using a bipolar attenuation factor of 4 dB/mm [27,28]. The same noise parameters used to obtain the maximum correlation obtained in the previous study (Figure 4) were used. The computed Tanimoto measure of similarity is shown in Figure 5. The results indicate that the Tanimoto measure decreases to approximately 0.1 at a distance of 2.5 mm from the stimulus. These results are comparable to those presented by Morse et al. [3] for independent analog noise signals, although those signals were applied simultaneously to multiple electrodes.

CONCLUSIONS

SR research has focused on inducing a regular response from a nonlinear system that is correlated with a weak driving signal. The psychophysical data presented here indicate that a "noisier" system may indeed be related to better speech recognition. The theoretical study indicates that modulating the amplitude of a pulse train in a random fashion induces a more natural neural response pattern that may de-correlate spatially. Future work will involve theoretical work towards the optimization of the noise spectrum and analysis of alternative approaches to the goal of driving the neural response towards a more natural response. In addition, theoretical predictions will be compared to psychophysical and neurophysiological data.

ACKNOWLEDGMENTS

The modeling work that was a precursor to this effort [19-21] was performed in collaboration with Dr. Charlie Finley, Dr. Mark White, and Dr. David Smith, and their continued collaboration on this effort is greatly appreciated. The authors would also like to thank the subjects for their time and assistance. This work was supported by the Whitaker Foundation (RG98-0161) and NIH (R03 DC 03136).

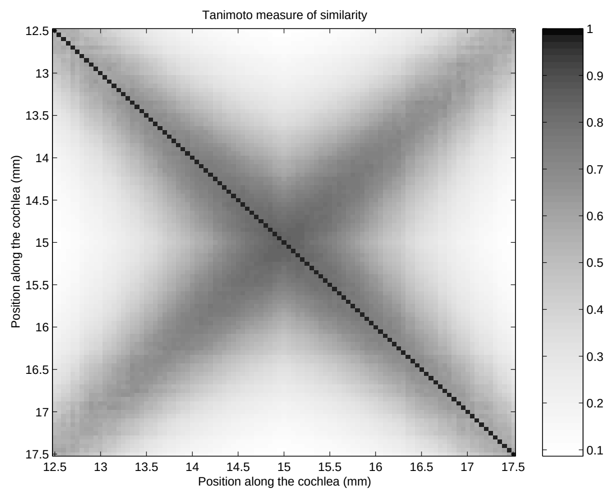

Figure 5. Tanimoto measure of similarity for a set of 100 fibers spaced in the +/- 2.5 mm around the center of the simulated cochlea.

REFERENCES

1. Rubinstein, J. T., Wilson, B. S., Finley, C. C., and Abbas, P. J., (1999). "Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation." Hearing Research, 127, 108-118.
2. Morse, R. P. and Evans, E. E., (1996). "Enhancement of vowel coding for cochlear implants by addition of noise." Nat. Med., 2(8), 928-932.
3. Morse, R. P., Meyer, G., and Evans, E. E. (1998). "Exploitation of current spread to obtain independent stimulation of cochlear nerve fibres by a cochlear implant." Br. J. Audiol., 32, 111-112.
4. Morse, R. P. and Evans, E. E., (1999a). "Additive noise can enhance temporal coding in a computational model of analogue cochlear implant stimulation." Hearing Research, 133, 107-119.
5. Morse, R. P. and Evans, E. E., (1999b). "Preferential and non-preferential transmission of formant information by an analogue cochlear implant using noise: the role of the nerve threshold." Hearing Research, 133, 120-132.

6. Wilson, B., Finley, C. Lawson, D., Zerbi, M., and van den Honert, C. (1997). "Speech processor for auditory prostheses. Seventh quarterly progress report, NIH Contract N01-DC-5-2103." Neural Prosthesis Program, NIH, Bethesda, MD.
7. Dawson, P. W., Blamey, P. J., Dettman, S. J., Barker, E. J., and Clark, G. M. (1995). "A clinical report on receptive vocabulary skills in cochlear implant users." Ear & Hearing 16, 287-294.
8. Collins, L. M., T. A. Zwolan and G. H. Wakefield (1997). "Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects." J. Acoust. Soc. Am. 101(1), 440-55.
9. Zwolan, T. A., L. M. Collins and G. H. Wakefield (1997). "Electrode discrimination and speech perception in postlingually deafened adult cochlear implant subjects." J. Acoust. Soc. Am. 102, 3673-3685.
10. Throckmorton, C. S. and Collins, L. M., (1999c). "The Effects of Additive Gaussian Noise on Speech Recognition in Cochlear Implant Subjects." 1999 Conference on Implantable Auditory Prostheses, September, 1999.
11. Henry, B. A., C. M. McKay, H. J. McDermott and G. M. Clark (1997). "The relationship between speech information perceived by cochlear implantees in different spectral regions and electrode discrimination." 1997 Conference on Implantable Auditory Prostheses, Asilomar Conference Center, Pacific Grove, California.
12. Chatterjee, M. and Shannon, R. V. (1998). "Forward masked excitation patterns in multielectrode electrical stimulation." J. Acoust. Soc. Am. 103, 2565-72.
13. Hartmann, R., Topp, G., Klinke, R., (1984). "Discharge patterns of cat primary auditory fibres with electrical stimulation of the cochlea." Hear. Res. 13, 47-62.
14. van den Honert, C. and P. H. Stypulkowski. (1984). "Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings." Hearing Res. 14, 225-243.
15. Weisenfeld, K. and Moss, F. (1995). "Stochastic resonance and the benefits of noise: From ice ages to crayfish and SQUIDs." Nature, 373, 33-36.
16. Gingl, Z., Kiss, L., and Moss, F., (1995). "Non-Dynamical Stochastic Resonance: Theory and Experiments with White and Arbitrarily Colored Noise." Europhys. Lett., 29(3), 191-196.
17. Nozaki, D., Mar, D., Grigg, P. and Collins, J. (1999). "Effects of Colored Noise on Stochastic Resonance in Sensory Neurons." Physical Review Letters, 82(11), 2402-2405.
18. Pei, X., Bachmann, K., and Moss, F., (1995). "The detection threshold, noise and stochastic resonance in the Fitzhugh-Nagumo neuron model." Physics Letters A 206, 61-65.
19. Ferguson, W. D., Collins, L. M., Finley, C. C., Pfingst, B. E., and Smith, D. W., "Prediction of the variance in behavioral thresholds using a stochastic model of electrical stimulation", Midwinter Research Meeting, Association for Research in Otolaryngology, February, 1998.
20. Ferguson, W. D., Collins, L. M., Finley, C. F., and Smith, D. W., (1999). "Prediction of psychophysical data using a stochastic model of electrical stimulation." 1999 Conference on Implantable Auditory Prostheses, September, 1999.
21. Collins, L.M., Ferguson, W. D., Finley, C. C., and Smith, D. W., "The relationship between variability in cat behavioral thresholds and predictions from a stochastic model of electrical stimulation", in preparation for submission to J. Acoustical. Society of America.
22. Green, D. and Swets, J. (1974). Signal Detection Theory and Psychophysics. Kreiger, New York.
23. Zwolan, T. A., Kileny, P. R., Ashbaugh, C. and Telian, S. A. (1996). "Patient performance with the Cochlear Corporation 20+2 implant: bipolar versus monopolar activation," Am. J. Otol. 17, 717-723.
24. Skinner, M. W., Holden, L. K., Holden, T. A., Dowell, R. C., Seligman, P., Brimacombe, J. A., and Beiter, A. L., (1991). "Performance of post-linguistically deaf adults with the Wearable Speech Processor (WSP III) and Mini speech Processor (MSP) of the Nucleus multi-electrode cochlear implant." Ear Hear. 12(1), 3-22.
25. Levitt, H. (1971). "Transformed up-down methods in psychoacoustics." J. Acoust. Soc. Am., 49, 467-477.
26. Shepherd, R. and Javel, E. (1997). "Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status." Hear. Res. 108, 112-144.
27. White, M. W., Finley, C. C. and Wilson, B. S. (1987). "Electrical stimulation model of the auditory nerve: stochastic response characteristics," Proc. Ninth Ann. Conf. IEEE Eng. Med. Biol. Soc. 4, 1906-1907.
28. Bruce, I. C., Irlicht, L. S., White, M. W., O'Leary, S. J., Javel, E., Dynes, S., and Clark, G. M. (1999). "A stochastic model of the electrically stimulated auditory nerve: single-pulse response." IEEE Trans. Biomedical Eng., In Press.
29. Rubinstein, J. T., (1991). "Analytical theory for extracellular electrical stimulation of nerve with focal electrodes 2: Passive myelinated axon". Biophys J. 60, 538-555.
30. Rose, J.E., Brugge, J.F., Anderson, D.J. and Hind, J.E. (1967). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey." J Neurophysiol 30: 769-793.
31. Knight, B. W., (1972). "Dynamics of Encoding in a Population of Neurons", The Journal of General Physiology, 59, p. 734-766.
32. Tou, J. T., and Gonzalez, R. C. 1974. Pattern Recognition Principles. Addison-Wesley Publishing Company, Reading, MA.