# MACHINE LEARNING AND THE AUDITORY NERVE

**Jeff Bondy[1], Ian C. Bruce, Sue Becker, and Simon Haykin**
[1]Dept. of Electrical and Computer Engineering, McMaster University, 1280 Main St. W., Hamilton, Ontario, Canada, L8S 4K1. email: jeff@soma.ece.mcmaster.ca

## 1. INTRODUCTION

The goal to our research is to produce a hearing-aid algorithm that would enhance the intelligibility of the hearing impaired in noisy conditions. For this we have been using machine learning techniques to derive the important statistical quantities describing the differences between the normal and impaired auditory systems auditory nerves. This leads to a technique we call Neurocompensation because it is essentially trying to re-establish the neural response of the auditory system after hair cell loss.

The main advantage from using the neural coding of the auditory nerve is it is the closest physical variable after the impairment. By encompassing the impairment, theoretically, the resulting algorithms should be better because they are free from simplifying assumptions. For example, by basing hearing-aid processing design on the audiogram alone, an implied assumption is that the loss of cochlear gain is the only important variable. This does not encompass the large differences in temporal and spectral properties between the normal and impaired auditory system.

## 2. METHOD

To apply machine learning to the auditory system we have a four component model:

1. A model of the normal auditory system up to the auditory nerve.
2. A model of the impaired auditory system that encompasses the hearing impairment, in this case we specifically look at processing lost with hair cell damage.
3. A processing block to train, a surrogate attempting to replicate the missing processing of the damaged system.
4. An error metric that is an intelligibility predictor, based on distortions to the auditory nerve.

An acoustic input is processed by the normal model to come up with a control signal, while the processing block preprocesses the same input before being passed to the impaired block which comes up with a distorted auditory response signal. The control and distorted signal are then compared to calculate how intelligible the distorted signal is. By maximizing the intelligibility of the distorted signal by training the parameters of the preprocessing block we are really building a hearing-aid processor the restores what is important on the auditory nerve.

## 3. DISCUSSION

The two auditory models were provided by Bruce et al. (2003). The error metric was first suggested in Bruce et al. (2002), and then improved into its useful form in Bondy et al. (2004). The metric is largely based on deriving a neural equivalent to the Articulation Index (AI); in fact the validation directly parallels Steeneken (1992), one of the modernizers of the AI, who suggested the now widely adopted Speech Transmission Index (STI). The error metric we derived is called the Neural Articulation Index (NAI) and showed a deviation of empirical intelligibility from prediction of about 8% versus the standard AI's deviation of 10%. This error was on a nonsense syllable, rippled filtering condition test, which has historically proven very hard to predict.

The last piece of the puzzle was the processing block. Several possible blocks and their responses to hearing impairment are given in Bondy et al. (2004b). The largest success was in predicting linear hearing-aid strategies. The Neurocompensator strategy showed that the NAL-R (Dillon, 2001), a strategy very close to returning optimal intelligibility, is in fact minimizing the differences between the normal auditory nerve representation and the impaired one. We show the extremely close correlation of the NAL-R with the values calculated by the Neurocompensator approach. An example of optimizing the gain per dB of threshold shift through the Neurocompensator method is shown in Figure 1.

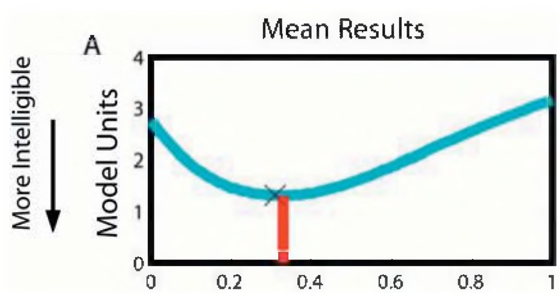


Fig. 1. The x-axis the gain in dB for every dB shift from an audiogram mimicking the NAL-R's gain ratio of 0.31. The X marks the NAL-R prescribed gain ratio, while the vertical bar is the Neurocompensator optimized value 0.32.

The Neurocompensator also has the added advantage of being able to be tailored to an individual, by maximizing the gains and shape for a specific hearing loss. It is also hoped that it could deal with the variance in intelligibility for people with similar audiograms, which can be derived from different ratios of inner and outer hair cell loss.

While the Neurocompensator did very well predicting normal, linear fitting strategies, it had problems when trying to derive the optimal non-linear parameters. We found no set of time constants, compression knees and rates or channel setups that produced a higher intelligibility prediction (Bondy et al. 2003). To try to come to terms with this we optimized the gain in different time windows of speech. The best possible values produced a scatter "function" versus input RMS; there was no connection between the input power and gain requirements, past the mean, "linear" response. Our initial trial listening tests produced quality and intelligibility deficits, subjects often complaining about odd artifacts. We decided to look closer at the differences between the normal and impaired response. In general we say the mode rate may be the same, but the shape of the activity was very different. Generally the mode discharge rate between the Normal auditory response and the impaired auditory response after NAL-R preprocessing looks similar. There are some points that the normal discharge rate is larger than the impaired response and some points when it is much smaller.

In describing the statistical differences between the normal and impaired auditory nerve responses we saw that there was a loss of contrast between different auditory landmarks. The well known lessening of the suppression effect in the impaired ear correlates neighbouring frequency bands, while the healthy cochlea produces a negative correlation. The loss of suppression reduces the peak-to-trough ratio of spectral information. Similarly, in the time domain, the loss of adaptation reduces the temporal contrast as shown in Figure 2.

The important peak-to-trough ratio of the impaired response is not as large as the normal response, it is not in the same place, cresting at different times over frequency and the adaptation response is much wider.

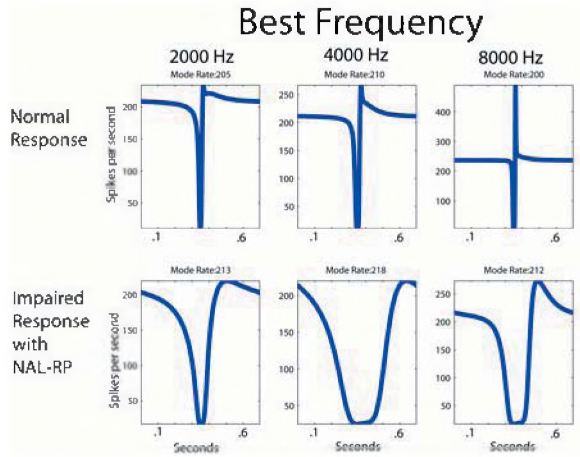


Fig. 2. Discharge rates for the normal and impaired auditory system with a 2 Hz 100% AM modulation applied to a tone at the best frequency.

That really defines an interesting formulation for the machine learning problem to address. Unlike our previous attempts which really were learning an average code and maybe not the important aspects of the AN responses, we are now looking at algorithms to reestablish suppression and onset/offset information in the impaired ear. We hope to address the fundamental question of segmentation enhancement for the hearing impaired. It is hoped that better segmentation will lead to more normal streaming, allowing the hearing-aid user the ability to unmask spectrally and temporally as well as a normal hearing person.

## REFERENCES


Bondy, J., Bruce, I.C., Rong Dong, Becker, S., and Haykin, S., (2003) "Modeling intelligibility of hearing-aid compression circuits", Signals, Systems & Computers, 2003 The Thirty-Seventh Asilomar Conference on ,Volume: 1 , 720-724.

Bondy, J., Bruce, I.C., Rong Dong, Becker, S., and Haykin, S., (2004) "Predicting Speech Intelligibility from a Population of Neurons", in Advances in Neural Information Processing Systems 16, Sebastian Thrun and Lawrence Saul and Bernhard Schoelkopf eds., MIT Press.

Bondy, J., Becker, S., Bruce, I.C., Trainor, L., and Haykin, S., (2004b) "A novel signal-processing strategy for hearing-aid design: Neurocompensation", Signal Processing, **84**(7), 1239-1253.

Bruce, I.C., Bondy, J., Becker, S., and Haykin, S., (2002) "A Physiologically Based Predictor of Speech Intelligibility", International Conference on Hearing Aid Research, 2002.

Bruce, I. C., Sachs, M. B., and Young, E. D. (2003). An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. J. Acoust. Soc. Am. 113, 369–388.

Dillon, H. (2001). Hearing Aids. New York: Thieme Medical Publishers.

Steeneken, H.J.M. (1992) "On measuring and predicting speech intelligibility", University of Amsterdam, 1992.