30th Annual International IEEE EMBS Conference
Vancouver, British Columbia, Canada, August 20-24, 2008

Hearing aid gain prescriptions balance restoration of auditory nerve mean-rate and spike-timing representations of speech

Faheem Dinath and Ian C. Bruce, *Senior Member*

***Abstract*— Linear and nonlinear amplification schemes for hearing aids have thus far been developed and evaluated based on perceptual criteria such as speech intelligibility, sound comfort, and loudness equalization. Finding amplification schemes that optimize all of these perceptual metrics has proven difficult. Using a physiological model, Bruce et al. [1] investigated the effects of single-band gain adjustments to linear amplification prescriptions. Optimal gain adjustments for model auditory-nerve fiber responses to speech sentences from the TIMIT database were dependent on whether the error metric included the spike timing information (i.e., a time-resolution of several microseconds) or the mean firing rates (i.e., a time-resolution of several milliseconds). Results showed that positive gain adjustments are required to optimize the mean firing rate responses, whereas negative gain adjustments tend to optimize spike timing information responses. In this paper we examine the results in more depth using a similar optimization scheme applied to a synthetic vowel /ε/. It is found that negative gain adjustments (i.e., below the linear gain prescriptions) minimize the spread of synchrony and deviation of the phase response to vowel formants in responses containing spike-timing information. In contrast, positive gain adjustments (i.e., above the linear gain prescriptions) normalize the distribution of mean discharge rates in the auditory nerve responses. Thus, linear amplification prescriptions appear to find a balance between restoring the spike-timing and mean-rate information in auditory-nerve responses.**

I. INTRODUCTION

Amplification prescriptions have a foundation in early empirical studies showing that the most comfortable gain at a particular frequency equals approximately half the hearing threshold shift at the same frequency. This is referred to as the "half-gain" rule. That is, for every 1 dB increase in hearing threshold, the most comfortable gain is increased by 0.5 dB [2]. Today, popular linear hearing aid prescriptions, including the National Acoustic Laboratories' Revision 1 (NAL-R) and the Desired Sensation Level (DSL) prescriptions, are based on modifications of the half-gain rule and on judgments of speech intelligibility, sound comfort, and loudness equalization [3].

The goal of this study is to find and analyze optimal single-band gain adjustments around the NAL-R and DSL prescribed gains by using the neural representation of speech rather than using perceptual feedback. The following work uses single-band gain adjustments, rather than multi-band gain adjustment in order to reduce the number of working variables, and therefore the complexity of adjustment. We will start by describing the auditory-periphery model in this study and show how it computes the speech neurogram. This is followed by a brief overview of the gain optimization strategy reported in [1] and then a detailed analysis of the optimal gain adjustments for a synthetic vowel /ε/ with discussion of the results.

This work is funded by NSERC Discovery Grant 261736.

F. Dinath (dinathf@mcmaster.ca) and I. C. Bruce (ibruce@ieee.org) are with the School of Biomedical Engineering, McMaster University, Hamilton, ON, L8S 4K1, Canada.

II. METHODS

A. Models

The auditory-periphery model used in this study was the cat auditory nerve model developed and validated against cat physiological data by Zilany and Bruce [4], [5]. The model describes the auditory pathway from the middle ear through to the auditory nerve. An outer ear filter [6] was utilized before the middle ear filer in this study.

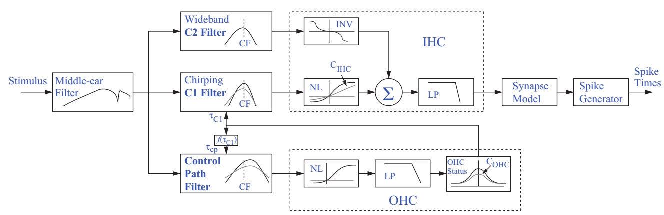


Fig. 1. Zilany and Bruce cat auditory nerve model. Reprinted from [4] with permission.

Input to the middle ear is an arbitrary sound waveform with instantaneous pressures, in Pascal, sampled at a rate of 100 kHz. In response, the model derives the spike timing information for an auditory nerve (AN) fiber with a specific preferred or *characteristic frequency* (CF).

Model parameters C_{IHC} and C_{OHC}, which control the level of inner and outer hair cell impairment, respectively, can be adjusted to provide a desired hearing threshold shift at a specific CF [4]. A C_{IHC} or C_{OHC} of 0 produces full impairment, whereas 1 produces normal function.

B. Amplification Schemes

In this study we use two popular and well established linear hearing aid amplification schemes. The first, the Australian National Acoustics Laboratories' Revision 1 hearing aid amplification scheme [7]. The NAL-R originated from improvements over the amplification scheme developed by Byrne and Dillon [8] and has kept true to the original aim of maximizing speech intelligibility for moderate sound pressure levels by equalizing perceived loudness over the frequency range important for speech (250–8000 Hz) [3]. NAL-R prescribes gain in terms of insertion gain (IG), that is, the gain provided by the hearing aid above the gain normally supplied by the outer ear's natural amplification [7].

Another widely used prescription in the Desired Sensation Level amplification scheme. Although the original version of the DSL algorithm was aimed for use in children, continued development and research, however, has expanded its role for use with adults. The DSL prescription differs from NAL-R in that it does not try to make speech equally loud, but rather comfortably loud. The desired sensation gains are based on data describing the most comfortable hearing level associated with speech presentation level for different hearing impairments [9]. In this scheme, gain provided is expressed in terms of the real ear aided gain (REAG), that is, the total gain supplied by the hearing aid.

C. Neural Output

We visualize the neural representation of speech in the auditory nerve by a "neurogram". A neurogram is similar to the spectrogram, except that it displays the neural response as a function of CF and time. Neurograms can include the spike timing information of the neural responses by maintaining a small time bin size (Fig. 2D), or the spike timing information can be excluded by computing the moving average with a window of several milliseconds to give only the average discharge rate as a function of time (Fig. 2C).

In this study, 30 CFs spaced logarithmically between 250 and 8000 Hz were modelled. The neural response at each CF is composed of 50 AN fiber responses. In accordance with Liberman and Kiang [10], 60% of fibers were chosen to be high spontaneous rate (>18 spikes/s), 20% medium (0.5 to 18 spikes/s), and 20% low (<0.5 spikes/s).

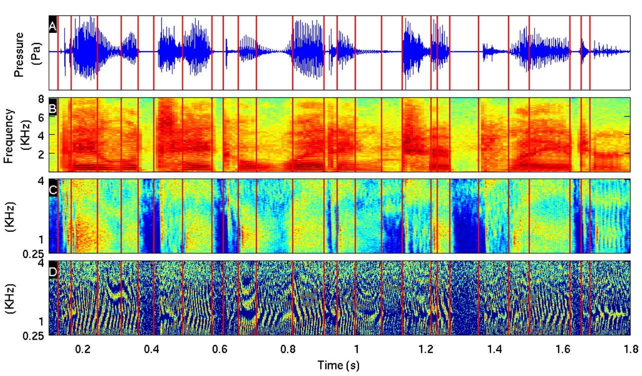


Fig. 2. An example sentence from the TIMIT database and the corresponding spectrogram and neurograms. (A) Time-domain pressure waveform; (B) Spectrogram; (C) Neurogram based on the average discharge rate; (D) Neurogram based on the spiking timing information. Phoneme boundaries are indicated by the vertical red lines.

D. Summary of the gain optimization strategy

The gain optimization strategy shown in Fig. 3 compares the neural responses to speech sentences on a phoneme-by-phoneme basis for the impaired and normal models [1]. In order to avoid the confounding and complicating effects of compression attack and release times, a constant gain adjustment was applied for the duration of each phone, using the known phone boundaries from the TIMIT transcriptions.

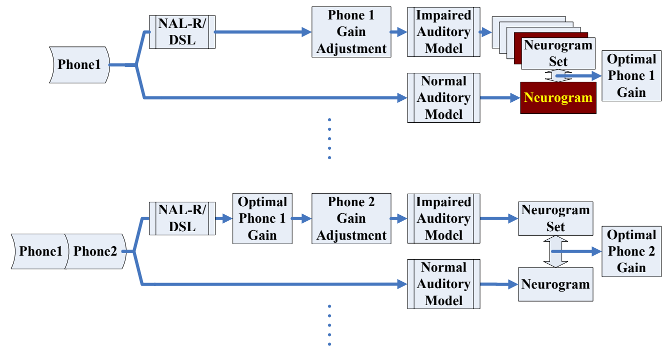


Fig. 3. Flow diagram of gain adjustment strategy. Reprinted from [1] with permission.

The strategy begins by passing the first phone through the normal model to derive the normal neurogram. In the impaired pathway, the phone is passed though either the NAL-R or DSL amplification prescription before a single-band gain adjustment is applied. Gain adjustments range from -40 to $+40$ dB in 5 dB increments resulting in 17 uniquely amplified phones. The phones are passed through the impaired model, producing a set of 17 neurograms. The gain adjustment that minimizes the mean absolute error between the normal and impaired neurograms is deemed the optimal gain adjustment for that phone.

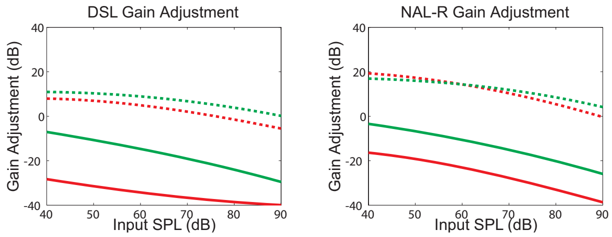
For each amplification prescription, optimal gain adjustments were found by comparing either the neurograms with spike timing information or the average discharge rate neurograms. The second and all subsequent phones are analyzed in the same manner as the first, however, due to adaptation in the auditory-periphery model, all prior phones are prepended. The range of gain adjustments is applied only to the current phone and all previous phones are amplified with their optimal gain adjustments.

E. Optimal gains adjustments for TIMIT sentences

The results in Fig. 4 were generated using test sentences from the TIMIT database. Two sentences were presented to the auditory model with mild hearing impairment and another two were presented to a model with mild-to-severe hearing impairment. Each sentence was delivered to the gain optimization strategy at 3 different mean sound pressure levels, thereby providing a number of diverse phone types and sound pressure levels for examination. The curves in Fig. 4 are quadratics fits to the results given in [1].

The results from [1] suggest that positive gain adjustments above the prescribed gains better restore the mean discharge rate representation of speech. However for the fine timing neurogram, the gain optimization strategy suggests negative gain adjustments. The discrepancy in optimal gain adjustments between the two neural representations is not clear and warrants further investigation. In particular, it is difficult from visual inspection to determine in exactly what ways gain adjustments affect the spike-timing neurograms.

In the following analysis, we used a synthetic /ɛ/ phone of known characteristics (100Hz fundamental frequency and



(a) (b)

Fig. 4. Optimal gain adjustments versus phoneme input sound pressure for mild (green) and severe (red) hearing loss types. Gain adjustments using the average discharge rate neurogram are represented with dashed lines and spike timing neurogram with solid. Optimal gains using the DSL amplification scheme is shown in (a) and NAL-R is shown in (b).

formants located exactly at 500, 1700, and 2500 Hz) to explore what changes are happening within the neurogram.

A test sentence containing 40 identical /ɛ/ phones over a duration of 400 ms with an SPL between 40 and 100 dB was passed through either a mildly or severely impaired or normal auditory periphery model. In the mildly impaired case, the sentence was processed by either the NAL-R or the DSL amplification scheme with overall gain adjustment from 40 dB below to 40 dB above the prescribed gains in 5 dB steps, and in the severe case, from -60 dB to 20 dB. Neural responses were obtained using the cat auditory-periphery model of Zilany and Bruce [4], [5].

F. Optimal gain adjustments for the synthetic /ɛ/ vowel using the mean discharge rate neurogram

We calculated histograms of occurrences of different mean spike rates at each CF over the duration of the mean-rate neurogram. Figure 5 shows the histogram response to an /ɛ/ phone presented at low to moderate SPL. With greater impairment of the auditory periphery, the distinctive features in the count distribution tend to fall, producing a uniform rate distribution across CFs in the impaired region. This abnormal distribution is brought closer to normal with amplification prescriptions.

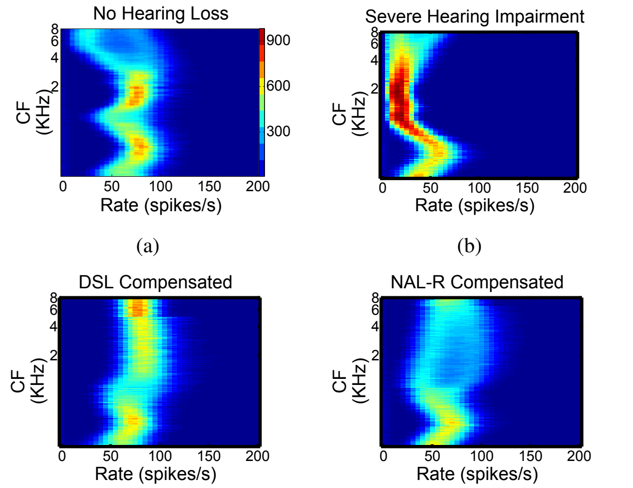
G. Optimal gain adjustments for the synthetic /ɛ/ vowel based on measures of synchronization to formant frequencies

Analysis in this section was based on measurements of synchrony to vowel frequency components and the relative degree of phase shift for a particular frequency in a population of fibers with CFs around that frequency.

Synchrony of an AN fiber to a particular frequency component in periodic speech is measured using the power ratio [see Fig. 6(a)], defined as the sum of power in the AN fourier response R at the frequency f_x and its harmonics, divided by the total power in the response:

$$\mathcal{PR}(f_x) = \frac{\sum_{m=1}^{u} R^2(m \cdot f_x)}{\sum_{n=1}^{v} R^2(n \cdot f_0)} \quad (1)$$

with $u < 4$, $u \cdot f_x \leq$ 5kHz, $v = 50$, and $v \cdot f_0 \leq$ 5kHz, where f_0 is the fundamental frequency [11]. Because phase locking in cats is not observed above 5kHz, the sums are limited to frequency components below 5kHz.

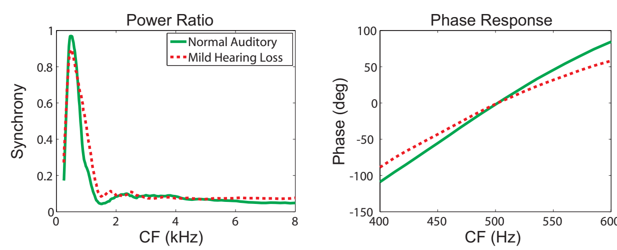


(a) (b)

(c) (d)

Fig. 5. The normal and impaired histogram of the average discharge rate neurogram are shown in subfigures (a) and (b) respectively. (c) and (d) show the DSL and NAL-R compensated response.

The phase response is a measure of the phase of the synchronized response at a particular frequency in fibers with CFs in the neighborhood of the frequency of interest. Phase is represented relative to the phase of the fiber with CF equal to the frequency of interest and is only measured for fibers with power ratios greater than 0.1 [see Fig. 6(b)].



(a) (b)

Fig. 6. Power ratio response to f = 500Hz (a) and phase response at f = 500Hz around CF = 500Hz (b) for a normal and mildly impaired auditory periphery using a synthetic vowel /ɛ/.

III. Results

Optimal gain adjustments for the average discharge rate comparison, shown in Fig. 7, were found by minimizing the mean absolute difference between the normal and impaired histograms of mean spike rates. In compensating for mild and severe hearing impairment, the DSL and NAL-R algorithms are fairly effective at restoring the impaired rate distribution count to normal. Consistent with the results for the TIMIT sentences, there seems to be a need for a slightly positive gain adjustment, particularly when dealing with the NAL algorithm since it gives less overall gain than DSL.

In our analysis of synchrony and phase responses, both mild and severe hearing impairment requires negative gains in both amplification schemes (Fig. 8), precisely what the gain optimization strategy for the TIMIT database suggests

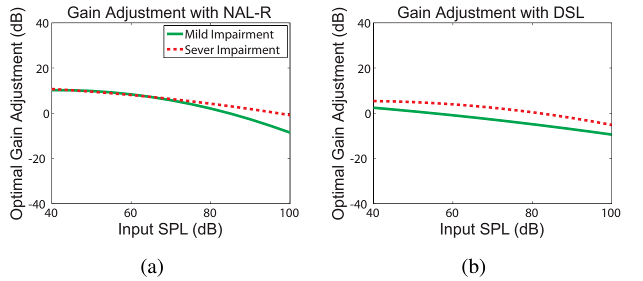


Fig. 7. Optimal gains using the average discharge rate neurogram and a synthetic vowel /ɛ/. NAL-R amplification scheme is shown in (a) and DSL shown in (b). Curves here are 3rd order polynomial fits.

(Fig. 4). When the neurogram includes spike timing information, the individual power ratios and phase responses (not shown) of the impaired auditory periphery broaden and flatten with respect to the normal auditory periphery with increasing gain adjustments, a phenomenon known as spread of synchrony [11]. With higher gains, phase deviations, in addition to spread of synchrony, become a driving factor in the error metric, leading to the requirement of negative gain adjustments for the fine timing neurogram.

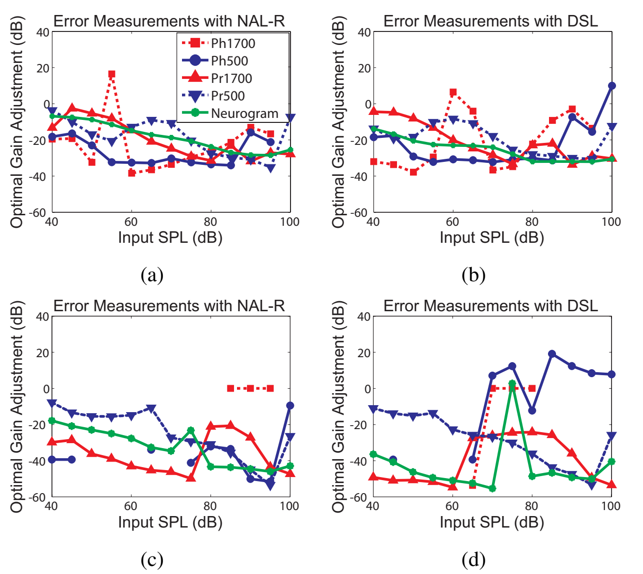


Fig. 8. Fine timing neurogram: optimal gains using a synthetic vowel /ɛ/ in the case of mild, (a) and (b), and severe, (c) and (d), hearing loss. Ph500 and Ph1700 show the optimal gain adjustment for the impaired phase response at $f = 500$ and $f = 1700$Hz (note that some data points are missing since the power ratio at these points is < 0.1 through the range of gain adjustments). Likewise, Pr500 and Pr1700 represent the optimal gain adjustments for power ratios measured at $f = 500$ and $f = 1700$Hz.

IV. CONCLUSIONS

The results with the synthetic vowel /ɛ/ strongly support the conclusions made in the gain optimization strategy for the TIMIT sentences, showing positive optimal gain adjustments for the average discharge rate, and negative optimal gain adjustments for the fine timing neurogram.

Optimal gains are determined by fundamentally different properties of the two different types of neurograms. Mean discharge rate neurograms are optimally restored in the impaired auditory periphery based on the patterns of mean discharge rate across CF. This is in contrast to optimal restoration of fine timing neurograms, which depends on the degree of synchrony fibers have to the different speech frequencies, particularly formant frequencies, and the relative phase of synchronized responses in fibers responding to those frequencies.

Discrepancies seen in optimal gain adjustments suggest that linear hearing aids, such as the NAL-R or DSL, may be trying to optimize both the fine timing and average discharge rate neurograms by presenting amplification gains in the region between optimal gains determined by the two types of neurograms. It is interesting to note, however, that the optimal gains using the average discharge rate neurogram (Fig. 7) indicate that DSL requires smaller gain adjustments than NAL-R, but the reverse is true from the optimal gains using the fine timing neurogram (Fig. 8). The difference here suggest that even though amplifications schemes may be optimizing for both fine and average discharge rate neurograms, DSL is weighted towards optimizing the average discharge neurogram whereas NAL-R is weighted towards optimizing the fine timing neurogram.

The presented work is an initial investigation into the differences between fine-timing and average discharge rate responses. Future avenues exist in exploring the effects of multi-band compression schemes on optimal gain adjustments and in understanding gain optimization for different categories of phonemes.

REFERENCES

[1] I. C. Bruce, F. Dinath, and T. J. Zeyl, "Insights into optimal phonemic compression from a computational model of the auditory periphery," in *Auditory Signal Processing in Hearing Impaired Listeners*. International Symposium on Auditory and Audiological Research, August 2007.

[2] S. F. Lybarger, "Selective amplification–a review and evaluation," *J. Am. Audiol. Soc.*, no. 6, pp. 258–266, 1978.

[3] H. Dillon, *Hearing Aids*. New York: Thieme Medical Publishers, 2001.

[4] M. S. A. Zilany and I. C. Bruce, "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," *J. Acoust. Soc. Am.*, vol. 120, no. 3, pp. 1446–1466, Sept. 2006.

[5] ——, "Representation of the vowel /ɛ/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats," *J. Acoust. Soc. Am.*, vol. 122, no. 1, pp. 402–417, July 2007.

[6] F. M. Wiener and D. A. Ross, "The pressure distribution in the auditory canal in a progressive sound field," *J. Acoust. Soc. Am.*, vol. 18, no. 2, pp. 401–408, Oct. 1946.

[7] D. Byrne and H. Dillon, "The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid," *Ear and Hearing*, vol. 7, no. 4, pp. 257–265, 1986.

[8] D. Byrne and W. Tonisson, "Selecting the gain of hearing aids for persons with sensorineural hearing impairments," *International Journal of Audiology*, no. 5, p. 5159, 1976.

[9] C. Kamm, D. D. Dirks, and M. R. Mickey, "Effect of sensorineural hearing loss on loudness discomfort level and most comfortable loudness," *Journal of Speech and Hearing Research*, no. 21, pp. 668–681, 1978.

[10] M. C. Liberman and N. Y.-S. Kiang, "Acoustic trauma in cats: Cochlear pathology and auditory-nerve activity." *Acta. Otolaryngol. Suppl.*, no. 358, pp. 1–63, 1978.

[11] R. L. Miller, J. R. Schilling, K. R. Franck, and E. D. Young, "Effects of acoustic trauma on the representation of the vowel /ɛ/ in cat auditory nerve fibers," *J. Acoust. Soc. Am.*, vol. 101, no. 6, pp. 3602–3616, 1997.