PHYSIOLOGICAL ASSESSMENT OF NONLINEAR HEARING AID AMPLIFICATION SCHEMES

Benedict Leung and Ian C. Bruce

Department of Electrical & Computer Engineering, McMaster University, Hamilton, ON, Canada

ibruce@ieee.org

AUDITORY ENGINEERING LABORATORY

Abstract

Nonlinear amplification schemes for hearing aids have been developed to deal primarily with the problem of loudness recruitment. The most commonly used form of nonlinear amplification is wide-dynamic-range compression (WDRC). Unfortunately, finding WDRC characteristics that satisfactorily deal with loudness recruitment while maintaining good speech intelligibility has proven difficult. An alternative nonlinear scheme, Advanced Dynamic Range Optimization (ADRO), has been shown in several studies to provide better speech intelligibility and listening comfort than fast-acting WDRC. ADRO uses a set of fuzzy-logic rules to make gain changes to optimize audibility, comfort, protection against loud sound, and noise attenuation. The "hearing protection" gain rule acts instantaneously, whereas the audibility and comfort rules adjust the gain slowly, such that ADRO provides linear amplification most of the time. The goal of this study was to examine the physiological basis for the relative performance of linear amplification, WDRC, and ADRO. Sentences from the TIMIT Speech Database were processed by each algorithm. In the case of WDRC, both single-channel and multi-channel schemes with fast and slow dynamics were tested. Speech signals were presented at 52, 62, 74, and 82 dB SPL (sound pressure level) with various noise levels and types, to simulate real-life environments. The simulations first use an auditory-periphery model to generate a "neurogram" of the auditory nerve's representation of the test speech material. The spectral and temporal modulations in the neurogram are then analyzed by a model of cortical speech processing. The effects of the background noise, the presentation level, the hearing loss and the amplification scheme are evaluated by comparing the cortical model response for a given condition (the "test" response) to the cortical model response to the same TIMIT sentence presented in quiet at 65 dB SPL to the normal-hearing model (the "template" response). From the difference between the test and template responses, a spectro-temporal modulation index (STMI) value is calculated. High STMI values predict good speech intelligibility, while low values predict poor intelligibility. Preliminary results show that ADRO is better at restoring the neural representation of speech than the other algorithms tested, even when the WDRC algorithms utilize slow time constants. In the case of no background noise, all the algorithms perform similarly well. However, when background noise is added, STMI values for higher SPLs drop notably for all the algorithms except for ADRO, which sustains a stable value throughout the range of SPLs test.

I. INTRODUCTION

- Computational models of speech processing in the ear and brain were used to investigate how compression algorithms affect the neural representation of speech.
- The goal of this study was to examine the physiological basis for the relative performance of linear amplification, WDRC, and ADRO [1, 2]. In the case of WDRC, both single-channel and multi-channel schemes with fast and slow dynamics were examined.
- For linear amplification, NAL-RP and DSL prescriptions were tested [3]. For WDRC, NAL-NL1 [4] and DSL *m*[i/o] [5] prescriptions were investigated.

II. METHODS

A. Models

- The auditory-periphery model used in this study (Fig. 1) was that of Zilany and Bruce [6, 7]. This phenomenological model describes the cat auditory pathway from the middle ear through to the auditory nerve.
- In this study, the real-ear unaided gain is modelled after the adult head-related transfer function described by Wiener and Ross [8].
- Input to the model consists of a sound waveform with instantaneous pressures in units of Pascal, sampled at a rate of 100 kHz, and the output is model AN fiber spike times.

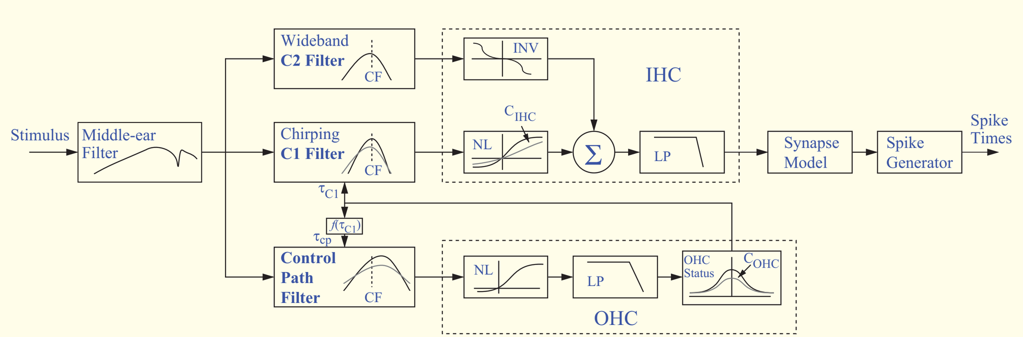

Figure 1: *Zilany and Bruce cat auditory nerve model [6, 7].*

- The model can incorporate outer hair cell (OHC) and inner hair cell (IHC) impairment to produce a range of hearing loss profiles. For this study, the model audiogram shown in Fig. 2 was utilized. For the threshold shift at each CF, 2/3 was created by OHC and 1/3 by IHC impairment.

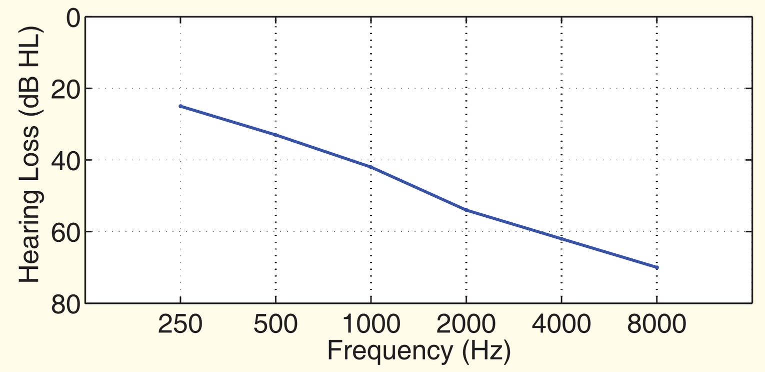

Figure 2: *Model audiogram used in this study.*

- Simultaneous outputs (discharge rates averaged over 8 ms with 50% overlap) from 128 AN fibers, CFs ranging from 0.18 to 7.04 kHz spaced logarithmically, make up the AN "neurogram", as shown in Fig. 3.
- A cortical model of speech processing [9] analyzes the AN neurogram to estimate the spectral and temporal modulation content, as shown in Fig. 3. It is implemented by a bank of modulation-selective filters ranging from slow to fast rates (2 to 32 Hz) temporally and narrow to broad (0.25 to 8 cyc/oct) scales spectrally.

B. Speech Intelligibility Predictor

- The Spectro-Temporal Modulation Index (STMI) of Elhilali et al. [9] is a measure of speech integrity as viewed by a model of the auditory system.
- The deviation between the template response (i.e. the expected response) and the test response at the cortical stage gives a measure of the STMI, as illustrated in Fig. 3.

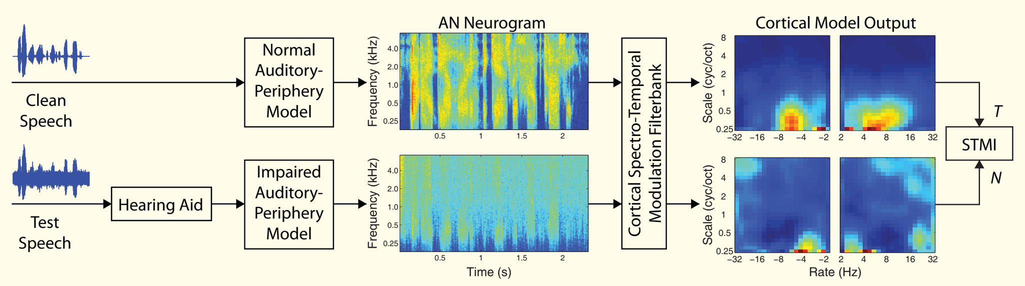

Figure 3: *Schematic of the STMI speech-intelligibility predictor computation. The clean and noisy speech signals are given as inputs to the auditory periphery model. The right panel shows the cortical output of both clean and noisy inputs for a short segment of the sentence, averaged over frequency. The cortical patterns are then used to compute the STMI.*

- Following [10], the template has been chosen as the output of the normal model (i.e. unimpaired) to the stimulus at 65 dB SPL (conversational speech level) in quiet.
- After analyzing the two-dimensional (time and frequency) AN neurogram with the modulation filter banks, the cortical output is a four-dimensional (time, frequency, rate and scale) complex-valued representation.
- Once the cortical output of the test stimulus, N, and the template, T, for that stimulus are computed, the STMI can be calculated as [10]:

$$\mathrm{STMI} = \sqrt{1 - \frac{\| T - N \|^2}{\| T \|^2}}, \qquad (1)$$

where $\|\cdot\|$ indicates the 2-norm of the corresponding signal.

C. Hearing Aid Schemes

- All of the hearing aid algorithms were tested using software simulations.
- The linear schemes and single-band compression schemes were realized using an FFT overlap-and-add filter implementation.
- Multi-band compression was implemented using a 4-band filterbank.
- A Simulink model of ADRO was supplied by Dynamic Hearing.
- Linear prescriptions were obtained using the NAL-RP formula and the DSL look-up table from [3]. Non-linear prescriptions were obtained from the NAL-NL1 v1.40 [4] and DSL *m*[i/o] v5.0a [5] software packages.
- For *fast compression*, the attack and release times were 5 and 25 ms, respectively, and for *slow compression* they were 120 and 500 ms.

D. Stimuli

- A sentence from the TIMIT Speech Database was used in this study.
- For good sound pressure level coverage at the input of the hearing aid algorithms, four different SPLs of 52, 62, 74 and 82 dB SPL were tested.

III. RESULTS

A. In Quiet

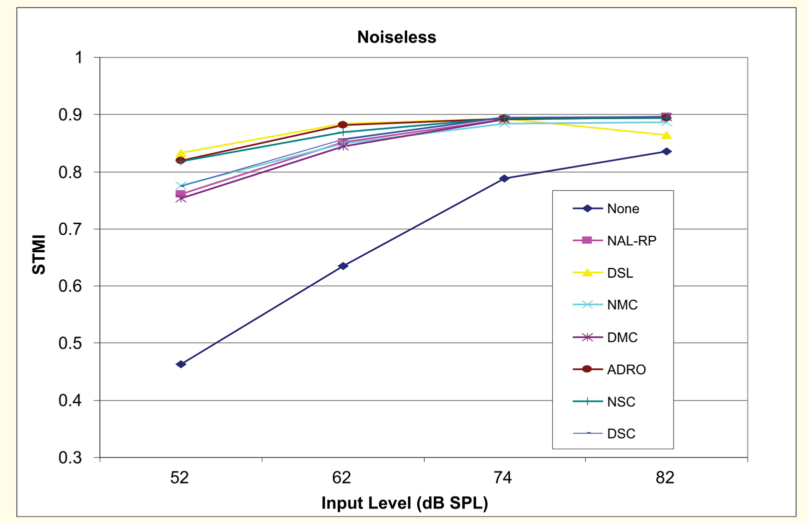

Figure 4: *Predictions of speech intelligibility in quiet for the different amplification schemes, as given in the legend. Abbreviations: NAL-NL1 Multi-Channel (NMC); NAL-NL1 Single-Channel (NSC); DSL m[i/o] Multi-Channel (DMC); DSL m[i/o] Single-Channel (DSC);*

B. In White Gaussian Noise and Babble Noise

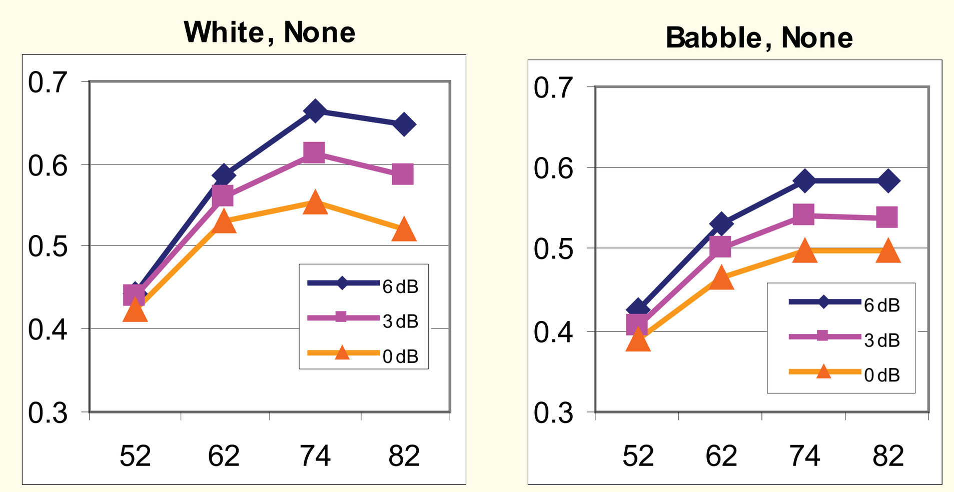

Figure 5: *Predictions of unaided speech intelligibility. The signal-to-noise ratio is given in the legend.*

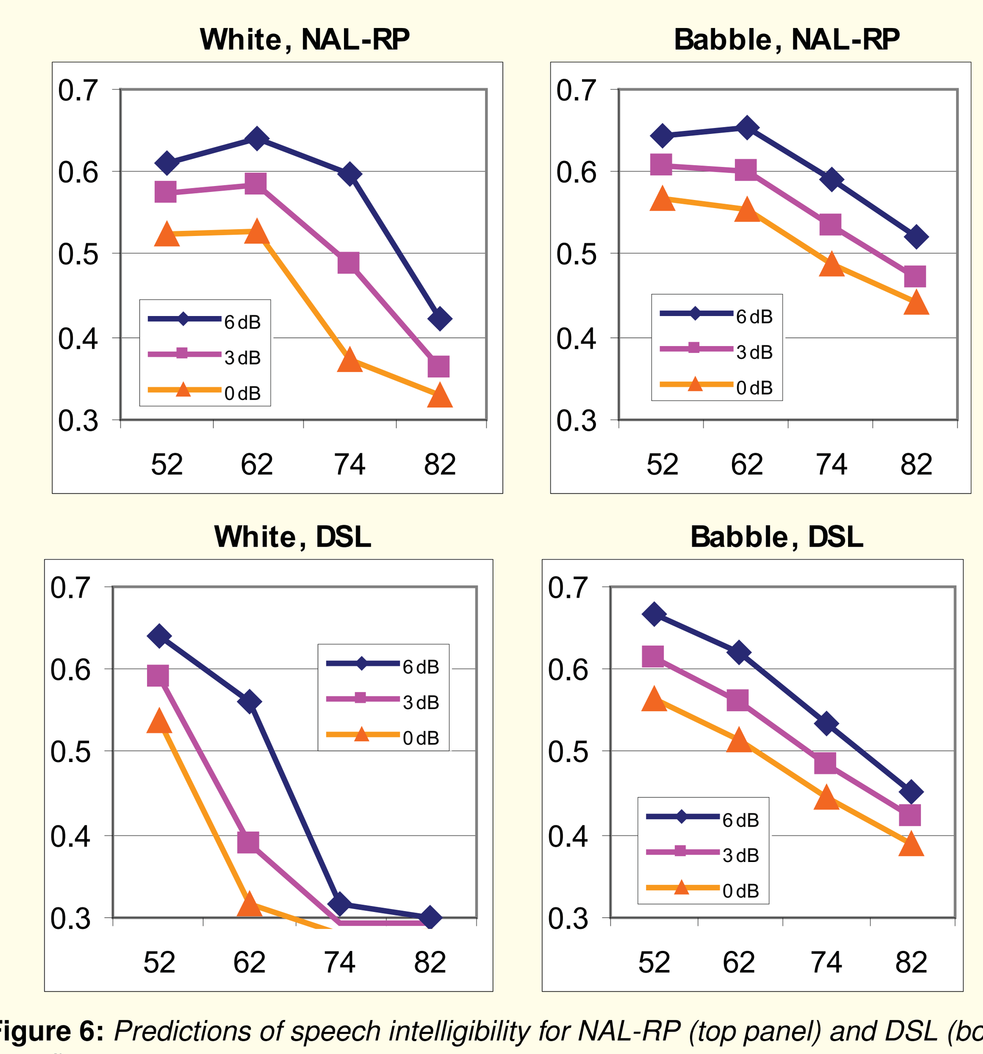

Figure 6: *Predictions of speech intelligibility for NAL-RP (top panel) and DSL (bottom panel).*

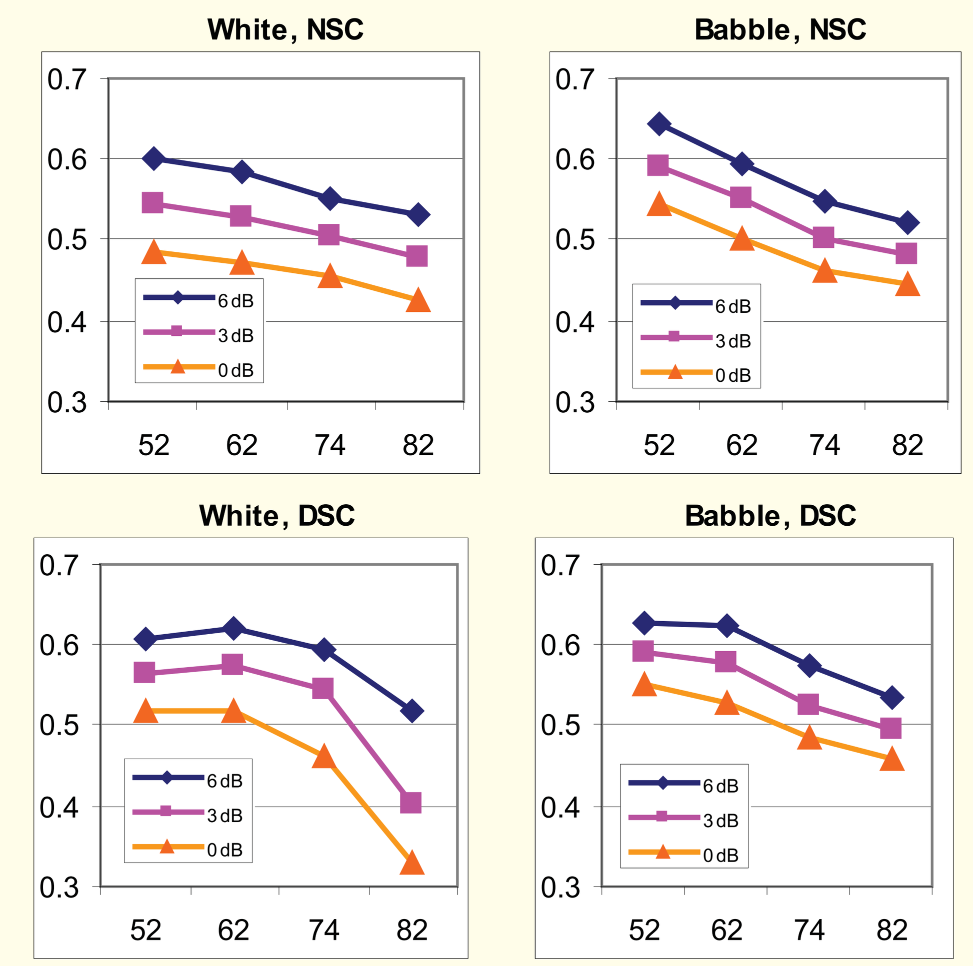

Figure 7: *Predictions of speech intelligibility for fast single-channel NAL-NL1 (top panel) and DSL m[i/o] (bottom panel).*

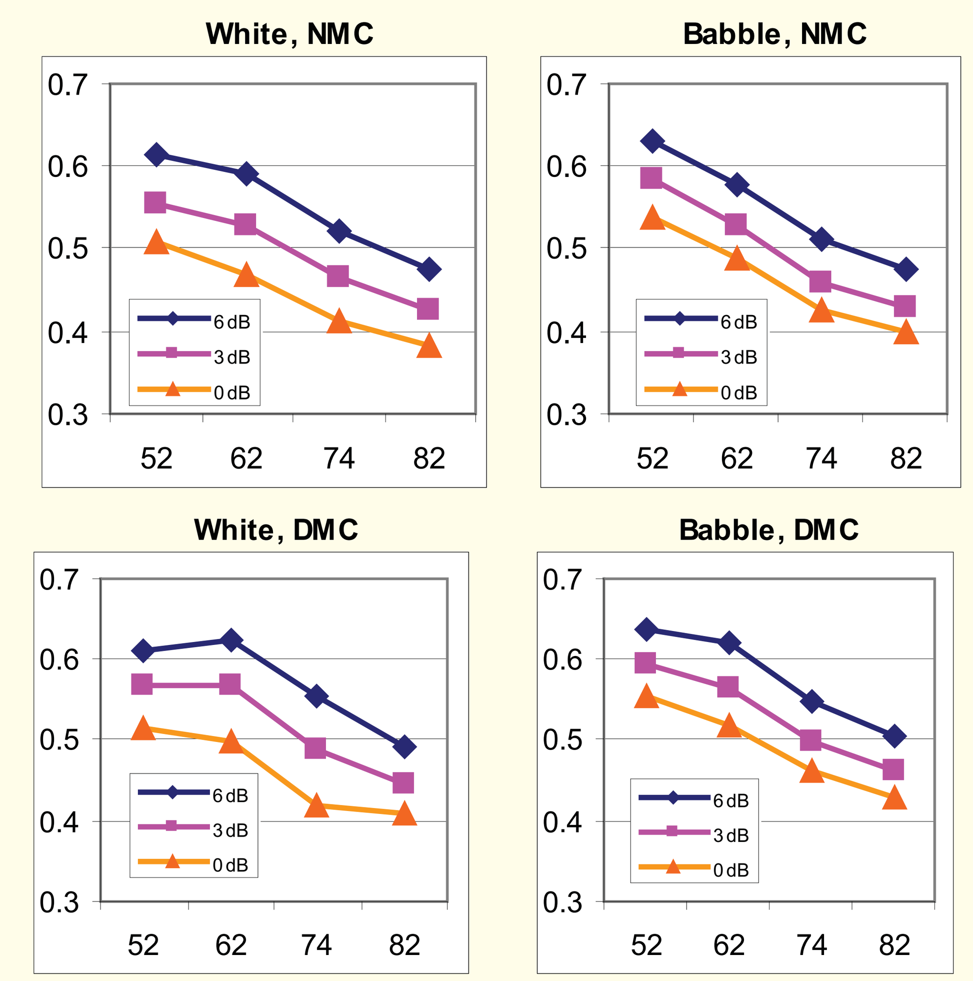

Figure 8: *Predictions of speech intelligibility for fast multi-channel NAL-NL1 (top panel) and DSL m[i/o] (bottom panel).*

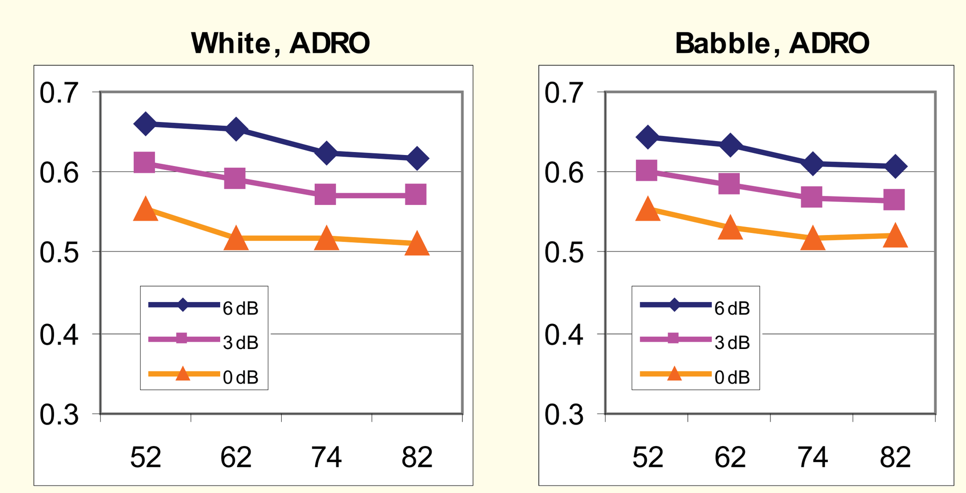

Figure 9: *Predictions of speech intelligibility for ADRO.*

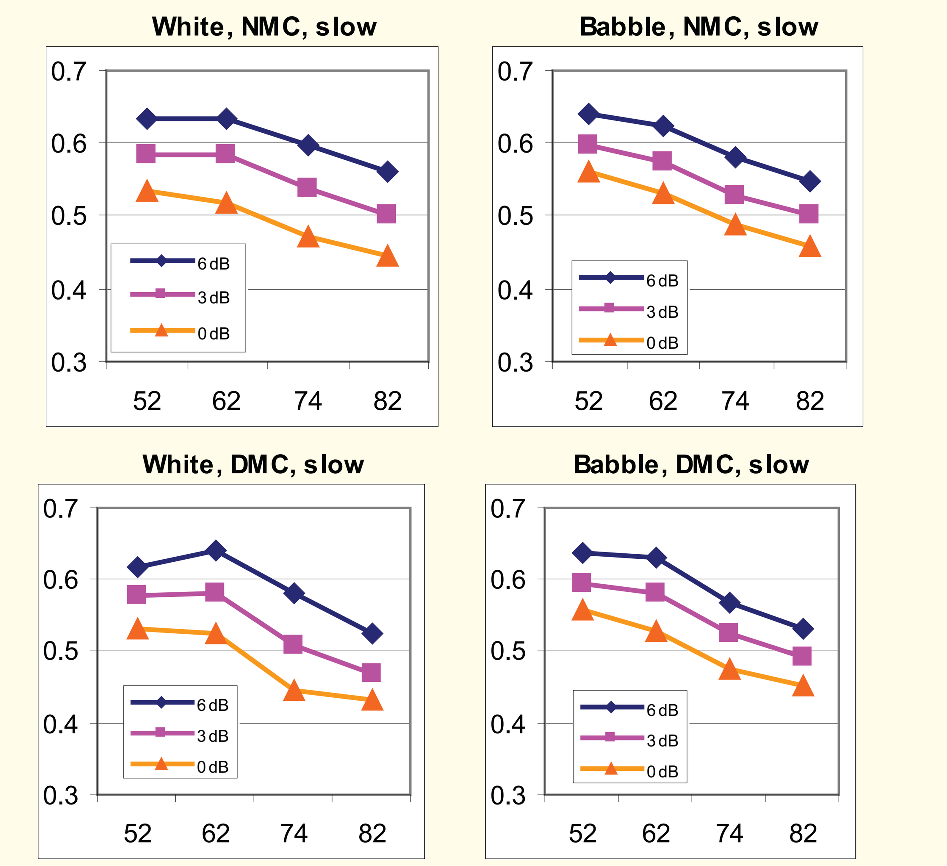

Figure 10: *Predictions of speech intelligibility for slow multi-channel NAL-NL1 (top panel) and DSL m[i/o] (bottom panel).*

IV. CONCLUSIONS

- All hearing aid algorithms gave very good speech intelligibility predictions for listening in quiet. Differences between the different algorithms become apparent in background noise. In general, linear aids work well for low SPLs and nonlinear aids work better for high SPLs. However, the predicted intelligibility at high SPLs for some amplification schemes was worse than the *unaided* intelligibility.
- Slower acting compression worked slightly better overall. However, it does not provide as much protection as faster acting compression because it does not reduce gain as fast when exposed to sudden high level sound.
- ADRO's performance was superior to that of slow multi-band WDRC at high SPLs.

V. ACKNOWLEDGMENTS

The authors thank Dynamic Hearing for providing a software simulation of the ADRO algorithm and the National Centre for Audiology at the University of Western Ontario for providing the DSL *m*[i/o] v5.0a prescription software.

References

[1] P. J. Blamey, "Adaptive dynamic range optimization (ADRO): A digital amplification strategy for cochlear implants and hearing aids," *Trends Amplif.*, vol. 9, no. 2, pp. 77–98, 2005.

[2] ——, "Adaptive dynamic range optimization for hearing aids," in *Proceedings of 9th Western Pacific Acoustics Conference (WESPAC IX)*, Seoul, Korea, June 26–28 2006.

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

[4] D. Byrne, H. Dillon, T. Ching, R. Katsch, and G. Keidser, "NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparison with other procedures," *J. Am. Acad. Audiol.*, vol. 12, pp. 37–51, 2001.

[5] S. Scollie, R. Seewald, L. Cornelisse, S. Moodie, M. Bagatto, D. Laurnagaray, S. Beaulac, and J. Pumford, "The Desired Sensation Level multistage input/output algorithm," *Trends Amplif.*, vol. 9, no. 4, pp. 159–197, 2005.

[6] 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.

[7] ——, "Representation of the vowel /ε/ 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.

[8] 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, 1946.

[9] M. Elhilali, T. Chi, and A. Shamma, "A spectro-temporal modulation index (STMI) for assessment of speech intelligibility," *Speech Comm.*, vol. 41, pp. 331–348, 2003.

[10] M. S. A. Zilany and I. C. Bruce, "Predictions of speech intelligibility with a model of the normal and impaired auditory-periphery," in *Proceedings of 3rd International IEEE EMBS Conference on Neural Engineeering.* Piscataway, NJ: IEEE, May 2007.