DOI: 10.61782/fa.2023.0205

PREDICTING SUPRA-THRESHOLD SPEECH PROCESSING DEFICITS IN HEARING-IMPAIRED LISTENERS USING A PHYSIOLOGICALLY INSPIRED AUDITORY MODEL

Johannes Zaar[1,2*] **Laurel H. Carney**[3]
[1] Eriksholm Research Centre, 3070 Snekkersten, Denmark
[2] Hearing Systems Section, Department of Health Technology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[3] Departments of Biomedical Engineering and Neuroscience, University of Rochester, Rochester, NY, 14642, USA

ABSTRACT

This study presents a new evaluation of a recently introduced speech-intelligibility prediction model, focusing on supra-threshold speech-processing deficits in hearing-impaired listeners. The model [Zaar and Carney, (2022), Hear Res., 426:108553] is based on a physiologically inspired auditory model that simulates the auditory periphery and on the recently proposed notion of across-frequency fluctuation profiles and their relevance for (speech) sound discrimination. The model has been shown to predict speech intelligibility in hearing-impaired listeners for several noise types, based solely on the stimuli and the audiogram. However, no individualized amplification was provided and the contributions of reduced audibility and supra-threshold deficits could thus not be completely disentangled. To investigate the model's predictive power with respect to purely supra-threshold aspects of speech reception, the present study evaluated the model using an additional speech-intelligibility data set collected with 9 hearing-impaired listeners using individualized amplification. In addition to the audiogram, supra-threshold measures of auditory function related to loudness perception were used to individualize the model predictions. The results indicate that while loudness-based estimates of outer-hair-cell loss did affect the model predictions, the predictive power did not improve. This result indicates that other measures of supra-threshold auditory deficits should be considered in future speech-intelligibility prediction efforts.

*__Corresponding author__: jozr@eriksholm.com.

Keywords: *Speech intelligibility, hearing loss, noise.*

1. INTRODUCTION

Speech-intelligibility prediction models face the challenge of performing effectively in a large range of acoustic conditions (e.g., various types of noise and other interfering sounds). Most models have been designed and evaluated using behavioral data from normal-hearing (NH) listeners. In recent years, a number of models have been proposed that attempt to predict speech intelligibility in listeners with hearing-impairment (HI), typically taking into account basic measures of auditory function, such as the pure-tone audiogram. A model by the authors [1, 2] was conceived based on the concept of neural fluctuation profiles in peripheral responses, which are translated into average rate profiles the midbrain [3, 4]. The model uses an auditory-nerve model (ANM) [5] followed by a bandpass modulation filter to represent processing at the level of the midbrain. The model was shown to predict the effects of signal-to-noise ratio (SNR) and different noise types in NH listeners, as well as the interaction of said effects with individual HI. The HI predictions were obtained based on individual audiograms, using assumptions about the contributions of the outer and inner hair cells (OHCs, IHCs) to the overall hearing loss (OHCL, IHCL). It was observed that a larger simulated contribution of IHCs to the hearing loss led to elevated predicted SRTs, and that different OHCL/IHCL combinations yielded optimal predictions for individual listeners. However, the stimulation paradigm did not guarantee perfect audibility at all frequencies and the OHCL/IHCL combinations were selected arbitrarily, i.e., without additional reference data beyond the audiogram. Therefore, to investigate the model with respect to "pure" supra-threshold speech processing and data-driven individualization beyond the audiogram, the present study simulated speech-in-noise data collected with linear individualized amplification and individual estimates of OHCL.

2. METHODS

2.1 Reference data

The behavioral data consisted of speech reception thresholds (SRTs) collected from 10 NH listeners (average age 24.7 years). SRTs were collected from a homogeneous group of 9 listeners (average age 71.2 years) with sensorineural hearing loss (see Fig.1) (Regev et al., personal communication).

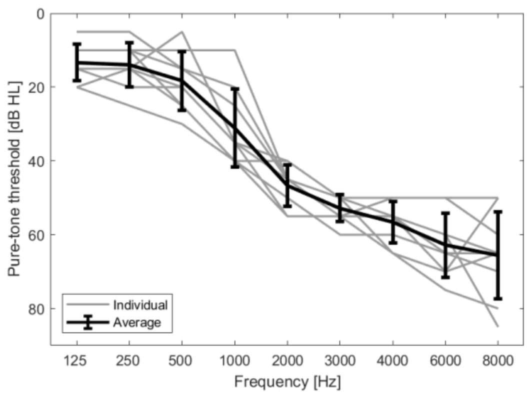

Figure 1. Pure-tone audiograms of HI listeners.

All listeners were tested monaurally via headphones using the Danish HINT sentences [6]. Four interferers were considered here: (i) *SSN* – speech-shaped stationary noise, (ii) *ICRA* - 3-band modulated SSN, (iii) *Cocktail* - a mix of two speakers (male and female) and cafeteria noise, with moderate reverberation added, (iv) *Speech* – a female Danish single speaker (average F_0: 192 Hz). SRTs were measured by adapting the level of the interferer and keeping the target speech constant at 60 dB SPL. Individualized linear amplification was provided for the HI listeners using the CAMEQ algorithm [7]. In addition to the audiograms, estimates of the proportions of OHC loss contributing to the total hearing loss were available at frequencies of 0.5, 1, 2, and 4 kHz, estimated using adaptive categorical loudness scaling in combination with a loudness model [8].

2.2 Model simulations

A subset of the stimulus speech material, with added interferers, was used for the simulations, with input SNRs ranging from -21 dB to 12 dB in 3-dB steps. The level of the noise-alone reference signal was adjusted to be identical to the level of the speech+noise mixture. The model was calibrated using a fitting function obtained by a nonlinear regression between percent-correct data measured for a range of SNRs in NH listeners in the *ICRA* condition and the corresponding model decision metrics. To predict HI subjects' SRTs, the configuration of the ANM was adapted according to the test-ear audiograms of the individual HI listeners by adjusting the ANM's OHC and IHC impairment parameters. Two simulations were conducted for the HI listeners: (i) *default OHCL/IHCL* assumed that 67% of the total hearing loss was related to OHC impairment and 33% to IHC impairment [9]; (ii) *individualized OHCL/IHCL*, assumed that the proportion of OHC and IHC impairment was adjusted based on the OHCL estimate (linearly interpolated up to and kept constant above 4 kHz). The same individualized linear amplification that was used in the experiment was used in the simulations to faithfully represent the experimental stimuli. The decision metric was converted to SI scores using the fitting function, and SRTs were calculated for all conditions as the 50-%-correct points on the predicted psychometric functions [see 1].

3. RESULTS AND DISCUSSION

The top panel of Fig. 2 shows a scatter plot of the measured (x-axis) and predicted (y-axis) SRTs for the *default OHCL/IHCL* case. The model predictions were significantly ($r=0.57$, $p<0.001$) correlated with the data when jointly considering noise conditions and listeners. The mean absolute error (MAE) was 3 dB. The bottom panel of Fig. 2 shows a similar scatter plot for the *individualized OHCL/IHCL* case. Comparison between the two panels suggests some changes in the predictions due to individualization of OHCL/IHCL; however, the predictive power of the model did not improve, but instead slightly deteriorated. This trend was driven by a single subject; when that subject was removed, correlations were higher (around 0.7) and the individualization led to a slight improvement. Nonetheless, the individualization had a limited effect, which may be due the homogenous selection of HI subjects (cf. Fig. 1).

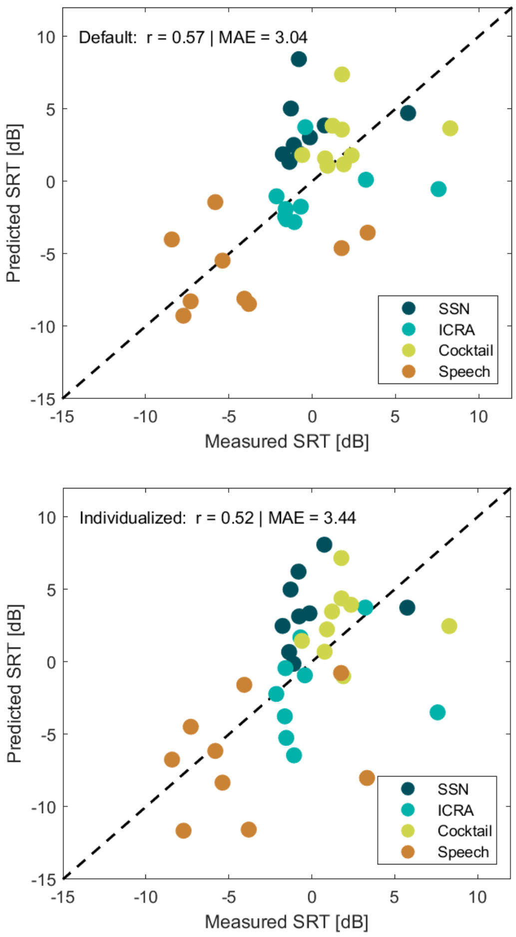

Figure 2. Predicted SRTs as a function of the HI subjects' measured SRTs for *default* (top) and *individualized* (bottom) OHCL/IHCL in the model.

Zaar and Carney [1] showed examples of OHCL/IHCL manipulations for which a larger contribution of IHCL translated to a deterioration of the predicted speech intelligibility, i.e., elevated SRTs. Figure 3 shows an assessment of this effect for the three noise types and OHCL/IHCL combinations considered in [1]. Consistent with the previous study, the predicted SRTs increased with increasing IHC contribution; the increase was steepest for the ISTS (speech-like) interferer.

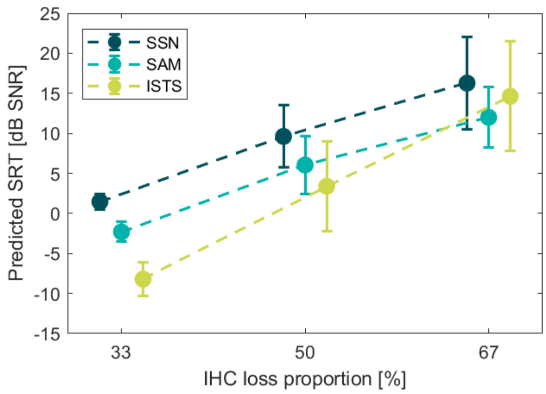

Figure 3. Across-listener average predicted SRTs from [1] for different OHCL/IHCL proportions and noise types.

Figure 4 shows the SRTs predicted in the present study as a function of the IHCL proportion for the default OHCL/IHCL case (along black dotted vertical line representing 33% IHCL) and for the individualized OHCL/IHCL configuration (with IHCL averaged across 2 and 4 kHz). The pair of SRTs for each subject are connected with a dashed line. Figure 4 demonstrates that (i) the individualization could either increase or decrease the predicted SRTs when the OHCL/IHCL ratio was manipulated with respect to the default assumption, and (ii) the trend of elevated SRT predictions for larger IHC contributions observed in [1] (Fig. 3) was also found in simulations of the present data set.

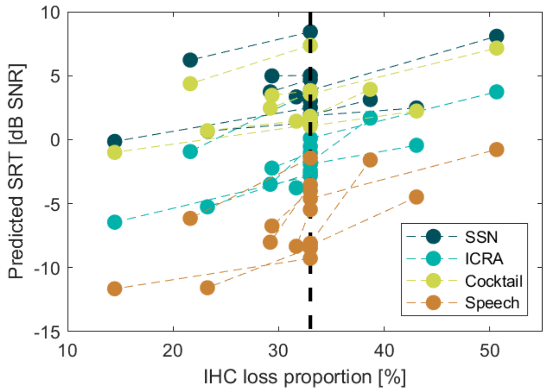

Figure 4. Predicted SRTs as a function of IHCL for all subjects and noise conditions. The dashed black line indicates the default IHCL proportion of 33%. Pairs of circles connected with dashed lines relate to a given subject.

4. CONCLUSIONS

The present study tested the model proposed in [1] using a small data set of SRTs measured in listeners with relatively homogenous HI, using individualized amplification and loudness-based estimates of OHC and IHC contributions. The model yielded satisfactory predictions of the SRT differences across listeners and conditions. Individualization of the model configuration using OHCL and IHCL estimates affected the model simulations as expected based on the previous study. However, the individualization did not markedly improve the model's predictive power. This result may be due to the homogeneity of hearing losses considered or the loudness-based individualization measures used. Future work should focus on additional data sets with a more heterogenous HI population and consider alternative measures of supra-threshold auditory deficits, such as spectro-temporal modulation detection [10, 11].

5. ACKNOWLEDGMENTS

We would like to thank Jonathan Regev for providing the stimuli and reference data used in this manuscript. Laurel Carney is supported by NIH-R01-DC-001641.

6. REFERENCES

[1] J. Zaar and L. H. Carney: "Predicting speech intelligibility in hearing-impaired listeners using a physiologically inspired auditory model," *Hearing Research*, vol. 426, 2022.

[2] C. Scheidiger, L. H. Carney, T. Dau and J. Zaar: "Predicting Speech Intelligibility Based on Across-Frequency Contrast in Simulated Auditory-Nerve Fluctuations," *Acta Acustica United with Acustica* vol. 104, pp. 914-917, 2018.

[3] L. H. Carney, T. Li, and J. M. McDonough: "Speech coding in the brain: representation of vowel formants by midbrain neurons tuned to sound fluctuations," *eNeuro*, vol. 2, no. 4, pp. 1–12, 2015.

[4] L. H. Carney: "Supra-threshold hearing and fluctuation profiles: implications for sensorineural and hidden hearing loss," *J. Assoc. Res. Otolaryngol.*, vol. 19, no. 4, pp. 1–22, 2018.

[5] M. S. A. Zilany, I. C. Bruce and L. H. Carney: "Updated parameters and expanded simulation options for a model of the auditory periphery," *J. Acoust. Soc. Am.,* vol. 135, no. 1, pp. 283–286, 2014.

[6] J. B. Nielsen and T. Dau: "The Danish hearing in noise test," *Int. J. Audiol.*, vol. 50, no. 3, pp. 202–208, 2011.

[7] B. C. J. Moore and B. R. Glasberg: "Use of a Loudness Model for Hearing-Aid Fitting. I. Linear Hearing Aids," *British Journal of Audiology*, vol. 32, no. 5, pp. 317-335, 1998.

[8] T. Brand and V. Hohmann: "An adaptive procedure for categorical loudness scaling," *The Journal of the Acoustical Society of America*, vol. 112, pp. 1597-1604, 2002.

[9] I. C. Bruce, A. C. Leger, B. C. J. Moore and C. Lorenzi: "Physiological prediction of masking release for normal-hearing and hearing-impaired listeners," *Proc. of Meetings on Acoustics* 19(1), 050178, 2013.

[10] J. Zaar, L. B. Simonsen, T. Behrens, T. Dau and S. Laugesen, "Investigating the relationship between spectro-temporal modulation detection, aided speech perception, and directional noise reduction preference in hearing-impaired listeners," *Proc Int Symp Auditory Audiol Res (ISAAR)*, Nyborg, pp. 182-188, 2019.

[11] J. Zaar, L. B. Simonsen, T. Dau and S. Laugesen, „Toward a clinically viable spectro-temporal modulation test for predicting supra-threshold speech reception in hearing-impaired listeners," *Hearing Research*, vol. 427, pp. 108650, 2023.