Bibliographical reference

Modeling Speech Intelligibility of Simulated Bimodal and Single-Sided Deaf Cochlear Implant Users, Ayham Zedan, Ben Williges and Tim Jürgens, Acta Acustica **vol. 104** (Number 5), 2018, pp. 918-921

DOI

https://doi.org/10.3813/AAA.919256

Acta Acustica united with Acustica
Vol. 104 (2018) 918 – 921

DOI 10.3813/AAA.919256

Modeling Speech Intelligibility of Simulated Bimodal and Single-Sided Deaf Cochlear Implant Users

Ayham Zedan[1], Ben Williges[1], Tim Jürgens[1,2]

[1] Medical Physics and Cluster of Excellence "Hearing4all", Carl-von-Ossietzky University, 26129 Oldenburg, Germany.

[2] University of Applied Sciences Lübeck, Germany.

[ayham.zedan, ben.williges, tim.juergens]@uni-oldenburg.de

Summary

A model of speech intelligibility (SI) for bimodal cochlear implant (CI) subjects was extended to include different amounts of aided impaired hearing contralateral to the CI. This was done by using a nonlinear hearing aid (HA) simulation, shadow filtering, and adding an electro-acoustic (EA) weighting function to address frequency overlap across ears. The extended model predicts speech-in-noise performance of simulated bimodal and single-sided deaf CI subjects with high accuracy, including their spatial release of masking for different spatial scenarios. The best fitting EA function indicates the importance of electrically versus acoustically stimulated ear to the model.

PACS no. 43.66.Ba, 43.66.Pn, 43.71.Ts, 43.71.Ky

1. Introduction

CI candidacy today includes subjects with varying degrees of acoustic hearing on the contralateral side. Therefore, many CI subjects' hearing ability consists of usable contralateral acoustic hearing that is either supported by a HA, which makes them "bimodal" CI subjects, or normally hearing (NH) if they were implanted after single-sided deafness (SSD). The access to contralateral acoustic hearing has been shown to be beneficial for SI performance compared to unilateral CI listening, e.g., [1]. However, the amount of benefit is highly variable across subjects, which indicates different abilities to integrate electric and acoustic information across ears.

To control some of the inter-subject variability, simulations of bimodal and SSD listening with NH subjects were done, e.g., [2]. Hereby, the CI was simulated with a vocoder that performs important aspects of CI signal processing. Hearing loss (HL) was simulated using two modules as proposed in [3]: The first module scales the input signal in frequency bands in a level dependent manner to simulate threshold and loudness recruitment including the hearing threshold. The second module simulates broader auditory filters using an iterative process on the narrowband temporal envelopes of the signal. These simulations were shown to replicate SI of actual CI subjects very well [2]. Thus, they are a first step to model bimodal and SSD CI subject's SI.

Williges *et al.* [4] modeled SI of electro-acoustic CI listeners by combining the vocoder simulation front-end with the binaural speech intelligibility model (BSIM) [5]. It limited bimodal CI listening to contralateral low-frequency profound unaided-impaired hearing without frequency overlap across ears.

This paper aims to generalize the model developed in [4] to predict SI of bimodal listeners with varying degrees of acoustic hearing contralateral to the CI. This ranges from NH, through aided HL, to deafness, using a CI vocoder, HL simulations and HA processing. In contrast to [4], this requires a new approach for the combination of acoustic and electric hearing in the form of an EA weighting function to address different amounts of frequency overlap across ears. A comparison to the speech reception thresholds (SRT)s of simulated bimodal and SSD CI subjects [2] assesses the goodness of the prediction.

2. Methods

2.1. Speech test and spatial scenarios

The model aims at predicting SRTs of the Oldenburg sentence test [6] in stationary speech-shaped noise. Three spatial scenarios were created by convolving speech and noise

Received 5 March 2018,
accepted 6 September 2018.

with head-related impulse responses of an anechoic room [7]: frontal (0°) speech with noise from the left (−90°), front (0°), or right (+90°) sides.

2.2. Simulation of CI and (aided) impaired acoustic hearing

CI listening was simulated on the right using a further development of the vocoder presented in [4], as shown in Figure 1a. The signal was processed according to the continuous interleaved sampling coding strategy with 3-ERB wide Gammatone filters into a sequential pulse pattern with 800 pulses per second for each of the 12 electrodes. Pulses were randomly shuffled within one stimulation cycle to imitate stochastic auditory nerve firing in response to high-rate electric stimulation. Each pulse was then convolved across channels with a double-sided exponentially decaying function to simulate current spread in the perilymph. The pulse pattern of each of the 12 channels was then sent through a 1-ERB wide Gammatone filterbank with center frequencies corresponding to the implanted electrode locations [8] of a 31mm MED-EL electrode array. The filterbank group delay was compensated and the bands' signals were summed to form an audible (mono) signal.

Four modes of contralateral hearing were simulated on the left side: NH, mild (MHL) and severe hearing loss (SHL) with the audiograms shown in Figure 1 (b), and deafness (no signal). MHL and SHL were aided using a 9-channel dynamic range compressor on the Open Master Hearing Aid platform [9] using the CAMFIT formula [10]. Since the HA formula is non-linear, the mixture of noise and speech is used to calculate the time-, frequency-, and input level- dependent gains, and the same gain is then applied to the noise and speech signals separately ("shadow filtering"). As a reference, binaural NH and monaural NH, SHL and MHL hearing modes were simulated.

2.3. BSIM

BSIM was used as described in [5] with the modification of using the CI/HA-vocoded stimuli as a front end. The model can predict SI in spatial scenarios for acoustic listeners due to using binaural processing and better-ear-listening.

The sketch of BSIM including the CI/HA/NH front-end for modeling bimodal and SSD-CI listeners is shown in Figure 2. BSIM needs speech (S_R, S_L) and noise (N_R, N_L) on left and right ear as separate inputs. These signals are processed by the CI vocoder and HA before being used in the BSIM. The four signals are fed into a 30-band Gammatone filterbank with center frequencies between 167 Hz and 8822 Hz. Next, NH or HL thresholds are simulated by adding uncorrelated Gaussian noise to the spectrum of N_R and N_L [5].

Then, equalization cancellation (EC) is used to produce a binaural/bimodal signal that "effectively" models binaural processing. The EC process enhances the signal to noise ratio (SNR) by minimizing the level of noise in the

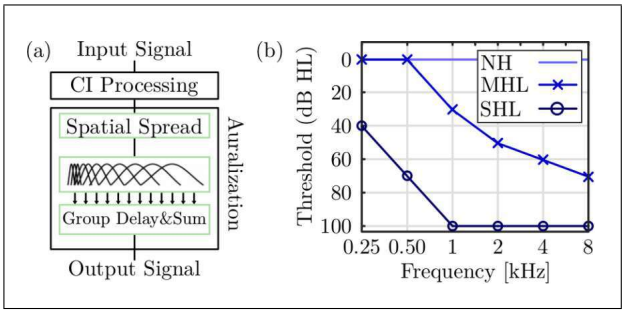

Figure 1. Left: Sketch of the CI-Vocoder, right: Simulated audiometric thresholds.

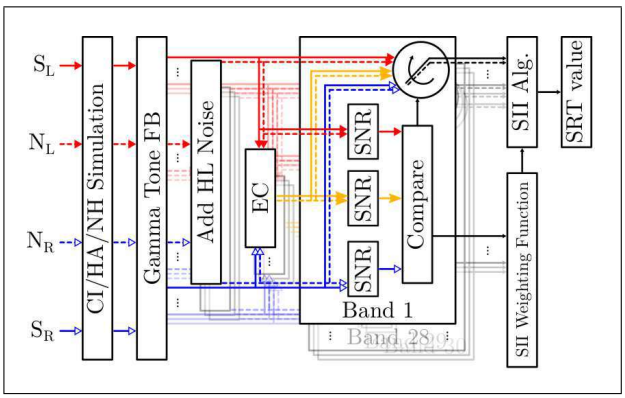

Figure 2. Model Diagram.

each frequency band independently based on optimal subtraction of the noises.

Afterwards, the model compares the SNRs of monaural left, right, and EC-processed signal in each frequency band, choosing the highest SNR of these three. Since the hearing threshold was implemented with frequency-shaped noise before, this model chooses the best *audible* SNR in each frequency band. The so-chosen speech and noise level as a function of frequency band are passed to the speech intelligibility index (SII). The SII will use an reference value (0.26 for acoustic hearing) that corresponds to a certain amount of speech intelligibility and is dependent on type of presentation and speech material.

2.4. The EA weighting function

The SII estimates SI using the power spectrum of noise and speech. This means that the model is not able to account for the information loss caused by the CI processing strategy that distorts the fine time structure of speech or the spectral smearing simulating the current spread in the CI. In order to account for this information loss, a second SII reference (0.41) for CI was chosen such that predicted and measured SRT of monaural CI matched for the spatial scenario with noise coming from the front, i.e., co-located.

It is currently unclear how CI and acoustic hearing are combined across ears. This model investigates how many acoustic or electric frequency bands are needed to best fit the data, by predicting SRTs across a range of EA weighting functions. Therefore, a SII reference value in-between

acoustic and CI reference values based on a skewed sigmoid function, as shown in Figure 3, is used. Since the relative importance of electric and acoustic hearing is unknown, several EA mid-point values (MV) were tested.

Figure 3 shows different possible EA weighting functions. Using for instance the green dashed line (MV of 21 bands), the overall SII reference value is 0.26 if 12 bands were having the highest SNR in the electric modality, 0.34 for 21, and 0.41 for 29 bands. This means that a poorer SRT (due to higher SII reference value) is only assigned if many electric bands are chosen. Thus, rightmost EA weighting functions assign more importance to acoustic hearing and vice versa. Slopes of curves were chosen identical for simplicity.

2.5. Measured SRTs

Modeled SRTs were compared to measured SRTs from [2] that were collected with eleven NH subjects tested with the same speech test, same spatial scenarios, and same simulations of CI with and without contralateral acoustic hearing.

3. Results

Figure 4 shows predicted SRTs (dots), obtained using the 21-band MV point EA weighting function, laid over measured SRTs (box plots) as a function of noise direction for frontal speech. The different hearing modes in the eight subpanels are indicated by in-set head symbols. Modeled and measured SRTs in the binaural NH mode show symmetrical decrease (improvement) of SRT when noise comes from the sides (−90°, +90°) and diagonal (asymmetric) SRT-patterns for monaural MHL and NH (top panels left to right). Measured SRTs of SHL show severe degredation in speech intelligibility along with extremely high variability between subjects. Predicted SRTs for SHL monaural (5 to 10 dB SNR, gray panel) were omitted because the SII was unable to predict such a severe hearing loss and would provide extremely elevated SRT predictions. The bottom panels show SRTs for monaural (right) CI and contralateral deafness, bimodal SHL, bimodal MHL and SSD (left to right). In both, modeled and measured SRTs, the diagonal pattern with monaural CI (bottom left) is slightly improved for contralateral SHL, turns over for MHL and shows lowest SRTs for contralateral NH. Model predictions are all within the inter-individual standard deviation of the measured SRTs. Mostly better-ear-listening can be observed, i.e., binaural SRTs are in line with the best monaural SRT. Combined benefit is between 1 and 2 dB.

Figure 5 shows predicted and measured spatial release from masking (SRM). SRM is extracted from the SRTs in Figure 4 as the most positive SRT-*improvement* when separating speech and noise. The model captures the SRM across hearing modes well, except for an underestimation at monaural CI.

Root-mean-square errors (RMSE)s and Pearson's correlation coefficients for the model's SRT predictions against the median of measured SRTs for the six EA weighting functions with different MVs, are shown in Table I. In general, correlation coefficients across all EA weighting functions are high (>0.9) and RMS-errors are less than 1.69 dB with an optimum of 1.23 dB RMS-error at a MV of 21 frequency bands.

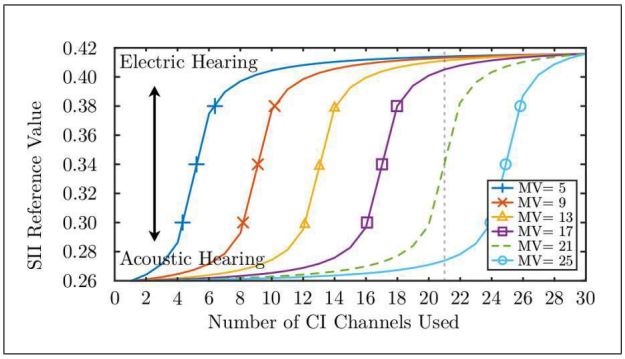

Figure 3. Electric-acoustic weighting function as a function of midpoint value (MV).

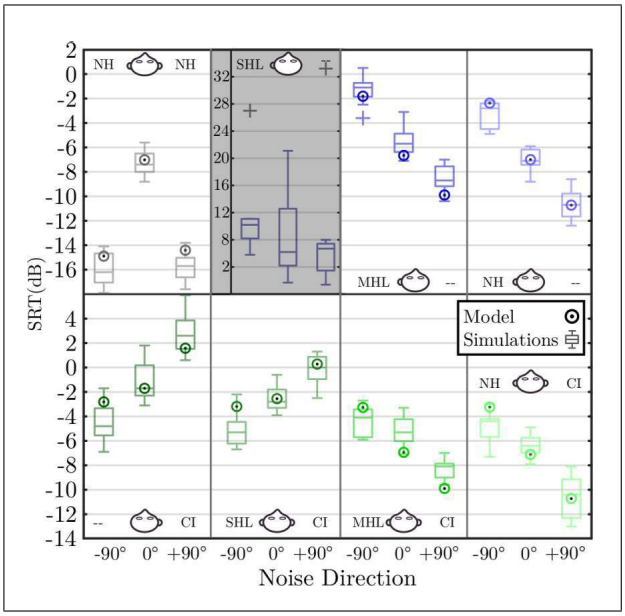

Figure 4. Predicted and measured SRTs for different modes of hearing (panels, denoted by in-set head symbols) and noise directions.

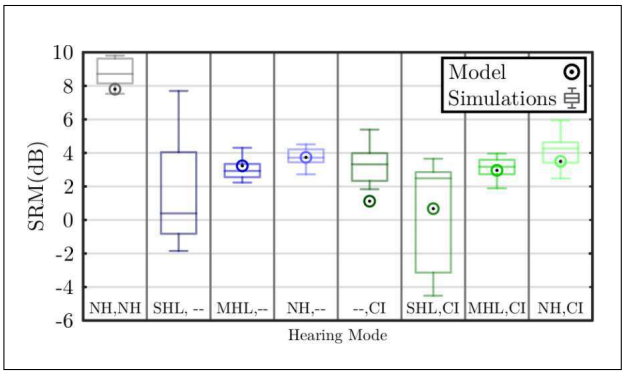

Figure 5. SRMs for different hearing modes, extracted from the measured and predicted SRTs of Figure 4.

4. Discussion

The most important additions to the model in [4] are that (1) SRTs for listeners with a systematic variation of frequency overlap across ears were predicted, and (2) a novel EA weighting function was used to integrate the electrical and acoustic information across ears. The EA weighting function used here automatically sets the SII reference value based on the number of electric or acoustic channels chosen in each spatial scenario, e.g. the SII value will differ for N0, N-90 and N90, because the number of useful CI/acoustic channels varies across these in the spatial scenarios. Therefore, it provides optimal information combination across ears.

Evaluating different EA functions shows that the model fits the SRT results best when the EA weighting function prefers acoustic hearing, more specifically, at an MV of 21 channels. This matches with studies showing benefit for CI subjects even with poor contralateral acoustic hearing [11].

The model was able to predict measured SRTs for simulated listeners in the monaural, SSD and bimodal hearing modes well across spatial scenarios. Both, better-ear listening and the limited amount of combined benefit over the best single-aided hearing mode were predicted with a relatively simple approach of using the best audible SNR in each frequency channel. However, the model tends to underestimate the combined benefit. The EC-process, and therefore binaural processing, although in principle available, was not utilized by the model when a CI was simulated due to the missing fine time structure in the signal. Instead, the model uses a better ear listening strategy to predict the SRTs of bimodal SRTs. The small combined benefit in the bimodal cases originates from the fact that the combined bimodal SII reference value is always slightly lower than that of the monaural CI. This is because of the addition of acoustic hearing in low frequency channels, where no electrical stimulation is assumed due to frequency-to-place mismatch of the CI.

The current CI simulation imitates only some, but not all characteristics of the signal processing and the electrode-nerve interface in actual CI subjects. One aspect that may be of particular importance for binaural processing is the representation of temporal fine-structure, which needs an appropriate coding strategy and a more complex representation of auditory nerve firing patterns and is thus not modeled here.

5. Conclusions

This study extended the model in [4] to include contralateral normal and (nonlinearly) aided moderate to severe hearing loss. The model was able to predict SRTs with high accuracy, replicating SRT and SRM patterns with the ability to adapt to different contributions of acoustic and electric hearing depending on the spatial scene. The EA weighting function proposed in the model automatically adapts the SII reference value based on the number of frequency bands used by electric and acoustic modality. A best match of predicted SRTs was found with an EA weighing function that dedicates acoustic hearing to be more important than electric hearing.

Table I. RMSE values of SRT predictions in dB for different electro-acoustic weighting functions.

MV	5	9	13	17	21	25
RMSE	1.69	1.45	1.29	1.24	1.23	1.62
Corr.	0.92	0.93	0.94	0.94	0.94	0.90
p <	1e-4	1e-5	1e-5	1e-5	1e-5	1e-4

Acknowledgements

This study was supported by DFG JU2858/2-1 and by the EU (EFRE) project VIBHear.

References

[1] T. Y. C. Ching, E. van Wanrooy, H. Dillon: Binaural-Bimodal fitting or bilateral implantation for managing severe to profound deafness: A review. Trends in Amplification **11**(3) (2007) 161–192.

[2] T. Jürgens, S. Kliesch, T. Wesarg, L. Geven, A. Radeloff, B. Williges: Räumliche Demaskierung von Sprache in simulierten Cochlea-Implantatträgern mit kontralateraler Normal- und Schwerhörigkeit. Annual meeting of the German Soc. for Audiology (2017) 1–4.

[3] T. Siebe, B. Williges, D. Oetting, V. Hohmann, T. Jürgens (2017). Evaluation einer modularen Auralisation von sensorineuraler Schwerhörigkeit. Annual meeting of the German Soc. for Audiology (2017) 1–4.

[4] B. Williges, M. Dietz, V. Hohmann, T. Jürgens: Spatial release from masking in simulated cochlear implant users with and without access to low-frequency acoustic hearing. Trends in Hearing. **19** (2015) 1–18.

[5] R. Beutelmann, T. Brand, and B. Kollmeier: Revision, extension, and evaluation of a binaural speech intelligibility model. J. Acoust. Soc. Am. **127**(4) (2010) 2479–2497.

[6] K. Wagener, V. Kühnel, B. Kollmeier: Entwicklung und Evaluation eines Satztests für die deutsche Sprache: Design, Optimierung und Evaluation des Oldenburger Satztests. Audiol. Acoust. **38** (1999) 4–95.

[7] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, B. Kollmeier: Database of Multichannel In-Ear and Behind-the-Ear Head-Related and Binaural Room Impulse Responses. EURASIP Journal on Advances in Signal Processing (1) (2009) 298605.

[8] D. M. Landsberger, M. Svrakic, J. T. Roland Jr, M. Svirsky. The relationship between insertion angles, default frequency allocations, and spiral ganglion place pitch in cochlear implants. Ear and Hearing, 36(5) (2015) e207–e213.

[9] The Open community platform for Hearing Aid Algorithm Research Project, OpenMHA – Open master hearing aid, http://www.openmha.org/, (2017).

[10] B. C. J. Moore, J. I. Alcantara, M. A. Stone, B. R. Glasberg: Use of a loudness model for hearing aid fitting: II. Hearing aids with multi-channel compression. Br. J. Audiol. **33**(3) (1999) 157–70.

[11] A. C. Neuman, S. B. Waltzman, W. H. Shapiro, J. D. Neukam, A. M. Zeman, M. A. Svirsky: Self-reported usage, functional benefit, and audiologic characteristics of cochlear implant patients who use a contralateral hearing aid. Trends in Hearing. **21** (2013) 1–14.