Model-based hearing aid gain prescription rule

STEPHAN D. EWERT AND GISO GRIMM

Medizinische Physik, Carl-von-Ossietzky Universität Oldenburg, Germany

In listeners with sensorineural hearing loss, loudness recruitment is typically observed and can be attributed to a loss or dysfunction of the outer hair cells (OHC). On the level of the basilar membrane (BM), OHC loss results in a reduced gain for low-level signals, changing the BM input-output function. This amount of low-level gain loss cannot be directly estimated from the overall hearing loss as characterized by pure-tone audiometric thresholds. However, from a modelling perspective, a hearing aid might be successful if it is able to compensate the gain loss and thus restore the compressive BM input-output function of a normal-hearing listener. Here, psychoacoustic temporal masking curves (TMC) and adaptive categorical loudness scaling data (ACALOS) of the same normal-hearing (NH) and hearing-impaired (HI) subjects were used to estimate gain loss. A linear model was fitted to predict gain loss from audiometric thresholds and from the steepness of the loudness function. Comparison of the predicted gain loss in HI and the gain in NH lead to a gain prescription rule. This prescription was tested with a conventional hearing-aid compressor and a model-based version, which compares simple NH and HI auditory models in real time.

INTRODUCTION

A reduced perceivable dynamic range between hearing threshold and uncomfortable level is often observed in listeners with hearing loss of cochlear origin (sensorineural hearing loss). This loudness recruitment phenomenon is thought to be related to a loss or dysfunction of outer hair cells (OHC), which facilitate basilar membrane motion at small signal levels in the normal ear. The gain attributable to the OHC persist up to a certain signal level around 30 to 40 dB SPL and is usually assumed to have vanished for levels around 85 dB SPL. Such behaviour is characterized by the basilar membrane input-output (BM I/O) function (e.g., Plack *et al*., 2004) for a certain signal frequency and basilar membrane site. A dysfunction of OHCs will reduce the gain and will thus require a higher external signal level to elicit the same amount of basilar-membrane motion as in the normal ear. Loudness perception is assumed to be based on the basilar membrane response at all frequency sites which consequently forms the initial stage of loudness models (e.g, Chalupper and Fastl, 2002). To counteract loudness recruitment, dynamic compression is widely applied in hearing aids. Depending on the exact compressor scheme and the time-constants involved, such algorithms do not aim at and cannot directly restore the BM-I/O function of the normal-hearing (NH) ear in a hearing-impaired (HI) listener. The fitting rationales for compression algorithms usually attempt to equalize loudness

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.

perception, to restore audibility, or to optimize higher-level measures such as speech perception (cf. Dillon, 2001).

In this study, an alternative approach towards a gain prescription rule was pursued. An individual model BM-I/O function was adjusted on the bases of clinically applicable audiometric methods. The aim was to directly restore the average NH BM-I/O function in the hearing aid compressor. The derived gains were then tested in a "conventional" compressor scheme with short time constants and in a model-based compressor that compared the output of the NH and HI BM-I/O function in real time.

GAIN PRESCRIPTION RULE

The underlying assumption was that the total hearing loss HL as measured by the pure-tone audiogram consists of two parts, the hearing loss attributable to loss or dysfunction of outer hair cells (HL_{OHC}) and the hearing loss attributable to inner hair cell damage (HL_{IHC}). For NH listeners, a model BM-I/O function was assumed with the following characteristics: (a) frequency-dependent linear gain (G_{NH}) for low levels according to Table 1, (b) power-law compression with an exponent of 0.25 (1:4 dB/dB compression ratio, Lopez-Poveda *et al*., 2003) for medium levels, (c) a fixed breakpoint at 85 dB SPL to return to linear behaviour (Plack *et al*., 2004; Plack and Oxenham, 2000). These assumptions result in a "broken-stick"-I/O function with frequency-dependent lower knee point depending on the maximum gain given in Table 1. The second assumption was that the maximum gain is reduced in a HI listener by HL_{OHC} which would directly manifest as gain loss (GL). The minimum resulting gain, corresponding to a total loss of OHCs, was assumed to be zero.

A recent study by Jürgens *et al*. (2011) determined GL from temporal masking curves (TMC, Nelson *et al*., 2001) in a mixed group of 5 NH and 11 HI listeners. Their data indicated that GL was about 10-15 dB lower than HL. A similar observation is also obvious for the relation of HL_{OHC} and HL in Jepsen and Dau (2011). In addition, Jürgens *et al.* found a high correlation of HL_{OHC} estimated by adaptive categorical loudness scaling (ACALOS, Brand and Hohmann, 2002) and GL estimated by TMC. Motivated by these findings, the following linear model with three free scalar parameters, *a*, *b*, *c*, was fitted to the data of Jürgens *et al.* 2011:

$$GL = a\ HL + b\ m_{low} + c\,. \qquad \text{(Eq. 1)}$$

GL was the TMC-based gain loss estimate and HL was the average of the audiometric threshold and the adaptive threshold measurement; m_{low} denotes the lower slope of the loudness function extracted from the ACALOS data (for details see Jürgens *et al*., 2011). The best fitting parameters were:

$$a = 0.58,\ b = 24.37,\ c = -8.54\ (r^2 = 0.92, \text{root-mean-square error} = 4.2\text{ dB})$$

With Eq. 1 at hand, one can now predict GL from the audiometric threshold measurement and the lower slope of the loudness function measured with ACALOS. As gain prescription rule, the NH BM-I/O function was compared to the BM-I/O function of the individual HI listener with residual gain $G_{HI} = G_{NH} - min(GL, G_{NH})$. The difference is the prescribed gain for any given frequency and signal level.

Frequency (Hz)	62.5	125	250	500	1000	2000	4000	8000
Gain G_{NH} (dB)	8*	10*	14	20	35	40	40**	30**

Table 1: Maximum gain assumed for the BM-I/O function of NH listeners based on Lopez-Poveda and Meddis (2001) and Plack and Oxenham (2000). The asterisk denotes extrapolation, two asterisks indicate lowered values to limit high-frequency gains resulting from the gain prescription rule.

HEARING AID ALGORITHMS

Two multi-band dynamic compressor schemes were tested. The reference algorithm (ref) used a linear-phase filterbank with 9 bands (1.33 bands per octave). The level estimator filtered the intensity with a first-order, 5-ms low-pass filter. The filtered intensity was converted into logarithmic levels and smoothed by a maximum tracker with a 50-ms release time constant. Level dependent gains were looked up in a gain table (resulting from the gain prescription rule described above), and applied to the respective filter band before re-synthesis. The filterbank was operating in the spectral domain (Grimm *et al.*, 2006). An additional gain limiter was set to a maximum of 40 dB insertion gain, and the maximum power output (MPO) was configured to not exceed 80 dB free-field sound pressure level in each band. In subjects where – despite of feedback control algorithms – feedback howling occurred, the maximum gain was further reduced.

The alternative, model-based algorithm (ohc) used a gammatone filterbank with complex-valued (quasi-analytic) output for frequency analysis and synthesis. The filterbank had 30 bands, one band per equivalent rectangular bandwidth (ERB) except for the four lowest bands where broader filters were used. For estimation of the level in frequency sub-bands, the absolute value of each filter output (Hilbert envelope) was filtered with a first-order, 5-ms low-pass filter, converted to a dB-scale, and smoothed by a filter with a 5-ms attack and 25-ms release time constant. For the re-synthesis, a fixed delay and phase alignment was applied before summation of the filter outputs. The gains were derived by a real-time comparison of the simple auditory BM-I/O model for NH and HI listeners as described above, including an additional stage for two-tone suppression (Hohmann and Kollmeier, 2007), see Fig. 1. To simulate the suppression of off-frequency components, the instantaneous frequency (IF) was measured in each filter and its deviation from the centre frequency in ERB was linearly approximated, $\Delta f = (IF - f_c)/f_{bw}$ (f_c = centre frequency, f_{bw} = bandwidth). IF is defined as the temporal derivative of the signal phase divided by 2π. It was low-pass filtered on the complex plane with a time constant of 5 ms for numerical stability. No suppression was applied for frequency deviations between 0 and 0.6 ERB (on-frequency components). Full suppression was applied for deviations below -1.5 ERB and above 2 ERB (off-frequency components). In the model, suppression was affecting the maximum gain. It was assumed that suppression decreases with increasing gain loss. For estimation of hearing aid gain in each frequency band, the HI model was inverted, and the required input level was looked up which created the same output as the NH model. The difference between the given input level and the required input level was the hearing aid gain. The same gain limiter and MPO as in the reference compressor were used.

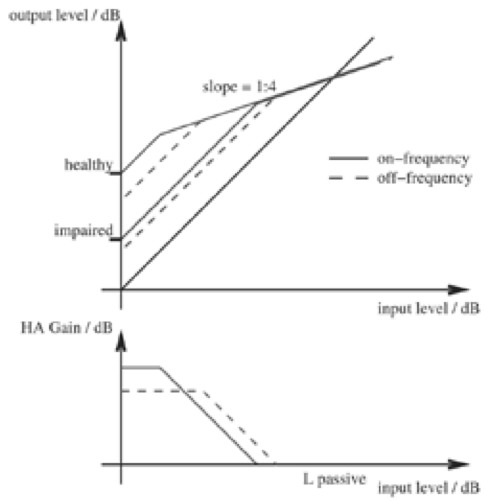

Fig. 1: Schematic I/O functions of the auditory model used by the ohc compressor (upper panel) and the derived gain (lower panel), for on-frequency (solid) and off-frequency (dashed) components.

METHODS

Speech reception thresholds

The algorithms were tested subjectively using speech reception thresholds (SRT) of sentences (OlSa, Wagener and Brand, 2005) and single phonemes in vowel-consonant-vowel (VCV) and consonant-vowel-consonant (CVC) logatomes taken from the Oldenburg logatome speech corpus (OLLO, Wesker *et al.*, 2005). The test consisted of a list with six VCV-logatomes and a list with four CVC-logatomes, see Table 2 for details. The logatomes of one list were presented interleaved in random order. The level of each logatome was controlled using a 1-interval, 2-alternative forced-choice, 1-up-1-down procedure. The step size in the measurement phase was 1 dB. SRTs of phonemes were measured in stationary speech-shaped noise, the

sentence tests were measured at stationary and modulated speech-shaped noise. Additionally, subjects were asked to rate the overall quality of the algorithm under different test conditions.

VCV	*[ɔtɔ]*	*[ɔdɔ]*	*[ɔkɔ]*	*[ɔgɔ]*	*[ɔsɔ]*	*[ɔtsɔ]*
German transcription	otto	oddo	ocko	oggo	osso	ozzo
CVC	*[tet]*	*[tit]*	*[tɛt]*	*[tɪt]*		
German transcription	teht	tiet	tett	titt		

Table 2: Phonetical description of logatomes in test lists and their german transcriptions.

Experiments and Subjects

The quality ratings were performed at two laboratories. In the first laboratory (Universitätsklinikum Gießen), 12 HI listeners participated, two female and ten male (average age 71, standard deviation 5.3 years). In the second laboratory (Jean-Uhrmacher-Institut, Köln), 10 subjects participated (four female, six male; average age 75, standard deviation 9.4 years). The quality rating data were pooled from both laboratories. At the phoneme SRT tests (performed at the Universität Oldenburg), 11 subjects participated (average age 66, standard deviation 15.2 years).

EVALUATION RESULTS

Comparison of Insertion gains with those of other gain prescription rules

The nominal real ear insertion gain (REIG) of the proposed gain prescription rule was compared with those of a commercially widely accepted gain prescription rule (NAL-NL2), and a gain prescription rule which compensates the loudness perception (characterized by ACALOS) of narrow-band stimuli (loudfit, Herzke and Hohmann 2005). An exemplary hearing loss taken from a pool of 15 typical symmetric sloping hearing losses was used for the comparison. The audiogram and ACALOS data (indicated by triangular shapes) are shown in the left panel of Fig. 2. The left ear (crosses and right part of the ACALOS triangles) was used. The tip of the triangles denotes the level corresponding to “inaudible” (L0), the broad end shows the level corresponding to “too loud” (L50). The line in the middle of the triangle represents the crossing point of a two-line linear fit (Brand and Hohmann, 2002), L_{cut}. The NAL-NL2 rule was configured for an experienced listener at age 70. The nominal target gains of a long term average speech signal (LTASS) at 50, 65, and 80 dB SPL were recorded and are shown in the right panel of Fig. 2. The model-based gains are generally higher than those resulting from the other two methods, except for high frequencies where loudfit results in even higher gains. In addition, REIG was analysed using the ISMADHA method for real signals and the proposed gain prescription for the model-based compressor (ohc) and the conventional reference compressor (ref). At low frequencies, ohc provided less gain then ref, which was caused by the suppression of off-frequency components.

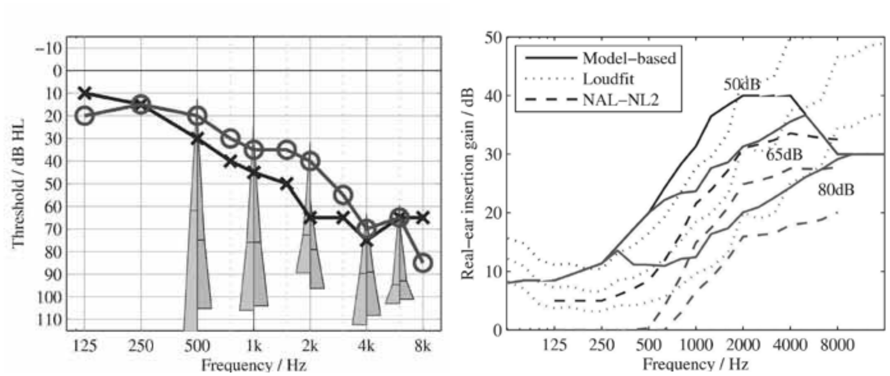

Fig. 2: Left: Audiogram with ACALOS data of the hearing loss used for simulation of real-ear insertion gains (REIG). Right: REIG simulation for a stationary, speech-shaped noise signal at 50, 65 and 80 dB SPL (see text).

Quality ratings and speech reception thresholds

The overall quality of the two compressor and and an commercial hearing aid was rated by the subjects in four test conditions. The test stimuli were a male speaker, a child, classical music and nature sounds (birds and a brook). The median rating and interquartile ranges are shown in the left panel of Fig. 3. The commercial hearing aid was rated as 'good'. The algorithms prescribed with the proposed model-based gain prescription rule were rated worse in all conditions except for nature sounds.

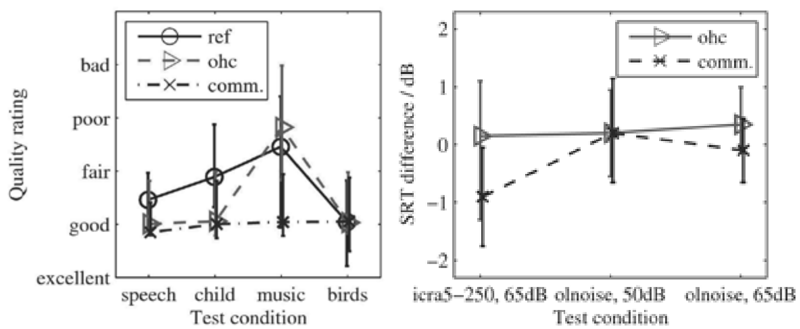

Fig. 3: Left: Quality rating. Right: Difference of individual speech reception thresholds (sentence test OlSa) between ref and the indicated algorithm. Negative values indicate better speech intelligibility.

Individual differences between speech reception thresholds measured with the two algorithms (ohc and commercial hearing aid) and the reference algorithm are shown in the right panel of Fig. 3. Lower values correspond to a better performance. The ohc algorithm performed similar to the ref algorithm, while the commercial system performed slightly better in the condition with modulated noise. Individual

differences of speech reception thresholds of phonemes between the ohc and the ref algorithm are shown in Fig. 4. No significant differences between the two algorithms were found except for the phoneme “oggo”, where the ohc algorithm performed significantly better.

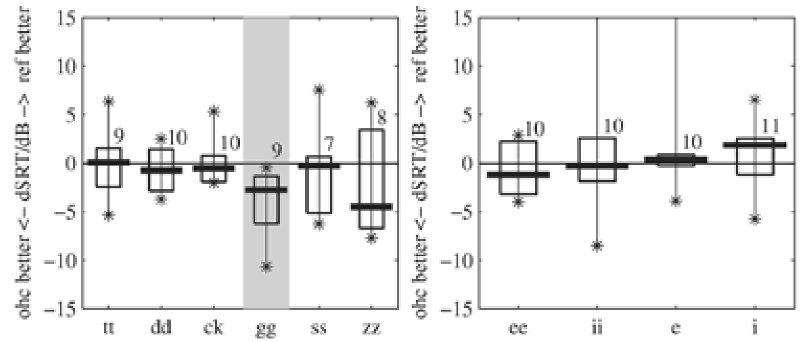

Fig. 4: Difference of individual speech reception thresholds (phoneme tests) between ohc and ref, for VCV phonemes (left panel), and CVC phonemes (right panel).

SUMMARY AND CONCLUSIONS

A gain prescription rule based on the basilar-membrane input-output function in the impaired and normal ear was suggested and evaluated in hearing-impaired listeners. The required estimate of gain loss (caused by outer hair cell damage) can be derived from audiometric thresholds and adaptive categorical loudness scaling using a simple linear model. The prescribed gains were used in a fast-acting “conventional” compressor scheme (gain table) and a model-based compressor scheme. The latter compares the normal and impaired basilar-membrane input-output function in real time (“model in the loop”) and requires a model-based fitting. The combination of model-based algorithm and fitting rule was able compete with an up-to-date commercial system which was fitted by the device vendor, based on the audiogram. The model-based gains were generally higher than those derived from other methods, likely related to the underlying model with independent basilar membrane filters. The results motivate an improved model-based compression algorithm with additional top-down control stages and refined across-frequency control.

AKNOWLEDGMENTS

We thank Linda Grugel (Jean-Uhrmacher-Institut, Köln) and Steffen Kreikemeier (Universitätsklinikum Gießen) for data collection in the evaluation phase. Volker Hohmann continuously contributed to the hearing aid algorithm development. This work was supported by the BMBF (“Modellbasierte Hörsysteme”).

REFERENCES

Brand, T. and Hohmann, V. (**2002**) "An adaptive procedure for categorical loudness scaling" J. Acoust. Soc. Am., **112**, 1597-1604.

Chalupper, J., Fastl, H. (**2002**). "Dynamic loudness model (DLM) for normal and hearing-impaired listeners" Acta Acustica United with Acustica 88, 378-386.

Dillon, H. (**2001**) *Hearing Aids*, Thieme Medical Publishers, New York, NY, USA

Grimm, G., Herzke, T., Berg, D., and Hohmann, V. (**2006**). "The Master Hearing Aid - a PC-based platform for algorithm development and evaluation. Acustica - acta acustica" , **92,** 618-628

Herzke, T., and Hohmann, V. (**2005**). "Effects of instantaneous multi-band dynamic compression on speech intelligibility" EURASIP Journal on Applied Signal Processing, **18**, 3034-3043

Hohmann, V., and Kollmeier, B. (**2007**). "A nonlinear auditory filterbank controlled by sub-band instantaneous frequency estimates" in B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey, editors, *Hearing: from sensory processing to perception* - 14th International Symposium on Hearing , pages 11-18. Springer, Berlin.

Jepsen, M.L., Dau, T., (**2011**). "Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss," J. Acoust. Soc. Am. **129**, 262-281.

Jürgens, T., Kollmeier, B., Brand, T., and Ewert, S. D. (**2011**). "Assessment of auditory nonlinearity for listeners with different hearing losses using temporal masking and categorical loudness scaling" Hearing Research, in Press, [doi:10.1016/j.heares.2011.05.016]

Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (**2003**). "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing" J. Acoust. Soc. Am. **113**, 951–960.

Nelson, D.A., Schroder, A.C., Wojtczak, M. (**2001**). "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners" J. Acoust. Soc. Am. **110**, 2045-2064.

Plack, C. J., and Oxenham, A. J. (**2000**). "Basilar-membrane nonlinearity estimated by pulsation threshold" J. Acoust. Soc. Am. **107**, 501–507.

Plack, C.J., Drga, V., Lopez-Poveda, E.A (**2004**). "Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss" J. Acoust. Soc. Am. **115**, 1684-1695.

Wagener, K., and Brand, T. (**2005**). "Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters", International Journal of Audiology, **44**, 144-156

Wesker, T., Meyer, B., Wagener, K., Anemüller, J., Mertins, A., and Kollmeier, B. (**2005**). "Oldenburg logatome speech corpus (ollo) for speech recognition experiments with humans and machines" In *Proceedings of Interspeech*, pp. 1273-1276, Lisbon, Portugal

The effect of a linked bilateral noise reduction processing on speech in noise performance

JORGE MEIJA[1,2], GITTE KEIDSER[1,2], HARVEY DILLON[1,2], CONG-VAN NGUYEN[1,2] AND EARL JOHNSON[3]

[1] *National Acoustic Laboratories, Sydney, Australia*

[2] *The Hearing Cooperative Research Centre, The University of Melbourne, Australia*

[3] *The USA Department of Veterans Affairs, East Tennessee State University*

Directional processing already provides tangible noise reduction benefits in hearing aids but further improvement is needed for hearing-impaired listeners to communicate as effectively as normal-hearing listeners in noisy environments. The objective of this study was to investigate if a binaurally linked beamformer could further improve the signal-to-noise ratio (SNR). Speech reception thresholds (SRT) and spatial perception were compared for bilaterally fitted cardioid microphones and two binaurally linked beamformer processing conditions; 1) a single audio stream output to the two ears, and 2) two audio stream outputs which preserved spatial cues. 10 normal-hearing and 22 hearing-impaired listeners were recruited for this study. The strategies were implemented on a real-time PC processing platform, wired to a pair of behind-the-ear devices via a sound interface. A speech-in-noise test was administered using the Bamford-Kowal-Bench (BKB) sentences targeting the SNR for which 75% correct keywords were identified in spatially separated multi-talker babble noise and room reverberation. The SNR level at which the listeners acquired 95% intelligibility from continues speech discourse material, using a male and a female talker, was also obtained. Sound amplification was provided according to NAL-NL2. Both beamformer conditions improved the SRTs relative to the conventional cardioids, but by a greater degree for the hearing-impaired listeners, and more convincingly at the higher SRTs.

INTRODUCTION

The understanding of speech in noisy listening situations is extremely challenging for hearing-impaired (HI) listeners. Assuming speech levels that are typical in environments with different noise levels (Pearsons *et al*., 1977), and assuming a typical noise spectrum (Keidser, 1995) speech intelligibility index calculations suggest that HI listeners with moderate losses (averaging 50 dB HL) experience a maximum of 50% intelligibility in moderate background noise levels (Dillon, H. 2010). The electro-acoustic amplification provided by hearing aids result in 90% intelligibility in background noise levels not exceeding 60 dBA. However at background noise levels greater than 70 dBA, electro-acoustic amplification provides no more than 10% intelligibility improvement over the unaided ear. When

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.