Frequency Dependency of Binaural Masking Level Differences in Normal-Hearing and Hearing-Impaired Listeners

Christopher Hauth[1], Thomas Brand[1], Birger Kollmeier[1]

[1] *Medizinische Physik and Cluster of Excellence Hearing4All, CvO Universität Oldenburg,*
26129 Oldenburg, E-Mail:christopher.hauth@uni-oldenburg.de

Introduction

In binaural tone-in-noise detection experiments humans are able to achieve substantially lower thresholds if either the tone or the noise has interaural differences in time (ITD) or level (ILD). This effect is named binaural masking level difference (BMLD) and was mainly investigated for 500 Hz (e.g. [1]). An effective model, which is able to describe the binaural unmasking, is the equalization-cancellation (EC) mechanism [2]. In the EC model, first, left and right ear signal are equalized in level and time based on the interfering signal, which is assumed to be the dominant source. In a second step, the left ear signal is subtracted from the right ear signal leading to an attenuation of the noise due to destructive interferences and an amplification of the target signal due to constructive interferences as long as target and interfering signal differ in their ITD and/or ILD. In the EC model, binaural processing inaccuracies are incorporated, which have been derived from binaural tone-in-noise detection experiments [3], in order to explain perceptual data and to prevent the model from yielding an infinite SNR improvement. The EC model has been combined with the speech intelligibility index (SII) [4] in order to predict binaural speech intelligibility (e.g. [5]). Moreover, individual hearing abilities were accounted for by incorporating information about the audiogram in the model [6]. Even though the audiogram is considered in the model, individual differences in supra-threshold deficits are not incorporated.

This study investigates whether or not binaural tone-in-noise detection experiments can be used to characterize a listener's ability to understand speech in binaural situations and to individualize the binaural processing inaccuracy in the EC model, which can be interpreted as a supra-threshold uncertainty in processing the left and right ear signals. In a first experiment, the frequency dependency of the BMLD is investigated, as so far the binaural inaccuracy was investigated only for 500 Hz. Furthermore, the experiment is conducted by hearing-impaired listeners to investigate whether or not differences in binaural tone detection occur even though the audiogram is similar across subjects. In a second step, binaural speech intelligibility experiments are performed and the results are predicted using the binaural speech intelligibility model by [5], where the EC stage is modified such that the processing inaccuracies can be individualized.

Method

Binaural tone-in-noise detection thresholds were determined for 10 normal-hearing listeners and 5 hearing-impaired listeners with high frequency hearing loss. In Figure 1, the individual audiograms are shown.

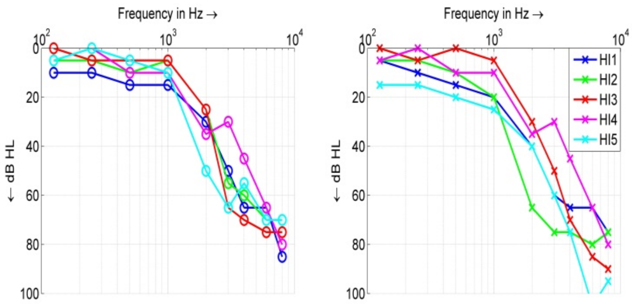

Figure 1: Audiograms of the hearing impaired-listeners. The left and right panels show the audiogram of the right and left ear, respectively.

The tested tone frequencies were 250, 500, 750, 1000, 1500, and 2000 Hz for listeners with normal hearing and 500, 750, 1000, and 1500 Hz for the hearing-impaired listeners. Tone detection was investigated in Gaussian noise, which was bandpass filtered between 100-4000 Hz. The tested interaural delays were selected to mirror fixed interaural phase differences (IPD) for each tested frequency ranging from 0 to 5π in steps of $\pi/2$ leading to frequency dependent ITDs. For a frequency of 500 Hz, ITDs ranging from 0 ms to 5 ms in steps of 0.5 ms were used. Tone detection thresholds were determined in a 3-AFC 1up-2-down procedure converging to the 70.7% correct point on the psychometric function.

In a second step, a binaural speech in noise experiment was conducted with the same listeners. Speech intelligibility experiments were conducted using the Oldenburg Sentence Test (OLSA, [7]) in speech shaped stationary noise. To determine the speech reception threshold (SRT) of 50% intelligibility, an adaptive procedure was used for controlling the level of the speech (Equation 9, [8]). The SRT was determined using test lists of 20 sentences. The test lists were randomly selected out of 45 lists. In order to test binaural processing capabilities, speech and noise were low pass filtered at 1500 Hz to minimize the influence of the audiogram for the hearing-impaired listeners. In total, 4 conditions were tested: monaural left/right, diotic, and dichotic (phase of the noise was inverted between both ears).

Binaural speech intelligibility model (BSIM)

The binaural speech intelligibility model (BSIM) [5] used in this study combines the EC mechanism with the SII to predict binaural speech intelligibility. In a first step, the incoming signal is bandpass filtered by a gammatone filterbank, mirroring the frequency selectivity on the basilar membrane. Afterwards, the EC mechanism is applied in each frequency channel, independently. In the EC model, processing inaccuracies, in level and time are incorporated, which were defined by [3]. Here, only the processing inaccuracy for the compensation of the ITD is shown in more detail, because the binaural tone-in-noise detection experiment with different ITDs of the noise particularly aims at determining the processing inaccuracy in time. In [3], the processing inaccuracy in time is defined as normally distributed random variable with zero mean (μ = 0) and a standard deviation, which is linearly increasing with increasing ITD according to

$$\sigma_\delta = \sigma_{\delta 0} \cdot \left(1 + \frac{|\tau|}{\Delta_0}\right),$$

with $\sigma_{\delta 0} = 65\ \mu s$ and $\Delta_0 = 1.6\ ms$ (see Figure 2).

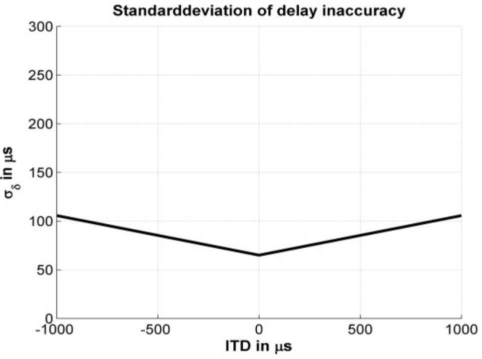

Figure 2: Standard deviation of the delay inaccuracy in µs as a function of ITD as defined by vom Hövel [3].

Results

In Figure 3, the BMLDs obtained in the binaural tone-in-noise detection task for the normal-hearing listeners and the predicted BMLDs using the EC model are shown. Please note that the BMLD shows the relative improvement to the diotic condition, which is set to 0 dB. The BMLD is high if the ITD of the noise results in a phase shift corresponding to odd (1, 3, 5) multiples of π of the tested frequency, and low if the ITD corresponds to phase shifts of even (2, 4) multiples of π. This results in a periodic pattern of BMLD, which can be observed in Figure 3. The largest BMLD of approx. 11 dB can be observed for an ITD of 1ms corresponding to a phase shift of π at 500 Hz, which is in line with literature. With increasing frequency, the maximum achievable BMLD gets lower. This is in line with our knowledge about phase locking, i.e. coding of temporal fine structure in the auditory system, which becomes less effective with increasing frequency. At the lower frequency of 250 Hz, the maximum BMLD is slightly reduced to 9.5 dB compared to 500 Hz and the periodic pattern is less prominent.

The EC mechanism is able to predict the BMLD for 500 Hz and also higher frequencies. For the lower frequency of 250 Hz, the BMLD is predicted to be in the range of 16 dB, where the delay corresponds to a phase shift of π. The measured BMLD in this condition, however, is around 9 dB, resulting in a difference of approx. 7 dB between predicted and measured BMLD. This difference also holds for the tested ITD of 6 ms (3π). While the measured BMLD tends to be smaller compared to the 500 Hz condition, the predicted BMLD is larger. A possible explanation for the reduced BMLD in the 250 Hz condition is a lower (better) diotic threshold compared to the 500 Hz condition, because the auditory filter centered at 250 Hz can be assumed to be narrower and, therefore, providing a better signal-to-noise ratio (SNR), while the dichotic threshold is not much affected.

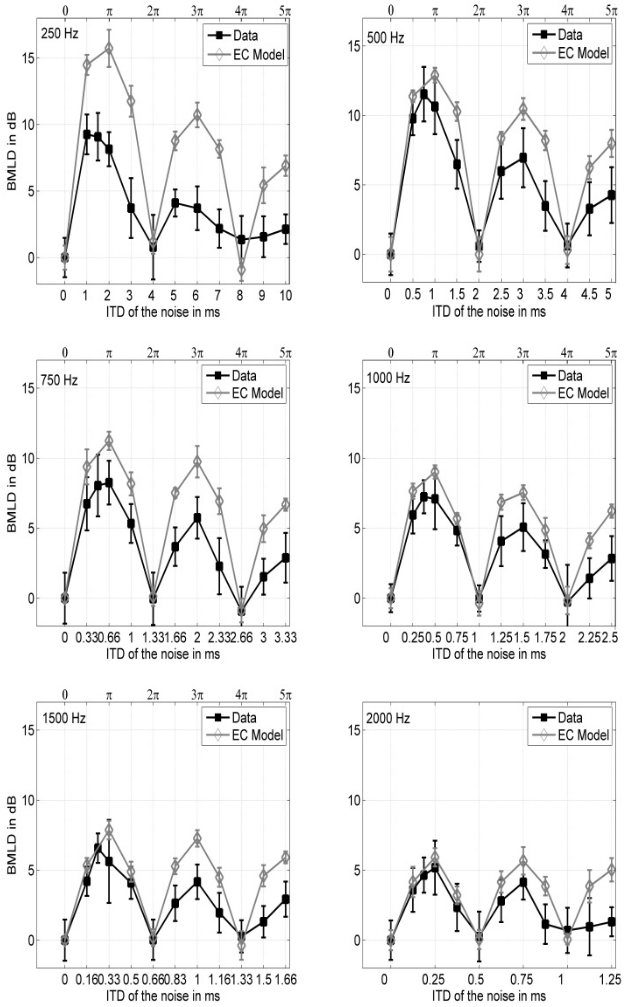

Figure 3: Measured (black) and predicted (gray) BMLDs in dB as a function of the ITD of the noise. Squares and error bars denote mean and standard deviation.

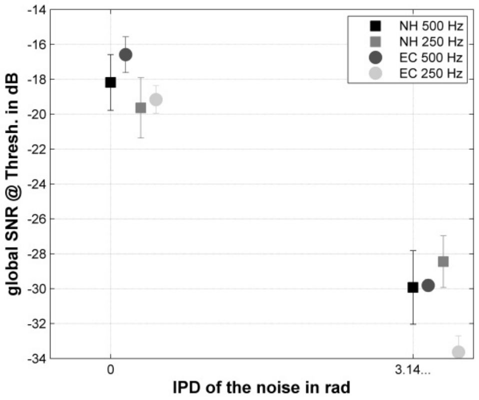

Figure 4: Measured and predicted thresholds in the diotic and dichotic condition for 250Hz and the 500Hz. Squares denote perceptual data, while circles denote model predictions.

In Figure 4, measured and predicted diotic and dichotic thresholds for 250 Hz and 500 Hz are shown. The diotic threshold is lower in the 250 Hz condition than in the 500 Hz condition. However, the EC mechanism shows a larger improvement in the NπS0 condition for 250 Hz than 500 Hz. This is contradicting the perceptual data, which show a slightly worse dichotic threshold for 250 Hz. Therefore, the EC model is overestimating the binaural benefit. These results suggest that the binaural processing inaccuracy results in predictions, which are in line with perceptual data, for frequencies larger than or equal to 500 Hz. On the other hand, predictions of the dichotic conditions overestimate the BMLD for frequencies below 500 Hz suggesting a larger binaural inaccuracy for frequencies below 500 Hz.

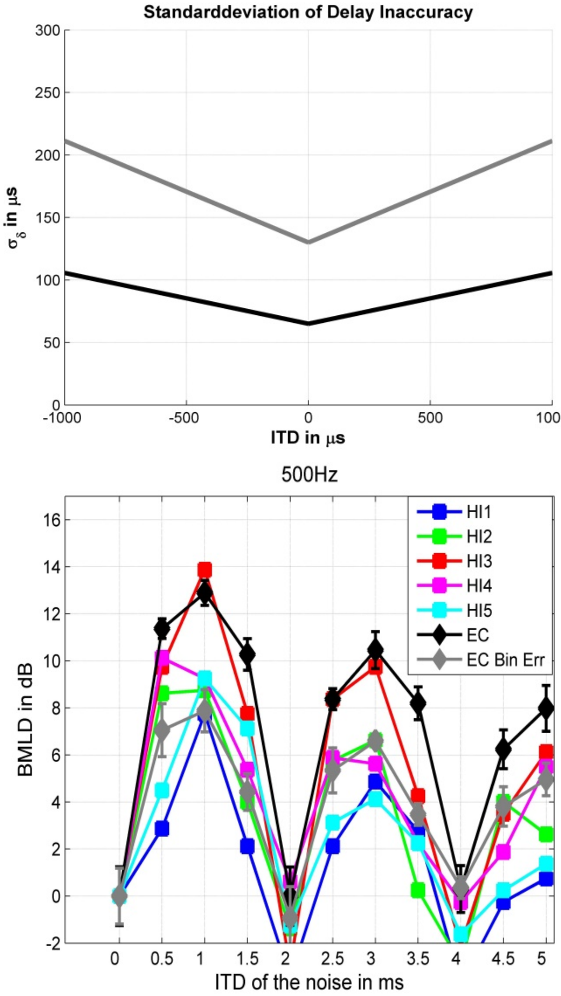

Figure 5: Upper panel: Standard deviation of the binaural processing inaccuracy for the normal-hearing (black) and four of the hearing-impaired subjects (gray). Lower panel: Measured and predicted BMLD in dB at 500 Hz for the hearing-impaired listeners. BMLD were predicted using either a normal-hearing processing inaccuracy (black) or a hearing-impaired processing inaccuracy (gray).

In Figure 5, measured and predicted BMLDs at 500 Hz are shown for the hearing-impaired subjects. Only one of the hearing-impaired listeners performs like a normal-hearing subject and achieves a BMLD of 14 dB for an ITD of 1 ms, even though all subjects can be assumed as normally hearing at 500 Hz. The other listeners show a reduced BMLD of only 8-9 dB in the same condition. The standard deviation of the binaural inaccuracy can be increased, which results in a reduction of the predicted BMLD and a better agreement with the BMLD of the remaining hearing-impaired listeners.

In a follow up experiment, binaural speech intelligibility experiments were performed by the same listeners. The results are shown in Figure 6.

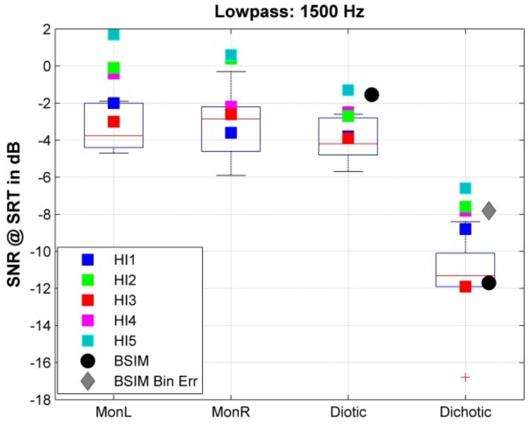

Figure 6: SRTs in dB of normal-hearing listeners (boxplots) and hearing-impaired listeners for monaural, diotic, and dichotic (NπS0) stimulation. Black circles denote standard BSIM predictions; the gray diamond indicates predictions with an increased binaural processing inaccuracy, which was fitted to the BMLDs.

SRTs obtained in normal-hearing listeners are shown as boxplots; SRTs obtained in hearing-impaired listeners are color coded. The median monaural thresholds of the normal-hearing listeners are in the range of -3 to -4 dB SNR. The median diotic threshold is around -4 dB SNR and the median dichotic threshold is around -11 dB SNR.

The results of the hearing-impaired listeners are in the range of the normal-hearing listeners in the diotic condition. In the dichotic condition, one listener (HI3) shows a dichotic thresholds, which can be interpreted as normal-hearing. The same listener also performed best in the psychoacoustical experiment. The remaining subjects show dichotic thresholds that are reduced by 3-4 dB. The diotic and dichotic thresholds are predicted using BSIM. The model was calibrated to the diotic, broadband condition, which results in SRTs of approx. -7.1 dB in normal-hearing listeners if the OLSA is used. The predicted dichotic thresholds fit the normal-hearing data very well, while the diotic threshold is underestimated by 2 dB.

The binaural processing inaccuracy that fits the normal-hearing data in psychoacoustics, is also able to describe the dichotic threshold of the normal-hearing and the best hearing impaired subject in the speech intelligibility experiment. However, the prediction overestimates the BMLD for the remaining hearing-impaired subjects. By adjusting the processing error to fit the psychoacoustical data of the remaining hearing-impaired subjects, predicted dichotic thresholds are decreased by 3 dB, leading to better agreement with perceptual data of the hearing-impaired subjects.

Discussion

Binaural tone-in-noise detection experiments were performed to investigate a frequency dependency of BMLD and to investigate whether or not an individual binaural processing inaccuracy can be derived from this data. It has been shown that the inaccuracy investigated for 500 Hz [3] can be used to predict BMLD also for higher frequencies. In the perceptual data it was observed that the binaural unmasking becomes less effective with increasing frequency. As an accurate coding of the temporal fine structure is required by the binaural system to achieve a high BMLD, the loss of phase locking with increasing frequency is assumed to limit the performance in the dichotic conditions of binaural tone-in-noise experiments. The same trend is captured by the EC model, even though phase locking is not explicitly accounted for in the model. However, the processing inaccuracies are implemented in the time domain. With increasing frequency, the effect of the inaccuracy gets larger as the same phase relation between both ears is achieved at a much smaller delay. Therefore, the impact of the delay inaccuracy gets larger with increasing frequency, leading to a less accurate equalization in time and, consequently, to a less effective cancellation of the interfering source or noise.

At 250 Hz, a larger deviation between the perceptual data and the predicted BMLD was observed. In the dichotic condition, the thresholds for 250 Hz are only slightly worse than in the 500 Hz condition. However, the EC mechanism predicts the thresholds to be even better than at 500 Hz. Therefore, the processing inaccuracy needs to be increased at 250 Hz. This suggests a bandpass characteristic of binaural unmasking rather than a low pass characteristic, which occurs if only phase locking is considered as limiting factor in binaural unmasking. The periodic pattern with increasing ITD, which can be observed at 500 Hz and higher frequencies, is less prominent at 250 Hz. This might be caused by the relatively large delays of more than 5 ms, which are tested in this condition.

Moreover, a large variance in BMLD data in hearing-impaired listeners was observed, who either showed a normal-hearing BMLD (one subject) or a reduced BMLD (four subjects). The processing inaccuracy in the EC model of BSIM was modified in order to account for the reduced BMLD in the hearing-impaired subjects. In the binaural speech intelligibility experiment, it was shown that those hearing-impaired listeners, who show a reduced BMLD in a tone detection task, also show reduced dichotic SRTs. The hearing-impaired listener with normal-hearing BMLDs also showed a normal-hearing dichotic SRT. By adjusting the processing inaccuracy to the psychoacoustic data, the predicted SRT in the dichotic condition is increased and approaches the measured SRT for the hearing-impaired listeners.

These results indicate that the binaural inaccuracy in the EC model can be used to improve individual speech intelligibility predictions using binaural tone-in-noise detection data.

Conclusion

In this study it was shown that the processing inaccuracy in EC processing suggested by [3], which was already shown to account for binaural tone-in-noise detection experiments at 500 Hz, also holds for higher frequencies probably as it implicitly mirrors the loss of phase locking in the auditory system with increasing frequency. For frequencies below 500 Hz, this does not hold. Here, the processing inaccuracy of the model needs to be increased to account for dichotic tone-in-noise detection thresholds leading to a bandpass characteristic of binaural unmasking in the EC process rather than a low pass characteristic. Furthermore, it was shown that binaural tone-in-noise detection tasks can be used to derive individual binaural processing inaccuracies that help to improve individual binaural speech intelligibility predictions.

Literature

[1] Egan, J.P. (1965). “Masking-level differences as a function of interaural disparities in the intensity of signal and noise,” The Journal of the Acoustical Society of America 38, 1043-1049.

[2] Durlach, N. I. (1963). “Equalization and Cancellation Theory of Binaural Masking Level Differences,” The Journal of the Acoustical Society of America 35(8), 1206–1218.

[3] vom Hövel, H. (1984). Zur Bedeutung der Übertragungseigenschaften des Aussenohrs sowie des binauralen Hörsystems bei gestörter Sprachübertragung.

[4] ANSI (1997). “Methods for the calculation of the speech intelligibility index,” American National Standard S3.5-1997, Standards Secretariat, Acoustical Society of America.

[5] Beutelmann, R., Brand, T., and Kollmeier, B. (2010). “Revision, extension, and evaluation of a binaural speech intelligibility model,” The Journal of the Acoustical Society of America 127(4), 2479-2497.

[6] Beutelmann R., Brand, T. (2006). “Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners,” The Journal of the Acoustical Society of America 120, 331-342.

[7] Wagener, K., Brand, T., Kühnel, V., and Kollmeier, B. (1999a). “Entwicklung und Evaluation eines Satztests für die Deutsche Sprache I: Design des Oldenburger Satztests (Development and evaluation of a sentence test for the German language I: Design of the Oldenburg sentence test),” Z. für Audiologie, Audiological Acoust. 38, 4–15.

[8] Brand, T., and Kollmeier, B. (2002). “Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests,” J. Acoust. Soc. Am. 111(6), 2801–2810.