Modeling the Effects of Sensorineural Hearing Loss on Sound Localization in the Median Plane

Robert Baumgartner[1], Piotr Majdak[1], and Bernhard Laback[1]

Abstract

Listeners use monaural spectral cues to localize sound sources in sagittal planes (along the up-down and front-back directions). How sensorineural hearing loss affects the salience of monaural spectral cues is unclear. To simulate the effects of outer-hair-cell (OHC) dysfunction and the contribution of different auditory-nerve fiber types on localization performance, we incorporated a nonlinear model of the auditory periphery into a model of sagittal-plane sound localization for normal-hearing listeners. The localization model was first evaluated in its ability to predict the effects of spectral cue modifications for normal-hearing listeners. Then, we used it to simulate various degrees of OHC dysfunction applied to different types of auditory-nerve fibers. Predicted localization performance was hardly affected by mild OHC dysfunction but was strongly degraded in conditions involving severe and complete OHC dysfunction. These predictions resemble the usually observed degradation in localization performance induced by sensorineural hearing loss. Predicted localization performance was best when preserving fibers with medium spontaneous rates, which is particularly important in view of noise-induced hearing loss associated with degeneration of this fiber type. On average across listeners, predicted localization performance was strongly related to level discrimination sensitivity of auditory-nerve fibers, indicating an essential role of this coding property for localization accuracy in sagittal planes.

Introduction

Monaural spectral cues enable sound localization where binaural cues are ambiguous. This ambiguity occurs approximately within sagittal planes, that is, vertical planes orthogonal to the interaural axis and thus concerns localization in the up-down and front-back directions (Macpherson & Middlebrooks, 2002; Wightman & Kistler, 1997). The extraction of spectral localization cues from the acoustic signal relies on proper functioning of the auditory periphery. Conductive hearing loss degrades localization performance of sounds mainly within horizontal planes (Noble, Byrne, & Lepage, 1994), whereas sensorineural hearing loss degrades localization performance of sounds especially within sagittal planes (Dobreva, O'Neill, & Paige, 2011; Otte, Agterberg, Wanrooij, Snik, & Opstal, 2013; Rakerd, Vander Velde, & Hartmann, 1998). An important

factor behind this degradation is the decrease in high-frequency sensitivity (Baumgartner, Majdak, & Laback, 2014; Best, Carlile, Jin, & van Schaik, 2005). By itself, however, this factor does not fully explain individual variations in localization performance of hearing-impaired listeners with similar hearing loss (Noble et al., 1994).

To better understand the effect of sensorineural hearing loss on localization performance in the sagittal planes, the present study aimed at simulating the consequences of this hearing loss by means of a computational

[1]Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria

Corresponding author:
Robert Baumgartner, Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, A-1040 Vienna, Austria.
Email: robert.baumgartner@oeaw.ac.at

auditory model. To this end, we integrated a phenomenological model of the auditory periphery (Zilany, Bruce, & Carney, 2014; Zilany, Bruce, Nelson, & Carney, 2009), which has already been successfully used as a front-end of models for human tone-in-noise detection (Mao & Carney, 2014) and speech intelligibility (Mamun, Jassim, & Zilany, 2015), into a model of sound localization in sagittal planes for normal-hearing listeners (Baumgartner et al., 2014).

Cochlear damage can involve dysfunction of both inner hair cells (IHCs) and outer hair cells (OHCs). In most etiologies, the OHC damage is more pronounced. It reduces the cochlear gain as well as the spectral resolution of the auditory system (Moore, 1995, p. 22), and naturally diminishes the beneficial modulatory effects of the descending efferent system on localization performance in noise (Andéol et al., 2011; May, Budelis, & Niparko, 2004). Besides cochlear damage, sensorineural hearing loss can also be associated with a degeneration of auditory-nerve (AN) fibers, often called auditory deafferentation, which may be selective to certain fiber types categorized by their thresholds or spontaneous rates (SRs; Furman, Kujawa, & Liberman, 2013; Liberman, 1978). A SR-selective degeneration of AN fibers changes the rate-level function of the AN fiber population and thus is supposed to affect spectral-shape representations (Reiss, Ramachandran, & May, 2011) and localization performance (Macpherson & Sabin, 2013).

In our study, we first evaluated the updated localization model using existing experimental data from normal-hearing listeners. Then, we simulated the effects of several types of OHC dysfunction (considering the IHCs to be intact) on median-plane sound localization and the particular role of specific AN fiber types. Finally, we assessed the correlation between the predicted average localization performance and the fiber's sensitivity in level discrimination because it is hypothesized that level discrimination is a good indicator for localization performance at various sound pressure levels (SPLs).

Methods

### Coordinate System and Measures of Localization Performance

To describe auditory localization, we use the interaural-polar coordinate system with the lateral angle relative to the median plane and the polar angle around the interaural axis within a sagittal plane (Morimoto & Aokata, 1984). The lateral angle ranges from $-90°$ on the left-hand side to $90°$ on the right-hand side. The polar angle ranges from $-90°$ to $270°$ with $0°$ corresponding to the front, $90°$ to the top, and $180°$ to the back of the listener.

Localization errors were solely evaluated in the polar-angle dimension and denote the differences between target and response polar angles. Errors were classified as local errors if they were smaller or equal to $90°$ or as quadrant errors if they were larger than $90°$. Quadrant errors thus refer to responses in the hemisphere opposite to the target and include confusions between front and back or up and down. According to these two error types, localization performance was quantified by the percentage of quadrant errors and the root mean square (RMS) of local errors, which comprises both response accuracy and response precision.

Localization Model

Basic functionality. The model from Baumgartner et al. (2014) follows a template-based comparison procedure— see Figure 1. First, the incoming target sound and the direction-specific templates of head-related transfer functions (HRTFs) are processed by an auditory-periphery model in the peripheral processing stage
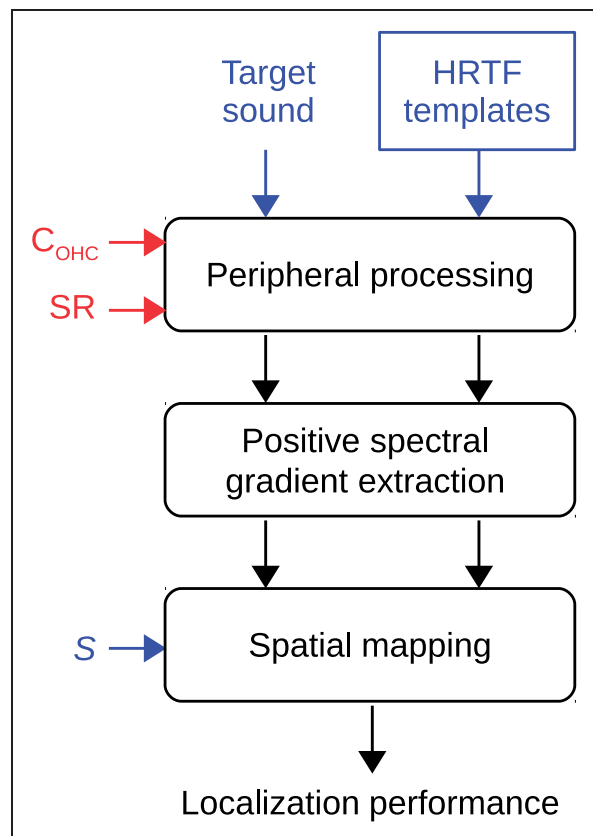
Figure 1. Structure of the localization model. Peripheral processing is approximated by the auditory periphery model from Zilany et al. (2009, 2014). Red: the cochlear gain (C_{OHC}) and the spontaneous rate (SR) of AN fibers were varied to study different aspects of sensorineural hearing loss. Blue: listener-specific HRTFs and values of the sensitivity parameter (S) were used to account for inter-individual differences.

whose output consists of temporally integrated firing rates. The across-frequency distributions of these rates are called rate profiles. Then, these rate profiles are transferred to positive spectral gradient profiles by a differentiation across frequency with a spacing of one equivalent rectangular bandwidth (ERB) and a nonlinear mapping that eliminates negative gradients. This gradient extraction stage is a functional approximation of the rising spectral edge sensitivity observed in the dorsal cochlear nucleus (Reiss & Young, 2005). In the following spatial mapping stage, the positive spectral gradient profile of the incoming target sound, called target profile, is compared with those of the template HRTFs, called template profiles. The more similar the target profile is to a certain template profile relative to all other template profiles, the higher is the predicted probability of the listener responding at the direction associated with that template. The comparison of the target profile with all template profiles consequently yields a probabilistic prediction of the listener's distribution of polar-angle responses to a target sound. This probability distribution can finally be used to calculate expectancy values of commonly used measures of localization performance.

The spatial mapping stage is controlled by a listener-specific sensitivity parameter S, which influences the predicted probability distribution in such a way that an increase of S results in a shallower distribution. Aside from the listener's HRTFs, S is the only listener-specific parameter of the model. Since shallower probability distributions are usually associated with poorer localization performance, S allows calibration of the model output to listener-specific localization performance (Majdak, Baumgartner, & Laback, 2014).

Modifications followed by integration of auditory-periphery model. In Baumgartner et al. (2014), we were able to predict several effects of HRTF modifications and spectral variations of the sound source on localization performance. In that study, we used a linear Gammatone filterbank to model the auditory periphery. A more realistic model of the auditory periphery is required for modeling the effects of individual hearing impairment. Thus, in the present study, we replaced the Gammatone filterbank with the humanized version of a nonlinear auditory-periphery model (Zilany et al., 2009, 2014). Humanization in this model version includes adjustments of the middle-ear filter, the basilar membrane tuning, a frequency-offset of the control-path filter representing the cochlear amplifier mechanism, and the relationship between latency and characteristic frequency (CF; Ibrahim & Bruce, 2010). The model allows for the simulation of various degrees of OHC dysfunction by adjusting the OHC gain C_{OHC}, where $0 \leqslant C_{OHC} \leqslant 1$ linearly scales the Q_{10} of the tuning curve between normal OHC function ($C_{OHC} = 1$) and complete OHC damage

($C_{OHC} = 0$; Bruce, Sachs, & Young, 2003). Furthermore, the auditory-periphery model distinguishes between low-, medium-, and high-SR fibers by adding different amounts of fractional Gaussian noise within its IHC-to-AN synapse approximation (Zilany et al., 2009).

Rate-level curves for the different fiber types are shown in Figure 2. To obtain these curves, we evaluated the temporally averaged firing rate of each fiber type at a CF of 4 kHz in response to Gaussian white noise at various SPLs. For all our simulations, we directly used the output of the synapse model (i.e., no spike generator involved) and set the internal sampling rate of the auditory-periphery model to 100 kHz. In Figure 2, the high-SR fiber shows relatively high spontaneous activity, a low threshold followed by a sharp increase up to about 60 dB SPL, and saturation across a wide range of SPLs. The medium-SR fiber shows little spontaneous activity below a threshold of about 30 dB SPL, a sharp increase between 40 and 90 dB SPL, and a mild saturation above. The activity of the low-SR fiber increases quite constantly with SPLs above about 40 dB and shows no saturation at the highest tested SPLs.

Rate profiles for localization predictions were evaluated across the relevant frequency range from 0.7 to 18 kHz (Algazi, Avendano, & Duda, 2001). This range was represented by 28 AN fibers of each type with CFs equally spaced on the equivalent rectangular bandwidth scale. Temporally averaged firing rates were averaged across fiber types according to their physiological prevalence of 61% high-, 23% medium-, and 16% low-SR fibers (Liberman, 1978).

The change of the peripheral processing stage of the localization model required adjustments also in
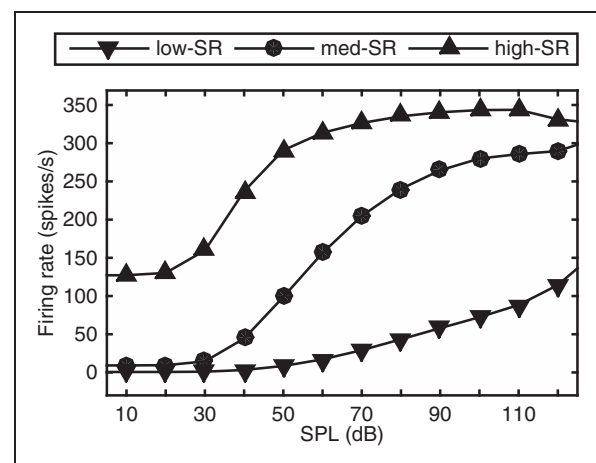
Figure 2. Rate-level curves of the three different fiber types represented in the auditory-periphery model. Firing rates were evaluated at a CF of 4 kHz in response to Gaussian white noise at various SPLs. Note that high-SR fibers saturate already at low SPLs, medium-SR fibers at moderate SPLs, and low-SR fibers not at all.

subsequent model stages. In the spatial mapping stage, availability of actual AN rate predictions allowed us to account for neural properties of rate discrimination. The discriminability between spectral cues is expected to decrease with increasing rate as the discharge variability, σ, of AN fibers (irrespective of fiber type or CF) increases with rate r according to the power law (May & Huang, 1997):

$$\hat{\sigma}(r) = 2.6r^{0.34}. \tag{1}$$

Following signal detection theory (Green & Swets, 1966), the discriminability between two normally distributed random variables x and y is given by

$$d'(x, y) = \frac{x - y}{\sqrt{\sigma(x)^2 + \sigma(y)^2}}. \tag{2}$$

In Baumgartner et al. (2014), the localization model compared target and template representations on the basis of differences between positive spectral gradients. Negative gradients were mapped to zero which introduces a strong bias in the distribution of gradients. To obtain statistical properties adequate for Equation (2), in the updated model, we mapped the gradient g to positive gradients g_p by using a mapping function that saturates for negative g and large positive g:

$$g_p = K \arctan(g/K - \pi/2) + \pi/2. \tag{3}$$

The constant K stretches the mapping function such that best sensitivity to gradient changes occurs at $K\pi/2$. K of 6 yielded the smallest prediction residues and was used throughout this study. The changes to the model also required a re-calibration of the listener-specific sensitivity parameter S.

Listeners and Stimuli

Model simulations represent 14 female and 9 male listeners aged between 19 and 46 years. At the time of the experiments, all listeners had absolute hearing thresholds within the 20-dB range of the average normal-hearing population in the frequency range between 0.125 and 12.5 kHz (ANSI, 2010; Goupell, Majdak, & Laback, 2010; Majdak, Walder, & Laback, 2013). The model was calibrated to each listener's localization performance as obtained in a baseline condition with 500-ms Gaussian white noise bursts filtered by listener-specific HRTFs. Presentation level was 50 ± 5 dB above the individual hearing threshold. We pooled the baseline data across four studies (Goupell et al., 2010; Majdak, Goupell, & Laback, 2010; Majdak, Masiero, & Fels, 2013; Majdak, Walder, et al., 2013) that all tested this particular

baseline condition in order to increase the number of listeners and the reliability of the calibration procedure. Calibration was performed by adjusting the listener-specific sensitivity parameter S of the model such that the difference between actual and predicted localization performance was minimized in a least-squared-error sense (Baumgartner et al., 2014). For our 23 listeners, we obtained sensitivity parameter values in the range of $1.2 < S < 2.5$. Figure 3 shows the actual and predicted baseline performance for all individual listeners after calibration. Pearson's correlation coefficients of 0.88 ($p < .001$) for the quadrant errors and 0.86 ($p < .001$) for the local errors indicate strong correspondence between actual and predicted baseline performance.

The stimuli used for the simulations of hearing impairments were Gaussian white noise bursts with a duration of 170 ms. Informal tests have shown that the rate profiles differed only marginally for longer stimuli. Targets were simulated in the median plane for polar angles between $-30°$ and $210°$.

Simulated Conditions of Hearing Impairment

Different degrees of OHC dysfunction were simulated, ranging from normal function ($C_{OHC} = 1$) to complete dysfunction ($C_{OHC} = 0$, Table 1). Two intermediate degrees of OHC dysfunction were selected according to their impact on absolute hearing thresholds. To estimate the hearing thresholds corresponding to OHC gains, we first determined the threshold firing rate for the normal-hearing condition ($C_{OHC} = 1$) as a function of CF by evaluating the firing rate predicted in response to a pure tone at the CF with a SPL matching the standardized hearing threshold (Table 6 in ANSI, 2010). Then, for OHC dysfunctions, the absolute hearing threshold was determined based on the SPL required to exceed the threshold firing rate determined in the normal-hearing condition. The estimated hearing thresholds in Figure 4 were selected as simulation conditions for
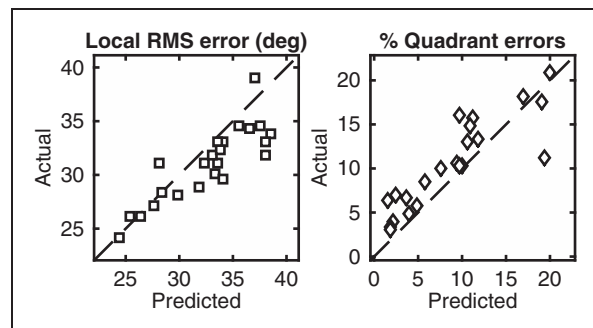
Figure 3. Correspondence between actual and predicted baseline performance for the 23 normal-hearing listeners after listener-specific calibration of the model's sensitivity parameter (*S*).

Table 1. Simulated Conditions of OHC Dysfunction, Estimated PTAs, and Corresponding Hearing Loss Categories.

OHC gain	OHC functionality	Low-f. PTA[a]	High-f. PTA[b]	Hearing loss category
$C_{OHC} = 1.0$	Intact	0 dB	0 dB	Normal
$C_{OHC} = 0.4$	Moderate dysfunction	14 dB	16 dB	Normal
$C_{OHC} = 0.1$	Severe dysfunction	28 dB	37 dB	Mild
$C_{OHC} = 0.0$	Complete dysfunction	37 dB	54 dB	Moderate

Note. OHC = outer hair cell; PTA = pure-tone averages.
[a]Low-f. (frequency) PTA evaluated at 0.5, 1, and 2 kHz.
[b]High-f. PTA evaluated at 4, 8, and 11 kHz.
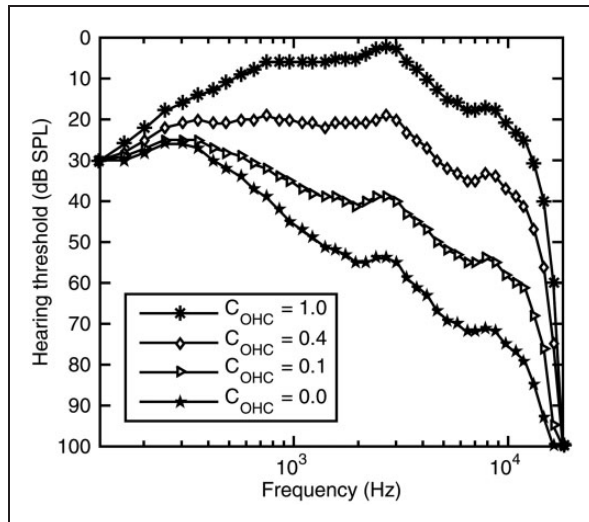
Figure 4. Hearing thresholds estimated for simulated OHC gains (C_{OHC}) within the range of 1 (normal active cochlea) to 0 (passive cochlea). The selected set of OHC gains results in approximately equal increments of high-frequency thresholds.

localization predictions because the OHC gains of $C_{OHC} = 0.4$ and $C_{OHC} = 0.1$ yield about equidistant intermediate levels of hearing loss between $C_{OHC} = 1$ and $C_{OHC} = 0$. Clinical categories of hearing loss were assigned to the OHC gains on the basis of pure-tone averages (PTAs; Goodman, 1965). We evaluated PTAs for the commonly used triplet of 0.5, 1, and 2 kHz as well as for a higher triplet of 4, 8, and 11 kHz (Otte et al., 2013). We finally used the high-frequency PTAs to assign the hearing loss categories because this frequency range is more important for sagittal-plane sound localization. If the low-frequency PTAs had been used instead, complete OHC damage would correspond to mild rather than moderate hearing loss. Correspondences for all other OHC dysfunctions would remain the same.

We simulated localization experiments for all fiber types combined and for each fiber type separately. The model templates were processed with the same OHC gains and fiber types used for the target sounds.

This represents the situation of a perfect adaptation of the auditory system to the hearing impairment.

To investigate the coding properties of the different AN fiber types, we evaluated their sensitivity in level discrimination based on rate differences. To this end, we simulated AN responses to broadband noise at various SPLs in steps of 10 dB and averaged the predicted firing rates across the frequency range from 700 Hz to 18 kHz. A fiber's discriminability $d'(r_i, r_j)$ at SPL i to a SPL change $i - j$ is, according to Equation (2), the difference in the firing rates r_i, r_j between two SPLs i and j, relative to the expected variance of the corresponding firing rates (see Equation (1)). The location of the maximum d' indicates the SPL providing best discriminability and the SPL range for $d' > 0$ indicates the dynamic range of the fiber type.

For the statistical analysis of main effects of OHC gain and fiber type activity on localization performance, we performed two-way repeated-measures analyses of variance (ANOVAs) with Greenhouse-Geisser correction for departure from sphericity. For post hoc analyses, we used Tukey's honest significance difference tests. All effects are reported as significant at the level of $p < .001$. Statistical analyses were carried out with the Statistics and Machine Learning Toolbox from MATLAB (The Mathworks, Natick, MA, USA).
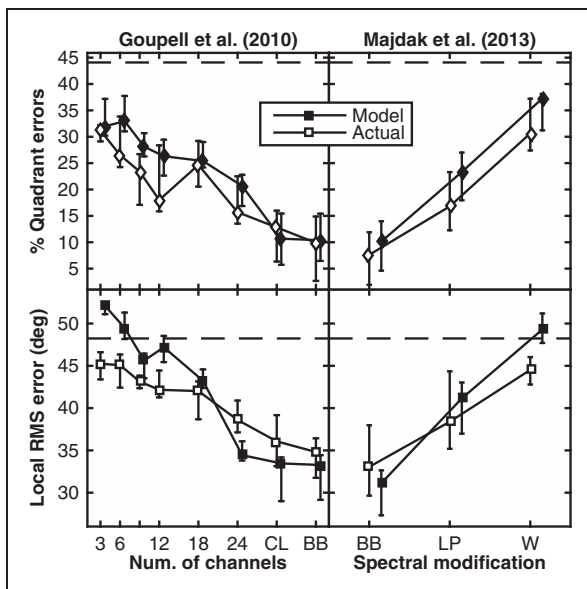
Results

Localization Model Predicts Effects of Spectrally Modified HRTFs for Normal-Hearing Listeners

We evaluated the model for two experiments with normal-hearing listeners whose localization performance in the polar-angle dimension was tested with spectrally modified HRTFs (Goupell et al., 2010; Majdak, Walder, et al., 2013). In both experiments, the stimuli were presented at a sensation level of about 50 dB relative to the hearing threshold for a frontal target sound. In Goupell et al. (2010), the effect of reduced spectral resolution was tested by varying the number of spectral channels used in a Gaussian envelope tone vocoder. Tested numbers of channels ranged from 3 to 24. As baseline conditions,

the listeners localized broadband noise bursts (BB) and click trains (CL). Both baseline conditions refer to an unlimited number of channels, that is, no reduction of spectral resolution. The CL condition provided the same long-term magnitude spectrum as the BB condition with phase characteristics similar to the vocoded stimuli. In the actual experiment, listeners performed worse as fewer channels were used, as shown by the open symbols in the left column of Figure 5. In Majdak, Walder, et al. (2013), listeners were tested with broadband HRTFs (BB; up to 16 kHz) and HRTFs band-limited up to 8.5 kHz, either by low-pass filtering (LP) or by warping (W) the frequency range between 2.8 and 16 kHz to 2.8 and 8.5 kHz. The listeners performed best with the broadband HRTFs, worse with the low-pass filtered HRTFs and worst with the warped HRTFs, as shown by the open symbols in the right column of Figure 5.
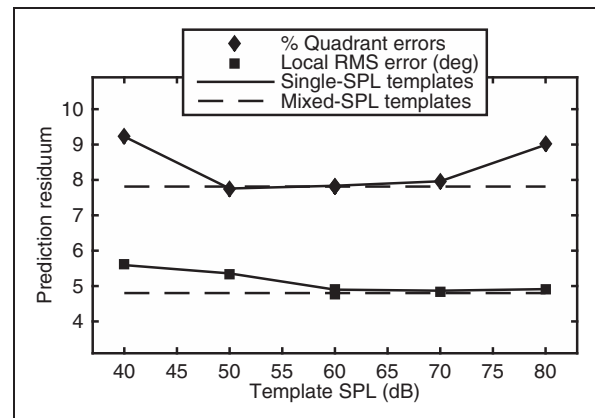
Simulations of these two experiments were performed with the actual participant's HRTFs and target directions. We assumed the absolute hearing thresholds for frontal targets to be at 10 dB SPL (Sabin, Macpherson, & Middlebrooks, 2005) and thus simulated the target stimuli at 60 dB SPL. To find the most reasonable SPLs for the template representation, we simulated the

two experiments for either representing the templates at a single SPL between 40 and 80 dB (in steps of 10 dB) or representing the templates as a mixture across these SPLs. The mixture was calculated by averaging the rate profiles evaluated at all the SPLs individually. For each template setting, the listener-specific parameter value S was calibrated (according to baseline performance), and the predictive power of the model was quantified by prediction residues, which evaluate the RMS differences between actual and predicted listener-specific performance measures. The residues were averaged across listeners and experimental conditions according to the number of trials in the corresponding actual experiments. Figure 6 shows the prediction residues as functions of the template SPL. Results for the mixed-SPL templates are shown by dashed horizontal lines. In general, the prediction residues were below 10% in quadrant errors and below 6° in local RMS error. This predictive power of the updated model is comparable to the normal-hearing model (Baumgartner et al., 2014). Moreover, the prediction residues varied only within a small range indicating that the choice of the template SPL was not crucial for the predictive power. Nevertheless, the mixed-SPL template yielded the smallest residues for both error measures and appears more plausible than templates tuned to a single SPL. We thus used the mixed-SPL template for further simulations.

The predictions (based on the mixed-SPL template) are shown as closed symbols in Figure 5. As indicated by the small prediction residues from Figure 6, predicted

Figure 5. Model evaluation for normal-hearing listeners tested on the effects of spectral resolution (by number of vocoder channels in Goupell et al., 2010) and spectral warping (Majdak, Walder, et al., 2013). Model data (filled circles) are compared with actual data (open circles) from the two studies. Error bars represent SDs. Symbols are slightly shifted along the abscissa for better visibility. BB = broadband noise burst; CL = broadband click train (infinite number of channels); LP = low-pass filtered at 8.5 kHz; W = HRTFs spectrally warped from 2.8 to 16 kHz to 2.8 to 8.5 kHz.


Figure 6. Effect of template SPL on predictive power of the model for the two studies (Goupell et al., 2010; Majdak, Walder, et al., 2013) shown in Figure 5. Predictions based on a single template SPL equivalent to the actual SPL of the target sounds of 60 dB result in similar prediction residues as based on templates mixed across a broad range of SPLs. Higher plausibility of the mixed-SPL templates was the reason to choose this representation for all further simulations (including predictions shown in Figure 5).

localization errors were quite consistent with the actual results, although it seems that the model tends to over-estimate the local RMS error, especially for very strong HRTF modifications, for example, induced by spectral warping or using very few spectral channels.

Predicted Localization Degrades With OHC Dysfunction and Depends on AN Fiber Types

Figure 7 shows the predicted localization performance for various combinations of OHC gain and AN fiber type in terms of quadrant error rates (top row) and local RMS errors (bottom row). The left-most column labeled "all SRs" shows simulations based on the whole population of AN fibers. The other columns refer to simulations based on either low-, medium-, or high-SR fibers alone. Within each column, OHC functionality is plotted as degrading from left to right. The normal-hearing baseline performance is depicted in the left-most condition, representing $C_{OHC} = 1.0$ with all SRs. The performances for the various conditions range from better-than-baseline (e.g., med-SR, $C_{OHC} = 1.0$) to chance (dashed horizontal line).

Significant main effects of the OHC gain were found for quadrant errors, $F(1.62, 35.6) = 716$; $p < .001$; $\eta_p^2 = .970$, and local errors, $F(2.09, 46.1) = 610$; $p < .001$;
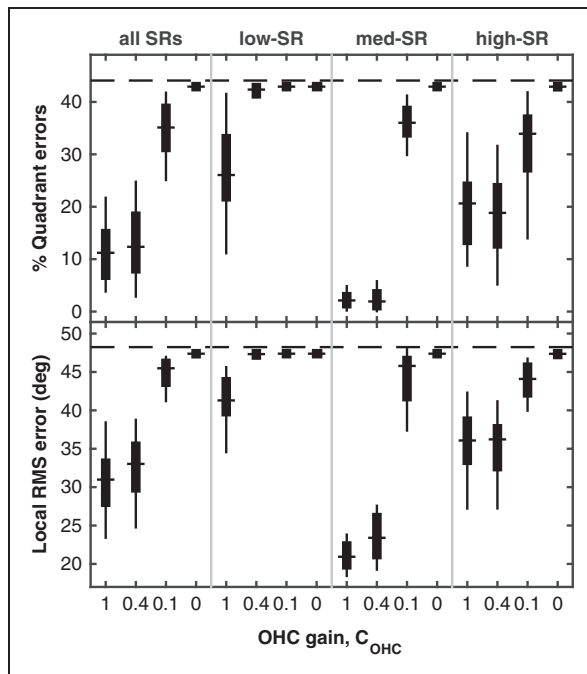
$\eta_p^2 = .965$. Post hoc tests revealed that both error types were significantly different between all levels of OHC gain and monotonically increased with decreasing gain. Significant main effects of fiber type activity were found for quadrant errors, $F(1.17, 25.8) = 370$; $p < .001$; $\eta_p^2 = .944$, and local errors, $F(1.49, 32.7) = 838$; $p < .001$; $\eta_p^2 = .974$. Post hoc tests showed significant differences between all combinations of fiber types. In particular, predicted quadrant error rates and local RMS errors were smallest (best) for medium-SR fibers, larger for all fibers, even larger for high-SR fibers, and largest (worst) for low-SR fibers.

The factors OHC gain and fiber type also showed a significant interaction, both for quadrant errors, $F(1.84, 40.5) = 187$; $p < .001$; $\eta_p^2 = .895$, and local errors, $F(2.23, 49.1) = 279$; $p < .001$; $\eta_p^2 = .927$. The degradation induced by OHC dysfunction was most pronounced for medium-SR fibers, less for all fibers combined, even less for high-SR fibers, and least for low-SR fibers. Since this order of interaction strength is in line with the main effect of fiber type activity, it may partly be due to ceiling effects.

Sensitivity in Level Discrimination Is Correlated With Predicted Average Localization Performance

Figure 8 shows the predicted sensitivities of AN fibers to SPL changes of 10 dB for the four different degrees of OHC dysfunction. Since the firing rates of high- and medium-SR fibers saturated at high SPLs, the sensitivity curves of these fiber types are bell-shaped and show a clear maximum. Compared with the high-SR fibers, the sensitivity curve of the medium-SR fibers generally appears slightly larger and shifted to higher SPLs. This resulted in a larger dynamic range and higher sensitivities at higher SPLs. In contrast to both high- and medium-SR fibers, the sensitivity of low-SR fibers is mostly lower, but gradually increases with SPL and thus provides a very broad dynamic range and best sensitivity for very high SPLs. The physiologically weighted combination of the three fiber types results in a broad, shallow sensitivity curve with a peak around 50 dB SPL (dashed line). A reduction of OHC gain shifts all sensitivity curves to higher SPLs, with the peak moving up to around 90 dB SPL. Note that the first incidence of positive sensitivity is equivalent to the previously estimated hearing thresholds (Table 1).

The gray area centered at 60 dB SPL displays the approximate range of the stimuli used for the simulations (±10 dB, c.f., Figure 2 from Baumgartner, Majdak, & Laback, 2013). Each fiber's sensitivity within this range appears to correspond to the predicted localization performance described earlier. For instance, medium-SR fibers are the most sensitive type and yield best localization performance for $C_{OHC} \geqslant 0.4$. In case of severe OHC

Figure 7. Effects of OHC dysfunctions and selective activity of AN fibers on predicted quadrant error rates (top) and local RMS errors (bottom). Thick bar: interquartile range (IQR). Thin bar: data range within 1.5 IQR. Horizontal line within thick bar: average. Dashed horizontal line: chance performance.
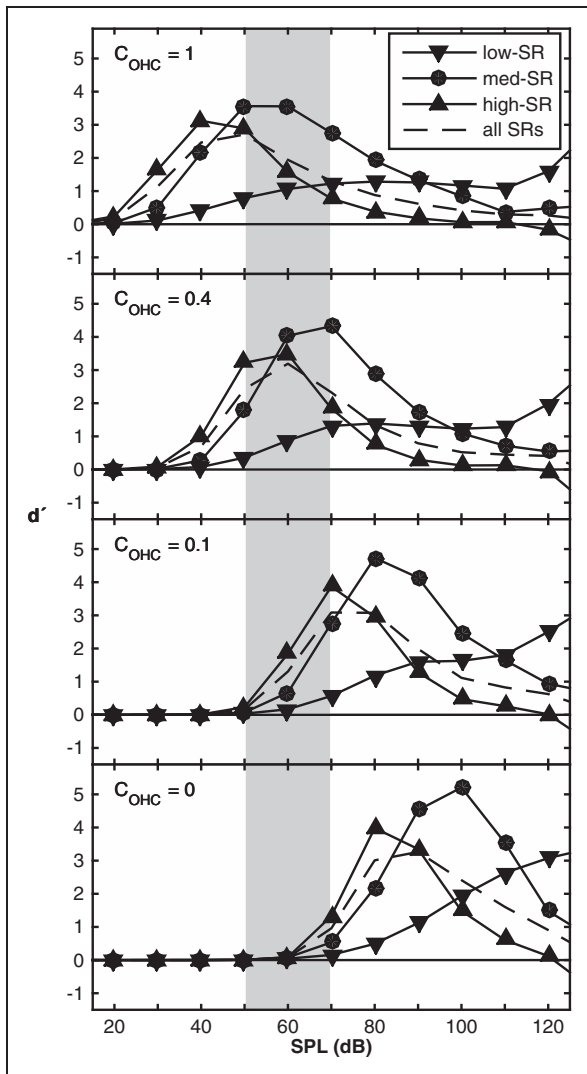
Figure 8. Sensitivity (d') of AN fibers in level discrimination as function of SPL predicted for different fiber types and OHC dysfunctions. Sensitivities were evaluated for SPL increments of 10 dB and averaged across 28 CFs from 0.7 to 18 kHz. Gray area: stimulus range of target sounds at 60 dB SPL.

dysfunction ($C_{OHC} = 0.1$), high-SR fibers provided highest sensitivity and, in line with that, predicted localization performance was best for this fiber type. In the case of complete OHC damage ($C_{OHC} = 0$), all fibers were almost insensitive to level changes and predicted localization performance was close to chance performance. Analysis of Pearson correlation coefficients between the sensitivity at 60 dB SPL and the across-listener average in predicted localization performance yielded $-.92$ ($p < .001$) for quadrant errors and $-.90$ ($p < .001$) for local errors and thus confirmed a strong relationship between the fiber's sensitivity in level discrimination and the predicted average localization performance.

Discussion

To study the effect of sensorineural hearing loss on median-plane sound localization in quiet, we integrated the auditory periphery model from Zilany et al. (2009, 2014) into the sagittal-plane sound localization model from Baumgartner et al. (2014). The model evaluation performed for normal-hearing listeners showed a good predictive power of the model for spectral modifications of HRTFs. Applied on simulations of OHC dysfunctions, localization accuracy was found to be relatively robust with respect to moderate OHC dysfunction, but severe OHC dysfunction drastically degraded the performance. Thus, the model predicted localization performance in accordance with the estimated hearing loss categories, that is, good performance for normal hearing (normal OHC function and moderate OHC dysfunction), degraded performance for mild hearing loss (severe OHC dysfunction), and chance performance for moderate hearing loss (complete OHC loss).

The predicted localization performance resembles the usually observed degradation in listener-specific localization performance induced by sensorineural hearing loss or comparable signal modifications. In particular, the predicted performance for moderate hearing loss being close to chance performance is consistent with several previous investigations (Dobreva et al., 2011; Noble et al., 1994; Otte et al., 2013; Rakerd et al., 1998). Macpherson and Sabin (2013) employed spectral contrast reduction, which simulated comparable signal degradations in normal-hearing listeners. For a spectral contrast factor reduced from 100% to 25%, their listeners' proportion of quasi-veridical responses (localization error $\leqslant 45°$) decreased from around 90% to 70%. According to their definition of spectral contrast, our low-SR condition with normal OHC function reduced the spectral contrast to about 20% and caused the predicted percentage of quadrant errors (localization error $> 90°$) to increase from 10% (baseline performance for all SRs) to around 30%, that is, 70% of responses were localized in the correct hemisphere (localization error $\leqslant 90°$). Hence, our predictions for the low-SR condition seem to be very consistent with the experimental results of Macpherson and Sabin (2013). In our medium-SR condition with normal OHC function, spectral contrast increased to about 130%, and the model predicted significantly improved localization performance. For comparable contrast expansions, Macpherson and Sabin found only a modest improvement. One reason for their modest improvement might be that their listeners performed better than ours in general, so that ceiling effects left no room for further improvement. Another reason might be related to the fact that we simulated perfect adaptation to changes in the auditory periphery, whereas in the experiments of Macpherson and Sabin, listeners
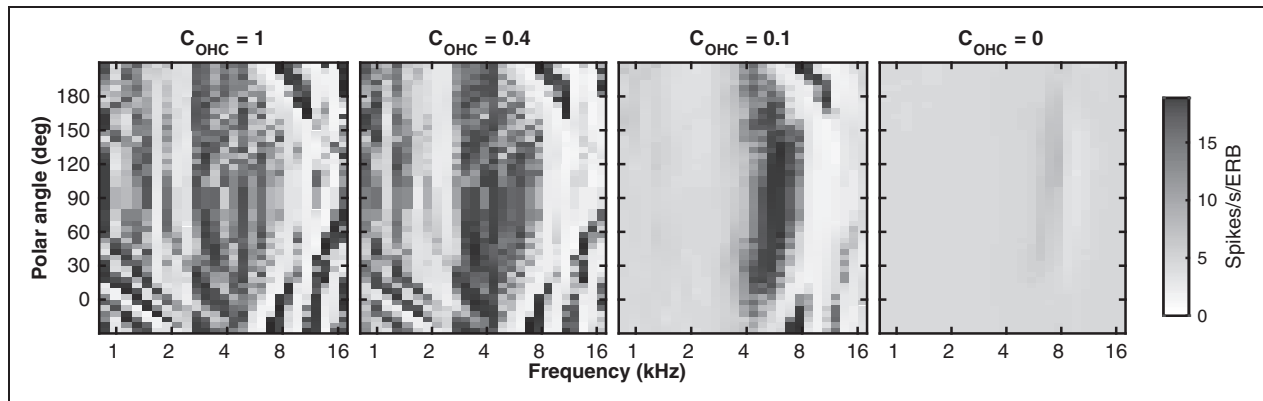
Figure 9. Effect of OHC dysfunction on positive spectral gradients. Exemplary median-plane HRTFs from one listener (NH46). Note the distinct direction-specific patterns for the normal and moderate OHC dysfunctions ($C_{OHC} \geqslant 0.4$), which are almost absent in the cases of the severe and complete OHC dysfunctions ($C_{OHC} \leqslant 0.1$).

were confronted with ad hoc spectral modifications. Contrast expansion in particular might have been less beneficial if it had not been represented in the internal spectral templates as well.

We thus further analyzed the general effect of OHC dysfunctions on the spectral cues represented in target sounds (and internal templates). Figure 9 shows the positive spectral gradient profiles as functions of the polar angle for all tested OHC gains. In line with localization performance being relatively robust to moderate OHC dysfunction, the profiles for normal ($C_{OHC} = 1$) and moderately reduced ($C_{OHC} = 0.4$) gain both appear very similar and reveal prominent direction-specific patterns. In contrast, there are smaller gradients with less directionality in the case of severe dysfunction ($C_{OHC} = 0.1$), and for the complete OHC damage ($C_{OHC} = 0$), almost no gradients can be identified. It should be noticed, however, that our investigations focused on listening conditions without any background noise. The presence of background noise might render proper OHC functionality even more important (May et al., 2004).

Predictions of localization performance based on a selective use of the different AN fiber types suggest that medium-SR fibers, although comprising a small percentage of AN fibers, are pivotal for good localization performance around the SPL of 60 dB. Interestingly, the exclusive activity of medium-SR fibers yielded even better performance than the combined activity of all fiber types (representing the actual performance of normal-hearing listeners with intact OHC function). It can thus be expected that localization performance would improve if the auditory system would be able to focus on the best coding fibers and ignore worse coding fibers, but to our current knowledge, there is no evidence for selective processing of different fiber types in the cochlear nucleus.

Our predictions of level discrimination show that in case of normal OHC function and for levels exceeding 70 dB SPL, the sensitivity was higher for low-SR fibers than for high-SR fibers. This is in agreement with results from Reiss et al. (2011), who recorded responses from the AN across a large range of SPLs and signal-to-noise ratios in the domestic cat and concluded that low-SR fibers provide more beneficial coding properties than high-SR fibers at higher SPLs and lower signal-to-noise ratios. We found a strong correlation between sensitivity in level discrimination based on rate differences and average localization performance predicted for the SPL of 60 dB. Consequently, level discriminability might be a good indicator for average localization performance also at other SPLs. Since the sensitivity of the low-SR fibers exceeded that of the high-SR fibers above 70 dB SPL and that of medium-SR fibers at 90 dB SPL, our simulations suggest that low-SR fibers are particularly important for accurate localization in the sagittal planes at higher SPLs. We, therefore, expect a marked deficit in localization performance at higher levels in listeners with noise-induced hearing loss which predominantly affects medium- and low-SR fibers (Furman et al., 2013).

Note that our model predictions are based on a specific set of three stereotypes of AN fibers. Since the auditory-periphery model we used has not been directly fitted to AN responses of human listeners, although humanization of a cat model was indirectly applied, some uncertainty about the spectral representations of incoming sounds in the AN remains. Hence, our model results for specific fiber types should be interpreted relatively rather than absolutely. Consequently, the prediction of optimal performance for medium-SR fibers might also indicate that lower-SR fibers are generally more important for localization than higher-SR fibers, particularly at higher levels.

With OHC dysfunction all AN fiber types became more sensitive in level discrimination at higher SPLs. This could compensate to some degree for the reduced discriminability at higher SPLs caused by a loss of lower-SR fibers if the sounds are not amplified by hearing-assistive devices. For lower-SPL stimuli, however, amplification improves localization performance (Rakerd et al., 1998), as it shifts the signal into a level region with better discriminability.

In the present study, we focused only on the OHC dysfunction and the contribution of different AN fiber types. The proposed model, however, can also serve as a framework for future investigations including the effect of IHC dysfunctions and gradual changes in the prevalence of AN fiber types. For the sake of reproducibility and accessibility, we incorporated the model baumgartner 2016 as well as the modeled experiments exp_baumgartner2016 in the Auditory Modeling Toolbox (http://www.amtoolbox.sf.net; Søndergaard & Majdak, 2013).

References

Algazi, V. R., Avendano, C., & Duda, R. O. (2001). Elevation localization and head-related transfer function analysis at low frequencies. *The Journal of the Acoustical Society of America*, *109*(3), 1110–1122. doi:10.1121/1.1349185

Andéol, G., Guillaume, A., Micheyl, C., Savel, S., Pellieux, L., Moulin, A. (2011). Auditory efferents facilitate sound localization in noise in humans. *The Journal of Neuroscience*, *31*(18), 6759–6763. doi:10.1523/JNEUROSCI.0248-11.2011

ANSI. (2010). *Specification for audiometers* (Standard No. ANSI/ASA S3.6-2010) (pp. 1–68). Melville, NY: Accredited Standards Committee S3, Bioacoustics

Baumgartner, R., Majdak, P., & Laback, B. (2013). Assessment of sagittal-plane sound localization performance in spatial-audio applications. In J. Blauert (Ed.), *The technology of binaural listening* (pp. 93–119). Berlin, Heidelberg: Springer.

Baumgartner, R., Majdak, P., & Laback, B. (2014). Modeling sound-source localization in sagittal planes for human listeners. *The Journal of the Acoustical Society of America*, *136*(2), 791–802. doi:10.1121/1.4887447

Best, V., Carlile, S., Jin, C., & van Schaik, A. (2005). The role of high frequencies in speech localization. *The Journal of the Acoustical Society of America*, *118*(1), 353–363. doi:10.1121/1.1926107

Bruce, I. C., Sachs, M. B., & Young, E. D. (2003). An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. *The Journal of the Acoustical Society of America*, *113*(1), 369–388. doi:10.1121/1.1519544

Dobreva, M. S., O'Neill, W. E., & Paige, G. D. (2011). Influence of aging on human sound localization. *Journal of Neurophysiology*, *105*(5), 2471–2486. doi:10.1152/jn.00951.2010

Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of Neurophysiology*, *110*(3), 577–586. doi:10.1152/jn.00164.2013

Goodman, A. (1965). Reference zero levels for pure-tone audiometer. *ASHA*, *7*, 262–263.

Goupell, M. J., Majdak, P., & Laback, B. (2010). Median-plane sound localization as a function of the number of spectral channels using a channel vocoder. *The Journal of the Acoustical Society of America*, *127*(2), 990–1001. doi:10.1121/1.3283014

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (1st ed.). New York, NY: Wiley & Sons, Inc.

Ibrahim, R. A., & Bruce, I. C. (2010). Effects of peripheral tuning on the auditory nerve's representation of speech envelope and temporal fine structure cues. In E. A. Lopez-Poveda, A. R. Palmer, & R. Meddis (Eds.), *The neurophysiological bases of auditory perception* (pp. 429–438). New York, NY: Springer.

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *The Journal of the Acoustical Society of America*, *63*(2), 442–455. doi:10.1121/1.381736

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*, *111*(5), 2219–2236. doi:10.1121/1.1471898

Macpherson, E. A., & Sabin, A. T. (2013). Vertical-plane sound localization with distorted spectral cues. *Hearing Research*, *306*, 76–92. doi:10.1016/j.heares.2013.09.007

Majdak, P., Baumgartner, R., & Laback, B. (2014). Acoustic and non-acoustic factors in modeling listener-specific performance of sagittal-plane sound localization. *Frontiers in Psychology*, *5*(319), 1–10. doi:10.3389/fpsyg.2014.00319

Majdak, P., Goupell, M. J., & Laback, B. (2010). 3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training. *Attention, Perception, & Psychophysics*, *72*(2), 454–469. doi:10.3758/APP.72.2.454

Majdak, P., Masiero, B., & Fels, J. (2013). Sound localization in individualized and non-individualized crosstalk cancellation systems. *The Journal of the Acoustical Society of America*, *133*(4), 2055–2068. doi:10.1121/1.4792355

Majdak, P., Walder, T., & Laback, B. (2013). Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer

functions. *The Journal of the Acoustical Society of America*, *134*(3), 2148–2159. doi:10.1121/1.4816543

Mamun, N., Jassim, W. A., & Zilany, M. S. A. (2015). Prediction of speech intelligibility using a neurogram orthogonal polynomial measure (NOPM). *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *23*(4), 760–773. doi:10.1109/TASLP.2015.2401513

Mao, J., & Carney, L. H. (2014). Tone-in-noise detection using envelope cues: Comparison of signal-processing-based and physiological models. *Journal of the Association for Research in Otolaryngology*, *16*(1), 121–133. doi:10.1007/s10162-014-0489-1

May, B. J., Budelis, J., & Niparko, J. K. (2004). Behavioral studies of the olivocochlear efferent system: Learning to listen in noise. *Archives of Otolaryngology–Head & Neck Surgery*, *130*(5), 660–664. doi:10.1001/archotol.130.5.660

May, B. J., & Huang, A. Y. (1997). Spectral cues for sound localization in cats: A model for discharge rate representations in the auditory nerve. *The Journal of the Acoustical Society of America*, *101*(5), 2705–2719. doi:10.1121/1.418559

Moore, B. C. J. (1995). *Perceptual consequences of cochlear damage* (vol. 28). Oxford, England: Oxford University Press.

Morimoto, M., & Aokata, H. (1984). Localization cues in the upper hemisphere. *The Journal of the Acoustical Society of Japan (E)*, *5*, 165–173.

Noble, W., Byrne, D., & Lepage, B. (1994). Effects on sound localization of configuration and type of hearing impairment. *The Journal of the Acoustical Society of America*, *95*(2), 992–1005. doi:10.1121/1.408404

Otte, R. J., Agterberg, M. J. H., Wanrooij, M. M. V., Snik, A. F. M., & Opstal, A. J. V. (2013). Age-related hearing loss and ear morphology affect vertical but not horizontal sound-localization performance. *Journal of the Association*

for Research in Otolaryngology*, *14*(2), 261–273. doi:10.1007/s10162-012-0367-7

Rakerd, B., Vander Velde, T. J., & Hartmann, W. M. (1998). Sound localization in the median sagittal plane by listeners with presbyacusis. *Journal of the American Academy of Audiology*, *9*(6), 466–479.

Reiss, L. A. J., Ramachandran, R., & May, B. J. (2011). Effects of signal level and background noise on spectral representations in the auditory nerve of the domestic cat. *Journal of the Association for Research in Otolaryngology*, *12*(1), 71–88. doi:10.1007/s10162-010-0232-5

Reiss, L. A. J., & Young, E. D. (2005). Spectral edge sensitivity in neural circuits of the dorsal cochlear nucleus. *Journal of Neuroscience*, *25*(14), 3680–3691. doi:10.1523/JNEUROSCI.4963-04.2005

Sabin, A. T., Macpherson, E. A., & Middlebrooks, J. C. (2005). Human sound localization at near-threshold levels. *Hearing Research*, *199*(1–2), 124–134. doi:10.1016/j.heares.2004.08.001

Søndergaard, P., & Majdak, P. (2013). The auditory modeling toolbox. In J. Blauert (Ed.), *The technology of binaural listening* (pp. 33–56). Berlin, Germany: Springer. doi:10.1007/978-3-642-37762-4_2

Wightman, F. L., & Kistler, D. J. (1997). Monaural sound localization revisited. *The Journal of the Acoustical Society of America*, *101*(2), 1050–1063. doi:10.1121/1.418029

Zilany, M. S. A., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *The Journal of the Acoustical Society of America*, *135*(1), 283–286. doi:10.1121/1.4837815

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., & Carney, L. H. (2009). A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics. *The Journal of the Acoustical Society of America*, *126*(5), 2390–2412. doi:10.1121/1.3238250