Binaural-cue Weighting and Training-Induced Reweighting Across Frequencies

Maike Klingel[1,2] (iD) and Bernhard Laback[2] (iD)

Abstract

During sound lateralization, the information provided by interaural differences in time (ITD) and level (ILD) is weighted, with ITDs and ILDs dominating for low and high frequencies, respectively. For mid frequencies, the weighting between these binaural cues can be changed via training. The present study investigated whether binaural-cue weights change gradually with increasing frequency region, whether they can be changed in various frequency regions, and whether such binaural-cue reweighting generalizes to untrained frequencies. In two experiments, a total of 39 participants lateralized 500-ms, 1/3-octave-wide noise bursts containing various ITD/ILD combinations in a virtual audio-visual environment. Binaural-cue weights were measured before and after a 2-session training in which, depending on the group, either ITDs or ILDs were visually reinforced. In experiment 1, four frequency bands (centered at 1000, 1587, 2520, and 4000 Hz) and a multiband stimulus comprising all four bands were presented during weight measurements. During training, only the 1000-, 2520-, and 4000-Hz bands were presented. In experiment 2, the weight measurements only included the two mid-frequency bands, while the training only included the 1587-Hz band. ILD weights increased gradually from low- to high-frequency bands. When ILDs were reinforced during training, they increased for the 4000- (experiment 1) and 2520-Hz band (experiment 2). When ITDs were reinforced, ITD weights increased only for the 1587-Hz band (at specific azimuths). This suggests that ILD reweighting requires high, and ITD reweighting requires low frequencies without including frequency regions providing fine-structure ITD cues. The changes in binaural-cue weights were independent of the trained bands, suggesting some generalization of binaural-cue reweighting.

Keywords

spatial hearing, interaural time difference, interaural level difference, plasticity, adaptation

Introduction

Spatial hearing is an important ability of the auditory system, as it allows the localization of sound sources and improves speech understanding in complex environments. This ability relies on the integration of information provided by different auditory cues (for a recent review on sound localization cues, see Stecker & Gallun, 2012). For azimuthal sound localization (i.e., in the horizontal dimension, referred to as lateralization for in-head localization judgements), which is the focus of this study, humans rely primarily on two binaural cues, namely interaural differences in time (ITD) and level (ILD).

Psychophysical and physiological studies (e.g., Grothe et al., 2010; Henning, 1980) have revealed that ITD cues are available in the temporal fine structure at low carrier frequencies (up to approx. 1.4 kHz) and in the temporal envelope at higher carrier frequencies. ILD cues, on the other hand, are much more pronounced at high frequencies, due

to the frequency dependence of the head shadow (e.g., Middlebrooks & Green, 1991). Consequently, there are disparate cues available at low (mainly ITDs) and high (mainly ILDs) frequencies, a phenomenon known as the duplex theory of sound localization (Strutt, 1907). To establish the perceived azimuth of a sound source, the auditory system weights the information provided by the binaural cues. Given that the information carried by ITD and ILD cues varies with frequency, it is an important question to

[1]Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Wien, Austria
[2]Acoustics Research Institute, Austrian Academy of Sciences, Wien, Austria

Corresponding author:
Maike Klingel, Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Liebiggasse 5, Wien, 1010 Vienna, Austria.
Email: maike.klingel@univie.ac.at

what extent these binaural-cue weights depend on the sound's frequency content. In fact, studies have shown that ITDs dominate for broadband sounds and at low frequencies, while ILDs dominate at higher frequencies (Ahrens et al., 2020; Macpherson & Middlebrooks, 2002; Wightman & Kistler, 1992). It is unclear, however, whether this binaural-cue weighting gradually changes with increasing frequency or whether it abruptly switches from being ITD dominant to being ILD dominant. Macpherson and Middlebrooks (2002) investigated the effect of delaying or attenuating the signal at one ear (while keeping the other localization cues intact) on localization of low-pass (0.5–2 kHz), high-pass (4–16 kHz), or wideband (0.5–16 kHz) stimuli, but did not test intermediate frequency bands. Ahrens et al. (2020) presented multiband stimuli consisting of up to 11 frequency bands which varied independently in either ITD or ILD while fixing the other cue at zero. They observed divergence for the lowest and highest frequency bands (i.e., stronger ITD weights for the lowest and stronger ILD weights for the highest band) but constant weights for in-between bands. This simultaneous presentation may, however, not show the weighting of each frequency band individually and instead be affected by simultaneous grouping effects, particularly binaural interference (Ahrens et al., 2020). The diverging weights for the lowest and highest bands appear to reflect an "edge effect" (i.e., stronger weighting of information at the edges), because the same pattern was observed in different edge bands when only intermediate frequency bands were presented ("removed" condition, Ahrens et al., 2020). Therefore, we were interested in whether binaural-cue weights measured similarly to Macpherson and Middlebrooks (2002) methods (i.e., by presenting each frequency band in isolation) show a comparable pattern to Ahrens et al. (2020) or whether they gradually change with increasing frequency. Such a gradual change would be more in line with the information carried by the binaural cues as discussed above as well as with previous literature on the basic sensitivity to ITD and ILD cues, assuming higher sensitivity (i.e., a lower detection threshold) correlates with stronger weighting of the respective cue. ILD thresholds tend to gradually improve between 1000 and 4000 Hz with increasing center frequency of narrow-band noise stimuli (Gabriel et al., 1992; Goupell & Stakhovskaya, 2018). In contrast, ITD thresholds worsen in a similar frequency range with increasing center frequency (Buchholz et al., 2018; Gabriel et al., 1992; Klumpp & Eady, 1956; Smoski & Trahiotis, 1986; Trahiotis & Bernstein, 1990). Note that the noise bands provide temporal fine-structure ITDs at lower frequencies and temporal envelope ITDs at higher frequencies. Therefore, if binaural cue weights are related to binaural cue thresholds, one would expect a gradual change in binaural-cue weighting with increasing frequency.

In addition to frequency, the binaural-cue weighting is influenced by other stimulus properties such as the overall intensity (David et al., 1959; Deatherage & Hirsh, 1959), the inter-click interval of click trains (Stecker, 2010), or the presence of reverberation (Rakerd & Hartmann, 2010) and shows substantial variation across listeners (Klingel et al., 2021; Klingel et al., 2020; Macpherson & Middlebrooks, 2002). Such a dependence on stimulus, environmental, and personal properties is not surprising, given that listeners adapt their processing of sound localization cues when exposed to cue alterations (see Carlile, 2014, for a review). Such adaptation can either be mediated by remapping (i.e., learning a new relationship between sound localization cues and corresponding locations in space; e.g., Shinn-Cunningham et al., 1998) or reweighting (i.e., a stronger relative weighting of unaltered or reliable cues compared to altered or unreliable cues). Such reweighting has been shown for monaural (spectral-shape) cues (resulting from the directional filtering of the outer ears) relative to binaural cues for horizontal sound localization (Keating et al., 2013; Kumpik et al., 2010; Van Wanrooij & van Opstal, 2007). Usually, monaural cues do not contribute to horizontal sound localization if binaural cues are available (Macpherson & Middlebrooks, 2002; Slattery & Middlebrooks, 1994).

Additionally, three recent studies have demonstrated reweighting of binaural cues. Kumpik et al. (2019) observed an increase in ILD weighting after the ITDs of broadband noise stimuli were randomized during the completion of a visual task, but no corresponding increase in ITD weighting after ILDs were randomized. Klingel et al. (2020) also observed an increase in ILD weighting after participants received response feedback consistent with the ILDs of a mid-frequency (2–4 kHz) narrow-band noise stimulus in an auditory discrimination task, but did not test a group receiving feedback consistent with the ITDs. Finally, Klingel et al. (2021) observed a comparably strong increase in either the ITD or the ILD weighting of the same 2–4-kHz noise stimulus, depending on which cue was visually reinforced during a lateralization task in a virtual audio-visual environment.

In summary, there is evidence that listeners can adjust their weighting of sound localization cues, and particularly their binaural-cue weighting. It is, however, still unclear whether such binaural-cue reweighting can be achieved for low- or high-frequency stimuli, for which either ITDs or ILDs are known to dominate. If the baseline weight of the reinforced cue is already strong, there might not be room for a further increase (i.e., ceiling effects). On the other hand, if the baseline weight of the reinforced cue is low, the dominance of the other cue might limit the access to the reinforced cue, thus preventing binaural-cue reweighting. Finally, we were interested in whether such reweighting is specific to the trained frequency band or whether it generalizes to untrained stimuli. Wright and Fitzgerald (2001), for example, observed different generalization patterns for ITD and ILD sensitivity training.

By conducting two experiments, the present study investigates the pattern of baseline binaural-cue weights across

different frequency regions (i.e., whether weights gradually or abruptly change from ITD- to ILD-dominant with increasing frequency), whether the weighting can be changed via visual reinforcement for different frequencies, and whether such binaural-cue reweighting generalizes to untrained stimuli. Experiment 1 addressed all of these questions but some questions about the potential roles of the stimulus' azimuthal range, the presentation mode of frequency bands (randomized vs. blocked), and the frequency range required for reweighting remained. These questions were addressed in experiment 2.

Experiment 1

The aim of this experiment was to test whether binaural-cue weights, measured similarly to Macpherson and Middlebrooks (2002), will show gradual weight changes across different frequency regions or whether they will show divergence for the lowest and highest frequency bands but constant weights in between, similar to the results reported by Ahrens et al. (2020). We further investigated whether binaural-cue reweighting can be induced for different frequency regions using the paradigm of Klingel et al. (2021; see description of methods below). Finally, we were interested in whether binaural-cue reweighting is specific to the trained frequency band or whether it generalizes to untrained frequency bands or to a broad-band stimulus.

Methods

Participants. Nineteen normal-hearing participants took part in experiment 1. All participants gave written informed consent before starting the experiment and received monetary compensation for their participation. The research protocol was approved by the Acoustics Research Institute's ethics committee. Participants were randomly assigned to two experimental groups: For the ITD group, ITDs were reinforced during training and for the ILD group, ILDs were reinforced during training. Ten participants (four female, mean age: 27.10 years, $SD = 4.33$, range 21–33) were assigned to the ITD group. The remaining nine participants (five female, mean age: 29.67 years, $SD = 7.25$, range 22–48 years) were assigned to the ILD group.

Apparatus and Stimuli. Participants stood on a platform surrounded by a circular railing inside a sound booth. Visual stimuli were presented binocularly via a head-mounted display (Oculus Rift, CV1). Participants' head position and orientation were tracked using the Oculus device and software, and the visual environment was rendered accordingly in real-time. The virtual visual environment was programmed in Unity (Unity Technologies, San Francisco, CA), connected to the Oculus device via the Oculus Integration tool[1], and communicated with our experimental software via custom scripts programmed in C#[2]. It consisted of a gray sphere with a reference position (red dot) straight

ahead and vertical stings of blue dots every 45° in azimuth for guidance. Additionally, a crosshair was displayed showing the azimuth of the head orientation, which was locked at 0° elevation to force participants to respond in the horizontal plane (while the rest of the visual environment followed vertical head rotations to encourage immersion).

Binaural auditory stimuli were generated using a computer and output via a digital audio interface (ADI-8, RME) at a 48-kHz sampling rate and presented via headphones (HD 580, Sennheiser). They were band-pass filtered white noise bursts (1/3 octave wide), randomly generated on each trial. Four frequency bands were used: A low-frequency band centered at 1000 Hz (low, 793.7–1259.9 Hz), a mid-low-frequency band centered at 1587.4 Hz (mid-low, 1414.2–1781.8 Hz), a mid-high-frequency band centered at 2519.8 Hz (mid-high, 2244.9–2828.4 Hz), and a high-frequency band centered at 4000 Hz (high, 3563.6–4489.8 Hz). These stimuli were chosen to have equally spaced frequency bands while minimizing physical and perceptual overlap between bands in a similar frequency range to Ahrens et al.'s (2020) "removed" condition. We did not test the full range presented in Ahrens et al. (2020) to keep the experimental time reasonable, since we presented the frequency bands sequentially. Additionally, a stimulus comprising all four bands (multiband) was used. The stimulus duration was 500 ms, including 50-ms raised-cosine on/off ramps. The mean overall sound pressure level (SPL) was 63 dB for an ILD of zero. For the multiband condition, the individual bands were equalized in level. To discourage participants from using differences in the absolute level rather than ILDs for lateralization, the overall level was roved randomly from trial to trial within a $\pm$ 2.5 dB range. The stimuli were not filtered with head related transfer functions (HRTFs). This was to ensure that they did not convey monaural spectral localization cues that are potentially informative about the stimulus azimuth (e.g., Hebrank & Wright, 1974), which might confound the binaural-cue weight estimation (although the narrow bandwidth and frequency range of the stimuli would have made this unlikely). Instead, ITDs ranging from −396 to + 396 μs and ILDs ranging from −14.95 to + 14.95 dB were imposed on these source stimuli. ITDs that matched azimuths ranging from −45° to + 45° with a 6° spacing were determined based on Xie's (2013) estimation using the HRTFs of the KEMAR head with DB-61 small pinna at a source distance of 1.4 m. In Xie (2013), ITD values were obtained via broadband cross-correlation of the left and right ear head-related impulse responses (HRIRs). ILDs matching the same azimuths were determined individually for each frequency band based on the mean HRTF magnitudes at the respective center frequencies of four participants that did not take part in this study (taken from the HRTF database[3] of the Acoustics Research Institute). For the multiband stimulus, the ILDs determined for each frequency band were used for the respective part of the stimulus. The azimuth range of $\pm$ 45° was chosen to ensure monotonically increasing ILDs with increasing azimuths

for all frequency bands. Figure 1 shows the dependence of ITDs (panel a) and ILDs (panel b) on azimuth as determined according to the described methods, with the used binaural cues marked by symbols up to the dashed black line (symbols beyond that line refer to experiment 2).

ITDs and ILDs were combined into "consistent-cue" and "inconsistent-cue" conditions. In consistent-cue conditions, the ITD and ILD cue of the auditory stimulus corresponded to the same azimuth, while they corresponded to disparate azimuths in inconsistent-cue conditions. Cue disparities (i.e., the difference between the azimuths corresponding to the binaural cues within each stimulus) were restricted to a maximum value of 24° to avoid the perception of split images which can occur in case of large cue disparities (Gaik, 1993). During training, either the ITD or the ILD cue was visually reinforced as described below. Reinforced-cue azimuths ranged from −21° to +21° (with 6° spacing) and azimuths corresponding to the unreinforced cue were uniformly distributed ± 24° (also with 6° spacing) around each reinforced-cue azimuth (Figure 1d in yellow). By symmetrically varying the unreinforced-cue azimuth around each reinforced-cue azimuth [resulting in a larger range of unreinforced-cue azimuths (±45°) than reinforced-cue azimuths (±21°)], the reinforced cue was more stable, which might encourage reweighting in addition to the visual reinforcement (Dahmen et al., 2010). In the pre- and posttest in which neither of the two cues was reinforced, azimuths in *both* the ITD and ILD dimension were uniformly distributed ± 24° around each azimuth from −21° to +21° (Figure 1c in yellow).

Procedure. The same task as in Klingel *et al.* (2021) was used. It involved lateralizing auditory stimuli in a virtual environment via head-turn. By asking participants to lateralize stimuli containing spatially inconsistent binaural cues and comparing the response azimuth to the azimuths of each binaural cue, it can be inferred how much each cue contributes to the azimuthal percept, thus, providing the binaural-cue weights (Macpherson & Middlebrooks, 2002; Stecker, 2010). The time course of tasks is shown in Table 1. The experiment started with a practice session in which participants familiarized themselves with the task, a pretest to measure the initial ITD and ILD weights, a training in which one of the binaural cues was visually reinforced as described below, and a posttest to remeasure the weights after the training. The experimental protocol was split into two sessions (either completed on consecutive days or with a one-day break in between). This was done to avoid fatigue and to potentially make use of consolidation overnight. In the first session, participants completed the practice session, the pretest, and the first training session. In the second session, participants completed the second training session and the posttest. The approximate duration was 2h for the first and 1.5h for the second session.

Practice. The practice session was performed at the beginning of the experiment and involved same task as the

training (see below). It consisted of 60 trials with spatially consistent ITD/ILD-combinations in which each azimuth from −33° to +33° (with 6° spacing) was presented once for each frequency-band condition (low, mid-low, mid-high, high, and multiband). The trials were presented in random order and, as in the training, visual reinforcement was provided.

Testing. The pretest and the posttest procedures were identical for the two groups and did not include visual reinforcement. On each trial, participants initiated the sound presentation by facing the reference position (straight ahead) and pressing a button. If they moved their head away from the reference position before the sound presentation finished, a warning tone was played and participants had to repeat the sound presentation. They then indicated the perceived azimuth via head turn and button press. The head orientation at the button press was recorded as the response azimuth. Participants then returned to the reference position to initiate the next trial. A total of 460 trials were presented, consisting of each ITD/ILD-combination shown in Figure 1c in yellow presented once for each frequency-band condition (low, mid-low, mid-high, high, and multiband). The trials were presented in random order and after each 115 trials, participants took a short break.

Training. The training procedure consisted of 6 steps (shown in Figure 2): 1) Initiating the sound presentation and listening to the auditory stimulus while facing the reference position, 2) indicating the perceived azimuth via head turn and button press, 3) receiving visual reinforcement (a rotating cube) at the reinforced-cue azimuth, 4) finding and confirming the reinforced-cue azimuth via head turn and button press, 5) returning to the reference position and initiating the presentation of the same auditory stimulus again while the visual reinforcement is still visible, and 6) confirming the reinforced-cue azimuth again via head turn and button press after which the visual reinforcement disappeared. In steps 1), 4), 5), and 6), the button-press was accepted only if the head-orientation was within ± 5° of the reference position or the reinforced-cue azimuth, respectively. If participants moved their head away from the reference position before the sound presentation finished, a warning tone was played, and participants had to repeat the sound presentation.

Auditory stimuli included both inconsistent and consistent ITD/ILD-combinations, as shown in Figure 1d in yellow. The training procedure was the same for the two groups except for which cue was visually reinforced and presented in a limited azimuth range. Thus, for the ITD group, ITD azimuths did not exceed ± 21° and for the ILD group, ILD azimuths did not exceed ± 21°. Each training session consisted of 432 trials presented in two blocks of 216 trials each. Within each block, trials were presented in a random order and each ITD/ILD combination shown in Figure 1d in yellow was presented once for the low, mid-high, and high bands. We chose the low and high bands for training to test for reweighting potential in frequency regions where one of
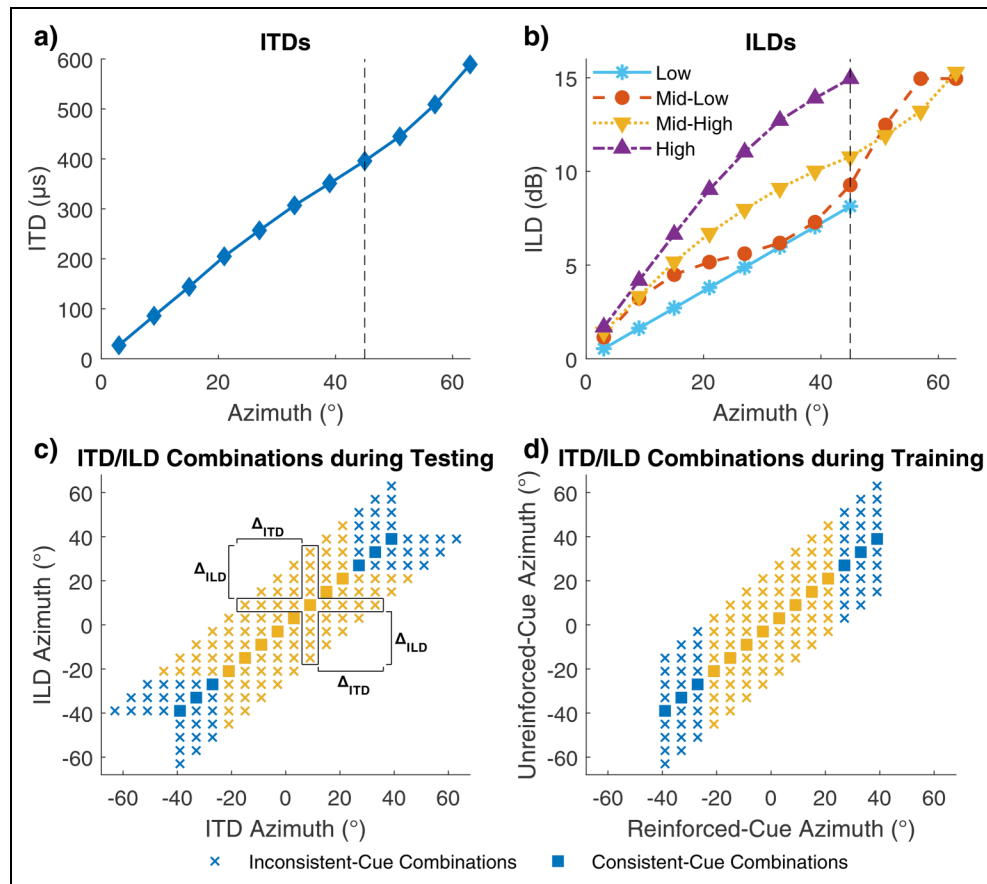
Figure 1. Experimental setup and stimuli. Panels a) and b) show the functional relation between the azimuth and the binaural cues. ITDs are derived by Xie (2013) based on broadband cross-correlation of the left and right ear head-related impulse responses (HRIRs) of the KEMAR head. ILDs are based on the mean HRTF magnitudes at the respective center frequency of four new participants. In experiment 1, binaural-cues only up to the black dashed line were used. Panel c) shows all the ITD/ILD-azimuth combinations in the pre- and posttest. In experiment 1, only yellow combinations were used. In experiment 2, both yellow and blue combinations were included. The frame indicates the azimuthal offsets Δ_{ITD} and Δ_{ILD} that were used to estimate the model parameters for the pre-/posttest data at one example azimuth (9°). Panel d) shows all cue combinations used in the training. For the ITD group, reinforced and unreinforced cues were ITD and ILD, respectively, and for the ILD group, reinforced and unreinforced cues were ILD and ITD, respectively.

Table 1. Time Course (top to Bottom) for Each Experiment. The Tested Frequency Bands are Shown in Parentheses. Tasks Involving Visual Reinforcement are shown in italic.

	Experiment 1	Experiment 2
Session 1	*Practice* (all bands)	*Practice* (mid-low)
		Pretest (mid-low)
	Pretest (all bands)	*Practice* (mid-high)
		Pretest (mid-high)
	Training (low, mid-high, high)	*Training* (mid-low)
Session 2	*Training* (low, mid-high, high)	*Training* (mid-low)
	Posttest (all bands)	**Posttest** (mid-low)
		Posttest (mid-high)

the cues clearly dominates. The mid-high band was chosen because it was close in center frequency to the stimuli used in Klingel *et al.* (2021), for which both ITD and ILD

reweighting was shown. Thus, the mid-low and the multi-band conditions were not presented during training to test whether the training effects generalize to them. After every 72 trials, participants took a short break.

Analysis. Analogous to Klingel et al. (2021), we estimated binaural-cue weights for each participant based on a regression analysis, which measured the effect of varying the ILD on the responses at each ITD azimuth and vice versa. The resulting regression coefficients were then combined to yield the normalized ILD weight (see below). A resulting ILD weight of 1 means that participants always responded at the ILD azimuth, an ILD weight of 0 means participants always responded at the ITD azimuth, and an ILD weight of 0.5 represents equal weighting of the two cues. Note that due to this normalization, the ITD weight can be directly inferred from the ILD weight (ITD weight = 1 – ILD weight), which is why we report only ILD weight values. The regression analysis was
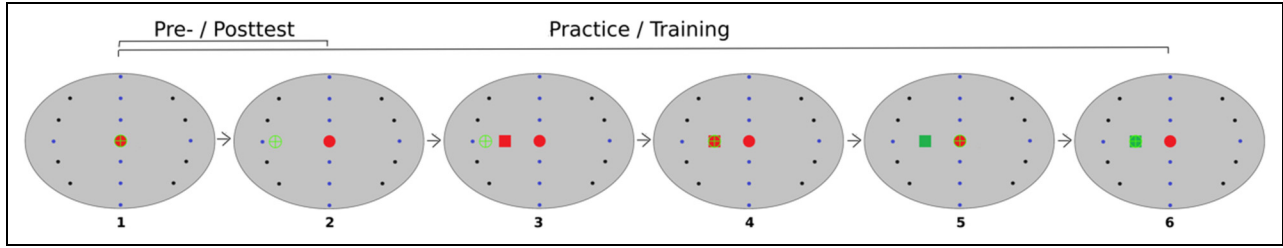
Figure 2. Time course of a trial during the pre- and posttest (panels 1-2) and the practice session as well as the training (panels 1-6). 1) Participants oriented towards the reference position (indicated by a red sphere) and pressed a button to elicit the sound presentation. 2) Participants turned their head (guiding a green crosshair) to the perceived azimuth and pressed the button (in this example, they turned their head to the left). 3) Visual reinforcement (a rotating red cube) appeared at the reinforced-cue azimuth. 4) The reinforced-cue azimuth was confirmed via a head-turn to the visual reinforcement and a button-press. 5) The visual reinforcement turned green, participants returned to the reference position, and elicited the second sound presentation (while the visual reinforcement was still visible) with another button-press. 6) Participants confirmed the reinforced-cue azimuth again via another head-turn and button-press.
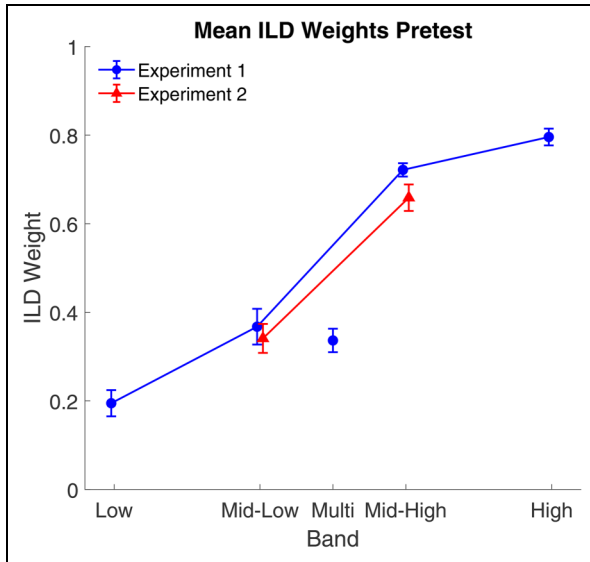

Figure 3. Pretest ILD weights (normalized, i.e., ILD weight = 1 − ITD weight) for each frequency band, averaged across azimuths. Blue circles show the results of experiment 1 and red triangles show the results of experiment 2. The averaged azimuths were restricted to the range that was tested in both experiments (3°-21°). Error bars show the standard error of the mean. ILD weights gradually and significantly increase for single-band conditions from low (794-1260 Hz), to mid-low (1414-1782 Hz), to mid-high (2245-2828 Hz), to high (3564-4490 Hz). These conditions are connected by a line. The multiband condition (multi, plotted as a separate data point) shows ILD weights similar to the mid-low condition.

fitted separately for each azimuth α (between 3° and 21° with a 6° spacing between azimuths) after mirroring the data across the midline. The regression model equations are as follows:

$$R_{ITD}(\alpha, \Delta_{ITD}) = k_{ITD}(\alpha) * \Delta_{ITD} + Q_{ITD}(\alpha)$$

$$R_{ILD}(\alpha, \Delta_{ILD}) = k_{ILD}(\alpha) * \Delta_{ILD} + Q_{ILD}(\alpha)$$

$$w_{ILD}(\alpha) = \frac{atan\left(\frac{k_{ILD}(\alpha)}{k_{ITD}(\alpha)}\right)}{\frac{\pi}{2}}$$

$$Q(\alpha) = \frac{Q_{ILD}(\alpha) + Q_{ITD}(\alpha)}{2} \qquad \text{(Eq.1)}$$

where R_{ITD} (R_{ILD}) is the participant's mean response azimuth in a trial for which the ILD (ITD) corresponded to azimuth α and the ITD (ILD) corresponded to azimuth $\alpha + \Delta_{ITD}$ ($\alpha + \Delta_{ILD}$). That is, Δ_{ITD} and Δ_{ILD} are not values in µs or dB but refer to the azimuth difference between α and the azimuth corresponding to the other cue. The parameters k_{ITD} and k_{ILD} are the estimated linear regression slopes at azimuth α (determining the individual binaural-cue contributions), and Q is the estimated response azimuth for consistent-cue stimuli corresponding to azimuth α. Parameter k_{ITD} (k_{ILD}) was estimated at each α by considering various azimuthal offsets (from −24° to +24° with 6° spacing) of the cue Δ_{ITD} (Δ_{ILD}) while setting the offset of the other cue, Δ_{ILD} (Δ_{ITD}), to zero. Thus, the model was fitted for each azimuth α, indicated by a yellow square in Figure 1c, by considering only items of the row (for k_{ITD}) and of the column (for k_{ILD}) that include that yellow square (indicated by the frame for an example azimuth of 9°). These estimates of k_{ITD} and k_{ILD} were then combined to derive the ILD weight, w_{ILD} (recall that $w_{ILD} + w_{ITD} = 1$). This differs from the weights reported by Macpherson and Middlebrooks (2002), who used the individual regression slopes (i.e., k_{ITD} and k_{ILD}) as weights. Instead, it is similar to the "binaural weighting index" reported by Kumpik et al. (2019), even though they forced the values to fall between −0.5 and +0.5 instead of 0 and 1. Finally, parameter Q was estimated as the average of the constants obtained in the regressions for ITD and ILD.

The data were analyzed using MATLAB R2018b (The MathWorks, Natick, MA). Statistical analyses were performed using SPSS Statistics 20 (IBM, Armonk, NY).

Results

Overall Binaural-Cue Weights. The blue circles in Figure 3 show the pretest ILD weights averaged across azimuths as a function of frequency band. The data were pooled across all participants, because up until the pretest, the procedure was identical between the two groups.

To test, whether ILD weights change gradually from low-to high-frequency bands, we ran a 5 (*frequency band*) x 4 (*azimuth*) ANOVA across the pretest data (to exclude any possible effects of training) of all participants. The ANOVA yielded a significant main effect of *frequency band* ($F(4,72) = 143.91$, $p < .001$, $\eta_p^2 = .889$). Follow-up pairwise comparisons showed significant differences in the ILD weight between all bands (all $p < .008$, Bonferroni-corrected) except between the mid-low and the multiband conditions ($p > .999$, Bonferroni-corrected). ILD weights gradually increased from the low ($M = .195$, $SEM = .030$) to mid-low ($M = .368$, $SEM = .040$) to mid-high ($M = .722$, $SEM = .015$) to the high band ($M = .796$, $SEM = .019$). The multiband condition ($M = .336$, $SEM = .026$) showed ILD weights similar to the mid-low band.

While the pretest ILD weight averaged across the four narrow bands was close to 0.5 ($M = .486$, $SEM = .019$), it was significantly smaller than 0.5 for the multiband ($M = .337$, $SEM = .026$; $t(18) = -6.17$, $p < .001$, two-tailed, $d_Z = -1.42$). This suggests that the low (i.e., ITD dominant) frequencies received more weight in the multiband stimulus.

The ANOVA additionally yielded a significant main effect of *azimuth* ($F(3,54) = 3.71$, $p = .017$, $\eta_p^2 = .171$). Follow-up pairwise comparisons showed significant differences in the ILD weight between 3° and 21° azimuth ($p = .015$, Bonferroni-corrected) with larger ILD weights at 3° ($M = .504$, $SEM = .019$) compared to 21° azimuth ($M = .458$, $SEM = .024$). This suggests that ILD weights were larger at central compared to lateral azimuths. Figure 4 depicts the detailed patterns of ILD weights across azimuths, separately for each band and group (blue circles denote data from experiment 1). Note how in most panels, the blue lines with filled circles (pretest results) trend downwards (indicating lower ILD weights) from central (left on the x-axis) to lateral (right on the x-axis) azimuths.

There was no significant *frequency band* x *azimuth* interaction ($F(5.38,96.87) = 1.56$, $p = .175$, $\eta_p^2 = .080$, Greenhouse-Geisser-corrected).

Binaural-Cue Reweighting. The results are visualized in Figure 4 (recall that blue circles denote data from experiment 1). The separate panels show the results in each frequency band (columns) and group (rows). The expected reweighting effect would be indicated by open symbols (posttest results) falling below (indicating lower ILD weight, therefore higher ITD weight) filled symbols (pretest results) in the ITD group (top row) and by open symbols falling above (indicating higher ILD weight) filled symbols in the ILD group (bottom row). We analyzed the data separately for each frequency band and group.

In the ITD group, a 2 (*pre- vs. posttest*) x 4 (*azimuth*) repeated measures (RM) ANOVA yielded no significant effects for the low band (all $p > .199$).

For the mid-low band, the 2×4 RM ANOVA yielded a significant main effect of *azimuth* ($F(3,27) = 6.70$, $p = .002$, $\eta_p^2 = .427$) as well as a significant *time* x *azimuth* interaction ($F(3,27) = 4.92$, $p = .007$, $\eta_p^2 = .353$), but no significant main effect of *time*. Follow-up pairwise comparisons showed that the *time* x *azimuth* interaction was driven by a significant decrease of ILD weights at 9° azimuth ($p = .022$, Bonferroni-corrected), but no difference between time points at other azimuths (all $p > .532$, Bonferroni-corrected). This suggests that participants reweighted the binaural cues in the expected direction, but only at 9° azimuth. The follow-up pairwise comparisons further showed significant differences in the ILD weights between 3° and 21° azimuth ($p = .005$, Bonferroni-corrected), as well as between 9° and 21° azimuth ($p = .049$, Bonferroni-corrected). ILD weights at 21° azimuth ($M = .286$, $SEM = .058$) were lower compared to 3° ($M = .422$, $SEM = .049$) and 9° azimuth ($M = .435$, $SEM = .051$). This suggests that ILD weights were larger at central compared to lateral azimuths.

For the mid-high band, the 2×4 RM ANOVA yielded no significant effects (all $p > .219$).

For the high band, the 2×4 RM ANOVA yielded a significant main effect of *azimuth* ($F(3,27) = 6.33$, $p = .002$, $\eta_p^2 = .413$) as well as a significant *time* x *azimuth* interaction ($F(3,27) = 3.08$, $p = .044$, $\eta_p^2 = .255$), but no significant main effect of *time*. Follow-up pairwise comparisons showed significant differences in the ILD weight between 3° and 21° azimuth ($p = .017$, Bonferroni-corrected) with lower ILD weights at 3° azimuth ($M = .796$, $SEM = .021$) compared to 21° azimuth ($M = .878$, $SEM = .024$), again suggesting larger ILD weights at central azimuths, but there were no significant weight changes from pre- to posttest at any azimuth (all $p > .107$, Bonferroni-corrected).

For the multiband condition, the 2×4 RM ANOVA yielded no significant effects (all $p > .086$).

In the ILD group, the 2×4 RM ANOVA for the low band yielded a significant main effect of *azimuth* ($F(3,24) = 4.25$, $p = .015$, $\eta_p^2 = .347$), but neither a significant main effect of *time* nor a *time* x *azimuth* interaction. Follow-up pairwise comparisons showed significant differences in the ILD weight between 3° and 21° azimuth ($p = .034$, Bonferroni-corrected) with higher ILD weights at 3° azimuth ($M = .239$, $SEM = .048$) compared to 21° azimuth ($M = .149$, $SEM = .033$). This again suggests that ILD weights were larger at central compared to lateral azimuths.

The 2×4 RM ANOVA yielded no significant effects for either the mid-low band (all $p > .200$) or the mid-high band (all $p > .289$).

For the high band, the 2×4 RM ANOVA yielded a significant main effect of *time* ($F(1,8) = 21.67$, $p = .002$, $\eta_p^2 = .730$) with smaller ILD weights in the pretest ($M = .756$, $SEM = .020$) compared to the posttest ($M =$
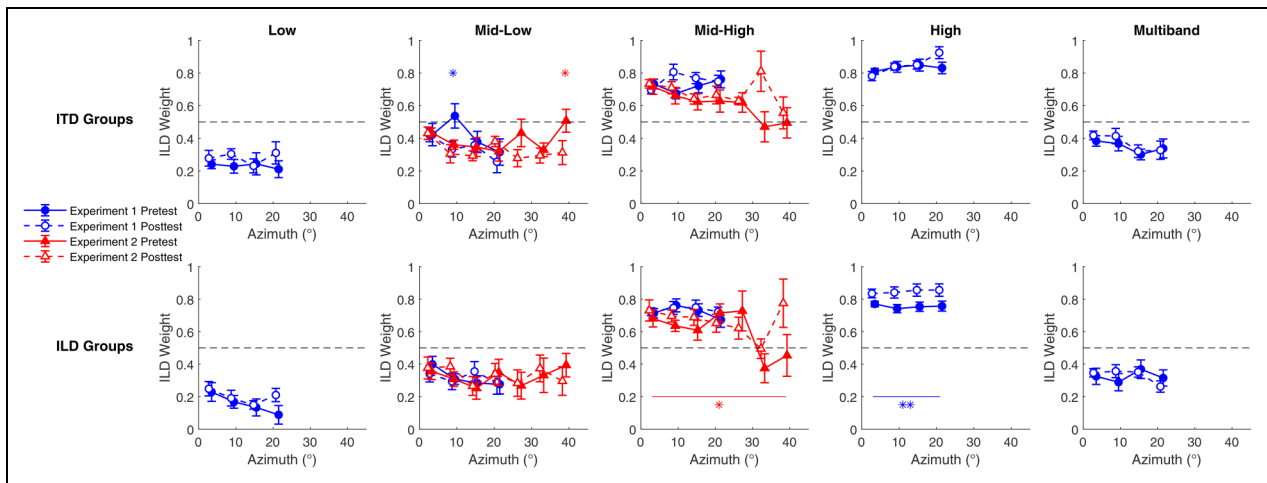
Figure 4. ILD weights (normalized, i.e., ILD weight = 1 − ITD weight) as a function of azimuth for each frequency band (columns) and group (rows). The top row shows results of the ITD groups (i.e., participants for whom ITDs were reinforced during training) and the bottom row shows results of the ILD groups (i.e., participants for whom ILDs were reinforced during training). Solid lines and filled symbols show pretest results, dashed lines and open symbols show posttest results. Blue circles show the results of experiment 1 and red triangles show the results of experiment 2. Error bars show the standard error of the mean. ILD weights gradually increase from low-, to mid-low-, to mid-high-, to high-frequency stimuli and on average are lower for lateral azimuths. Significant weight changes, indicated by the asterisks in the respective colors (* denotes $p < .05$, ** denotes $p < .01$), from pre- to posttest were observed for the ILD group in experiment 1 for the high-frequency band as well as in experiment 2 for the mid-high-frequency band. Additionally, a significant weight change from pre- to posttest was observed for the ITD group in experiment 1 for the mid-low-frequency band at 9° azimuth as well as in experiment 2 at 39° azimuth.
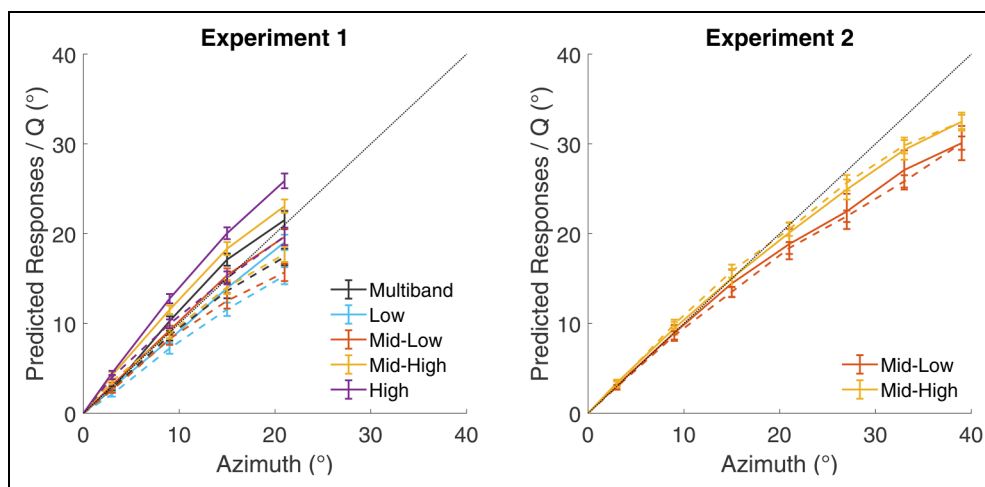

Figure 5. Predicted responses for consistent-cue combinations by the regression analysis (i.e., factor Q) as a function of azimuth. Solid lines show pretest results and dashed lines show posttest results. Error bars show the standard error of the mean. Lateralization slopes are steeper for higher frequencies. Experiment 1 shows response compression from pre- to posttest (i.e., lateralization slopes are shallower in the post- compared to the pretest) while in experiment 2, responses are already compressed in the pretest and therefore no difference between time points is observed.

.847, $SEM = .029$), but neither a significant main effect of *azimuth* nor *time* x *azimuth* interaction. This suggests that participants reweighted the binaural cues in the expected direction.

For the multiband condition, the 2×4 RM ANOVA yielded no significant effects (all $p > .090$).

Lateralization Functions. We additionally analyzed the lateralization pattern independent of binaural-cue weights for the different frequency bands and tests, because Klingel et al. (2021) observed a compression of responses from pre- to posttest. To that end, we fitted linear regressions to the responses for consistent-cue conditions predicted by the

regression analysis (i.e., the factor Q). Figure 5 shows the factor Q as a function of azimuth. Shallower functions in the post- compared to the pretest would indicate such a response compression. Using the factor Q is a more reliable estimate than fitting linear regressions to the responses in consistent-cue conditions alone, as the regression analysis includes all responses (i.e., also inconsistent-cue conditions).

We then submitted the resulting lateralization slopes of each participant to a 5 (*frequency band*) x 2 (*pre- vs. posttest*) RM ANOVA. The ANOVA yielded significant main effects of *frequency band* ($F(2.77, 49.81) = 39.24$, $p < .001$, $\eta_p^2 = .686$, Greenhouse-Geisser-corrected) and *time* ($F(1,18) = 39.76$, $p < .001$, $\eta_p^2 = .688$), with steeper lateralization slopes in the pretest ($M = 1.08$, $SEM = 0.04$) compared to the posttest ($M = 0.86$, $SEM = 0.05$). Additionally, there was a significant *frequency band* x *time* interaction ($F(4,72) = 8.93$, $p < .001$, $\eta_p^2 = .332$). Follow-up pairwise comparisons showed significant differences between the pre- and posttest for all frequency bands (all $p < .002$, Bonferroni-corrected). This suggests a compression of responses from pre- to posttest. They further showed significant differences between all frequency bands (all $p < .003$, Bonferroni-corrected) except between the multiband and the mid-low ($p = .286$, Bonferroni-corrected) as well as the mid-high bands ($p = .904$, Bonferroni-corrected), and between the low and mid-low bands ($p = .230$, Bonferroni-corrected). Lateralization slopes gradually increased from the low band ($M = 0.832$, $SEM = 0.04$) to the mid-high band ($M = 1.03$, $SEM = 0.04$) to the high band ($M = 1.13$, $SEM = 0.04$). The multiband ($M = 0.97$, $SEM = 0.04$) showed lateralization slopes similar to the mid-low band ($M = 0.88$, $SEM = 0.04$) and the mid-high band.

We further tested if lateralization slopes differed significantly from 1. A lateralization slope of 1 would indicate the expected lateralization behavior, smaller slopes would indicate compressed responses, and larger slopes would indicate expanded responses. In the pretest, lateralization slopes did not differ significantly from 1 for the low and mid-low band as well as the multiband (all $p > .053$, two-tailed). However, they were significantly larger than 1 for the mid-high ($t(18) = 4.09$, $p = .001$, two-tailed, $d_Z = 0.94$) and the high band ($t(18) = 6.97$, $p < .001$, two-tailed, $d_Z = 1.60$). In the posttest, slopes were significantly smaller than 1 for the low ($t(18) = -5.18$, $p < .001$, two-tailed, $d_Z = -1.19$), mid-low ($t(18) = -4.01$, $p = .001$, two-tailed, $d_Z = -0.92$), and mid-high band ($t(18) = -2.37$, $p = .029$, two-tailed, $d_Z = -0.54$) as well as the multiband ($t(18) = -2.71$, $p = .014$, two-tailed, $d_Z = -0.62$), but not for the high band.

Summary of Results

In this section we summarize the results of Experiment 1, which will then be discussed together with those of experiment 2 in the General Discussion. The baseline (i.e. pretest) results showed that binaural-cue weights gradually

changed from being ITD dominant to ILD dominant from low-, to mid-low-, to mid-high-, to high-frequency-band conditions. The multiband stimulus was weighted similarly to the mid-low band, meaning that ITDs dominated the percept when both ITD and ILD dominant bands were combined into one stimulus. On average across frequency bands, we further observed higher ILD weights at central compared to lateral azimuths, although deviations from this effect are observed for some conditions in band-specific analyses that included the posttest data. Additionally, we found steeper lateralization functions for higher frequencies and shallower lateralization functions in the post- compared to the pretest.

The results regarding training-induced changes in binaural-cue weights from pre- to posttest were unexpected. We expected to see an increase in the reinforced-cue weight and a respective decrease in the unreinforced-cue weight from pre- to posttest (i.e., reweighting) at least for the trained mid-high band, since it is similar in center frequency to the stimulus used in Klingel *et al*. (2021), where reweighting was shown in both the ITD and ILD groups. However, for the mid-high band, no reweighting was found in either group. There were, however, reweighting effects for some other conditions. The ILD group showed reweighting for the trained high band, which did not generalize to untrained conditions (i.e., mid-low and multiband). The ITD group, on the other hand, showed reweighting for the mid-low band, but only at 9° azimuth. This is surprising, as the mid-low condition was *not* trained. For all other bands, the ITD group showed no reweighting effects.

A closer look at the mid-high band showed that while it was similar in center frequency to the frequency band tested in Klingel et al. (2021), the baseline (pretest) binaural-cue weighting differed between the two studies. The ILD weight reported in Klingel et al. (2021) was lower than the ILD weight for the mid-high band in the present study. Instead, it closely matched the ILD weight for the mid-low band. This is likely due to the different bandwidths of the stimuli: In Klingel et al. (2021), stimuli were 1-octave wide while they were only 1/3-octave wide in the present study to minimize spectral overlap between frequency bands. As ITDs are known to dominate for broadband stimuli (Macpherson & Middlebrooks, 2002), it is likely that the lower-frequency part of the 1-octave wide band in Klingel et al. (2021) dominated the overall binaural-cue weighting of that stimulus.

Considering the design of the present study and of Klingel et al. (2021), we identified three potential reasons for the lack of reweighting in most conditions. 1) We had to restrict the azimuth range to ensure monotonic ILDs for all frequency bands. In Klingel et al. (2021), however, the largest effect was observed for more lateral azimuths (although the *time* x *azimuth* interaction did not reach significance). 2) We presented the different frequency bands randomly in one block. As they all had very different baseline weights, this may have been confusing, especially during training where visual

reinforcement was provided. And 3), we did not train participants with the mid-low band, which was closest in baseline ILD weight to the stimuli in Klingel et al. (2021). If the baseline ILD weight predicts the reweighting potential, training with this stimulus might increase chances for reweighting. To address these issues, we conducted a follow-up experiment. In experiment 2, we only included the mid-low- and mid-high bands, for which ILDs were monotonic up to 63° azimuth. This allowed us to use a wider range of azimuths than in experiment 1 (reinforced-cue azimuths up to 39° and unreinforced-cue azimuths up to 63° while experiment 1 included reinforced-cue azimuths only up to 21° and unreinforced-cue azimuths up to 45°). To avoid confusion, the two frequency bands were presented in blocks during practice as well as the pre- and posttest and only one frequency band (mid-low) was used during training.

Experiment 2

Based on the considerations discussed above, a follow-up experiment was performed. Experiment 2 was conducted identically to experiment 1 except for the changes described in the following.

Methods

Participants. Twenty new participants were recruited to participate in experiment 2. All participants gave written informed consent before starting the experiment and received monetary compensation for their participation. Ten participants (four female, mean age: 27.70 years, $SD = 5.29$, range 22–37) were randomly assigned to the ITD group. The remaining ten participants (four female, mean age: 25.60 years, $SD = 2.76$, range 20–30 years) were assigned to the ILD group.

Apparatus and Stimuli. Only the mid-low and the mid-high bands were presented. The reinforced-cue azimuth range was increased to $\pm 39°$ (and therefore the unreinforced-cue azimuth range to $\pm 63°$). As a consequence, ITDs ranged from -589 to $+589$ μs and ILDs ranged from -15.30 to $+15.30$ dB. The binaural cues used are shown in Figure 1a and b (all azimuths). For training, only the mid-low band was used, as its binaural-cue weighting closely matched the weighting reported in Klingel et al. (2021).

Procedure. To avoid confusion by presenting all frequency bands randomized in one block, the mid-low and the mid-high bands were now presented in separate blocks. The time course of tasks is shown in Table 1. The experiment started with a practice session with the mid-low band followed by the pretest with the mid-low band. Then, a practice session with the mid-high band was performed, followed by the pretest with the mid-high band. Afterwards, the training (with the mid-low band) was performed, followed by the posttests, first with the mid-low band and then with the mid-high band. The experimental protocol was again split into two sessions (either completed on consecutive days or

with a one-day break in between). In the first session, participants completed both practice sessions, both pretests, and the first training session. In the second session, participants completed the second training session and both posttests. The approximate duration was 1.5h for the first and 1h for the second session.

Practice. Each practice session (one for each frequency band) consisted of 32 trials with spatially consistent ITD/ILD-combinations in which each azimuth from $-45°$ to $+45°$ (with 6° spacing) was presented twice. The trials were presented in random order and, as in the training, visual reinforcement was provided.

Testing. Each pre- and posttest (one for each frequency band) consisted of 146 trials, each ITD/ILD-combination shown in Figure 1c (both yellow and blue) presented once. The trials were presented in random order and visual reinforcement was not provided.

Training. Each training session consisted of 252 trials presented in two blocks of 126 trials each. Within each block, trials were presented in a random order and each ITD/ILD combination shown in Figure 1d (both yellow and blue) was presented once. After every 63 trials, participants took a short break.

Analysis. The analysis was performed identically to experiment 1, except that binaural-cue weights were estimated for each azimuth α between 3° and 39° (i.e., additionally for azimuths of 27°, 33°, and 39°).

Results

Overall Binaural-Cue Weights. Figures 3 and 4 show a similar pattern in overall binaural-cue weights in the two experiments (blue circles for experiment 1 and red triangles for experiment 2).

Binaural-Cue Reweighting. The red triangles in Figure 4 show the ILD weights as a function of azimuth in separate panels for each frequency band and group. As in experiment 1, the expected reweighting effect would be indicated by open symbols (posttest results) falling below (indicating lower ILD weight, therefore higher ITD weight) filled symbols (pretest results) in the ITD group (top row) and by open symbols falling above (indicating higher ILD weight) filled symbols in the ILD group (bottom row). We analyzed the data separately for each frequency band and group.

In the ITD group, a 2 (*pre- vs. posttest*) x 7 (*azimuth*) RM ANOVA yielded a significant *time* x *azimuth* interaction for the mid-low band ($F(6,54) = 2.36$, $p = .042$, $\eta_p^2 = .208$), but neither a significant main effect of *time* nor *azimuth*. Follow-up pairwise comparisons showed that the interaction was driven by a significant decrease in ILD weight from pre- ($M = .508$, $SEM = .070$) to posttest ($M = .312$, $SEM = .073$) at 39° azimuth ($p = .021$, Bonferroni-corrected), but no significant differences at other azimuths (all $p > .127$, Bonferroni-corrected). This suggests that participants reweighted the binaural cues in the expected direction, but only at 39° azimuth.

For the mid-high band, the 2×7 RM ANOVA yielded a significant main effect of *azimuth* ($F(6,54) = 4.09$, $p = .002$, $\eta_p^2 = .312$), but neither a significant main effect of *time* nor *time* x *azimuth* interaction. Follow-up pairwise comparisons showed a significant difference between 3° and 15° azimuth ($p = .024$, Bonferroni-corrected) with larger ILD weights at 3° ($M = .725$, $SEM = .030$) compared to 15° azimuth ($M = .632$, $SEM = .032$). This suggests that ILD weights were larger at central compared to lateral azimuths.

In the ILD group, the 2×7 RM ANOVA yielded no significant effects for the mid-low band (all $p > .407$).

For the mid-high band, the 2×7 RM ANOVA yielded a significant main effect of *time* ($F(1,9) = 5.17$, $p = .049$, $\eta_p^2 = .365$), with smaller ILD weights in the pre- ($M = .600$, $SEM = .046$) compared to the posttest ($M = .666$, $SEM = .040$). This suggests that participants reweighted the binaural cues in the expected direction. There further was a significant main effect of *azimuth* ($F(2.45,22.08) = 5.97$, $p = .006$, $\eta_p^2 = .399$, Greenhouse-Geisser-corrected). Follow-up pairwise comparisons showed significant differences between 33° and 3° ($p = .039$, Bonferroni-corrected), 15° ($p = .030$, Bonferroni-corrected), as well as 21° azimuth ($p = .018$, Bonferroni-corrected). ILD weights were smaller at 33° azimuth ($M = .435$, $SEM = .058$) compared to 3° ($M = .706$, $SEM = .046$), 15° ($M = .649$, $SEM = .050$), and 21° azimuth ($M = .684$, $SEM = .051$). This suggests that ILD weights were again larger at central compared to lateral azimuths. There was no significant *time* x *azimuth* interaction.

Lateralization Functions. We again determined lateralization slopes based on the factor Q from the regression analysis and submitted them to a 2 (*frequency band*) x 2 (*pre- vs. posttest*) RM ANOVA. The ANOVA yielded a significant main effect of *frequency band* ($F(1,19) = 7.83$, $p = .011$, $\eta_p^2 = .292$) with steeper lateralization slopes for the mid-high band ($M = 0.90$, $SEM = 0.03$) compared to the mid-low band ($M = 0.82$, $SEM = 0.04$), but no significant main effect of *time* nor *time* x *frequency band* interaction. This suggests that there was no response compression from pre- to posttest.

All lateralization slopes were significantly smaller than 1. In the pretest, both the mid-low band ($M = .825$, $SEM = .064$; $t(19) = -2.73$, $p = .013$, two-tailed, $d_Z = -0.61$) and the mid-high band ($M = .892$, $SEM = .034$; ($t(19) = -3.17$, $p = .005$, two-tailed, $d_Z = -0.71$) had lateralization slopes smaller than 1. This suggests that responses were already compressed in the pretest. In the posttest, again both the mid-low band ($M = .804$, $SEM = .021$; $t(19) = -9.37$, $p < .001$, two-tailed, $d_Z = -2.09$) and the mid-high band ($M = .906$, $SEM = .026$; $t(19) = -3.58$, $p = .002$, two-tailed, $d_Z = -0.80$) had lateralization slopes smaller than 1.

Summary of Results

In experiment 2, we observed binaural-cue reweighting in the ILD group at the mid-high band, which was not trained, but not for the trained mid-low band. This is surprising, given that we did not observe reweighting at the mid-high band in experiment 1, even though it was trained. In the ITD group, we observed binaural-cue reweighting for the trained mid-low band only at 39° azimuth, which is in line with the expectation based on Klingel et al. (2021) that a stronger reweighting effect occurs for more lateral azimuths. However, we did not observe reweighting at 9° azimuth as we did in experiment 1. This reweighting did not generalize to the mid-high band.

We suspected that the trial-by trial switching between frequency bands might have contributed to the overall small reweighting effect in experiment 1 and, therefore, used blocked presentation during testing and included just one band during training in experiment 2. However, as only the *untrained* mid-high band in the ILD group showed more reweighting in experiment 2 compared to experiment 1, there is no evidence that trial-by-trial switching between frequency bands had an impact on reweighting.

Consistent with the results of experiment 1, we observed stronger ILD weights and steeper lateralization functions for the mid-high compared to the mid-low band. However, there was no response compression from pre- to posttest. Instead, especially at the larger azimuths tested in experiment 2, responses were already compressed in the pretest (i.e., pretest lateralization slopes were significantly smaller than 1 and shallower than in experiment 1). This might be attributable to the larger azimuth range presented in experiment 2. Namely, participants may have tended to respond inside of the visual markers of $\pm 45°$, even though the unreinforced cue extended that range. Note that while a larger azimuth range was also used in Klingel et al. (2020), who observed compression only in the posttest, they used visual markers every 15° and participants were therefore accustomed to respond beyond the markers.

General Discussion

The present study investigated whether binaural-cue weights gradually change from low- to high-frequency bands, whether binaural-cue reweighting can be induced in different frequency regions, and whether binaural-cue reweighting is specific to the trained frequency band or whether it generalizes to untrained frequency bands.

Binaural-Cue Weights Across Spectral Regions

We observed gradually increasing ILD weights with increasing spectral region of narrow-band noise stimuli. These baseline weights appear to be robust, as similar pretest weights were obtained in both experiments involving different participants and presentation modes (randomized vs. blocked). The finding of gradually increasing ILD weights with increasing frequency extends Macpherson and Middlebrooks (2002) results involving a low- and a high-pass stimulus but differs from Ahrens *et al.*'s (2020) observation, who

measured the contributions of different frequency bands in a multiband stimulus to binaural-cue weighting. Ahrens et al. (2020) found significant differences mainly for the lowest and the highest frequency band and similar weights for in-between frequency bands, irrespective of the absolute frequency range of presented bands. These different results likely resulted from differences in the methodology: While the frequency bands were presented in isolation in the present study, they were presented as part of a multiband stimulus in Ahrens et al. (2020), leading to an "edge effect" (i.e., stronger weighting of information at the edges). Instead, the present results are more in line with the sensitivity pattern to ITD and ILD as a function of center frequency for narrow-band noise stimuli. Compared to ITD thresholds, ILD thresholds are fairly constant across frequencies, but tend to decrease with increasing center frequency (Gabriel et al., 1992), although there are some local peaks in the threshold function (Goupell & Stakhovskaya, 2018). Additionally, ILDs are physically smaller for low- compared to high-frequency stimuli (leading to a shallower slope of the cue-versus-azimuth function) and therefore more informative at high frequencies. In contrast, ITDs are physically almost frequency-independent and ITD thresholds increase with increasing center frequency from approx. 800 Hz towards higher (at least up to 4000 Hz) frequencies (Buchholz et al., 2018; Gabriel et al., 1992; Klumpp & Eady, 1956; Smoski & Trahiotis, 1986; Trahiotis & Bernstein, 1990), because of decreasing access to fine-structure ITD.

For the multiband stimulus, comprising all tested bands, ITDs dominated the percept, similar to Macpherson and Middlebrooks (2002) broadband stimuli. While the pretest ILD weight averaged across the four frequency bands of experiment 1 was close to 0.5, it was significantly lower than 0.5 for the multiband stimulus. This suggests that for broadband sounds, the auditory system performs a weighted integration of binaural cues across frequencies, where ITDs receive more weight than ILDs.

Training-Induced Binaural-Cue Reweighting Across Spectral Regions

In the ILD group, a training-induced change in binaural-cue weights was observed for the trained high-frequency band (experiment 1) and for the untrained mid-high-frequency band (experiment 2), but not for the other conditions. In the ITD group, binaural-cue reweighting was observed for the untrained mid-low-frequency band (experiment 1, only at 9° azimuth) and for the same band when it was trained (experiment 2, only at 39° azimuth). While it may seem like these results are driven by noise in the pretest (given that participants get more accustomed with the setup and task over time), it should be noted that the chosen correction method for the post-hoc comparisons (Bonferroni) is very conservative and similar noise in other conditions did not

yield significant results, suggesting that the reported results are meaningful. There was no significant reweighting effect for the other frequency bands.

Given that both trained and untrained stimuli showed some effects in both groups, band-specific training does not seem to be crucial. This may suggest some across-frequency generalization of binaural-cue reweighting. Instead, the occurrence of reweighting seems to depend on an appropriate match between the stimulus' spectral region and the binaural cue that is reinforced. In the ILD group, effects were only observed for the mid-high and high bands, while in the ITD group, effects were only observed for the mid-low band (at specific azimuths). This suggests that inducing an increase in ILD weighting only works for sufficiently high frequencies: The strongest effect was observed for the high band (centered at 4000 Hz) and a less robust effect (only present in experiment 2, close to the significance threshold) occurred for the mid-high band (centered at 2520 Hz). Similarly, sufficiently low frequencies seem to be needed to induce an increase in ITD weighting, as only the mid-low band (centered at 1587 Hz) showed some effects in the ITD group. However, if the stimuli included the low-frequency temporal fine structure region that provides exquisite ITD sensitivity (i.e., for the low band centered at 1000 Hz and the multiband), no change in binaural-cue weights was observed. This suggests that ITD reweighting was based on envelope-ITD information. Compared to Klingel et al. (2021), the reweighting effects observed in the present study are relatively weak. In the ITD group, the effect for the mid-low band was so small that it only reached significance for specific azimuths. In the ILD group, the effect for the mid-high band reached significance in only one of the two experiments. Solely the effect for the high band in the ILD group showed comparable strength to the effects reported in Klingel et al. (2021). Nevertheless, the present results agree with the results reported by Klingel et al. (2021), who used bandpass-filtered (2–4 kHz) noise and found a reweighting effect in both directions (i.e., for both their ILD and ITD group). The 2–4 kHz stimulus both included "sufficiently high" and "sufficiently low" frequencies but did not touch the frequency region providing highest ITD sensitivity (i.e., the fine-structure region), and thus seems to be optimally suited to induce binaural-cue reweighting.

Additional Effects Independent of Binaural-Cue Reweighting

On average across frequency bands, we observed higher ILD weights at central compared to lateral azimuths. This is surprising, as the opposite pattern was observed in Klingel et al. (2020), who used a discrimination task to measure binaural-cue weights. In Klingel et al. (2021), who used the same auditory stimuli as Klingel et al. (2020) and the same

lateralization task as the present study, no significant effect of azimuth was observed. Therefore, the effect of azimuth does not appear to be robust and may depend on the task and auditory stimuli used. The pattern of higher ILD weights at central azimuths observed in the present study may be due to different sensitivity patterns for ITDs and ILDs across azimuths. Variations of ITD sensitivity across azimuths appear to be fully accounted for by the azimuth dependence of the ITD cue itself rather than by decreasing ITD sensitivity with increasing reference ITD (Smith & Price, 2014). Given that for the azimuth range included here ITDs increased relatively linearly with increasing azimuth (see Figure 1a), ITD sensitivity is expected to be constant across azimuths. ILD sensitivity, on the other hand, decreases with increasing reference ILD in addition to the physical dependence of ILD magnitude on azimuth (Brown et al., 2018). For the mid-low, mid-high, and high bands, the ILD magnitude increases less strongly (i.e., the slope of the cue-versus-azimuth function is shallower) for azimuths between approximately 21°–39° compared to azimuths between 3°–15° (see Figure 1b). The combined effect of decreasing ILD sensitivity with increasing reference ILD as well as the reduced increase in ILD magnitude for more lateral azimuths (leading to a reduced acoustic "goodness" of the cue) may, therefore, have contributed to the stronger ILD weighting at central compared to lateral azimuths.

Additionally, we found steeper lateralization functions for higher frequencies. This could be explained by two factors regarding ILDs, which dominate the percept at higher frequencies. First, for natural sounds arriving at the outer ears (at least for the frequency range used in this study), ITDs are almost frequency-independent, whereas ILDs are larger at higher frequencies (e.g., Middlebrooks & Green, 1991). Second, Bernstein and Trahiotis (2011) showed that even for a constant ILD magnitude, higher frequencies produce a larger extent of laterality than lower frequencies. The steeper lateralization functions at higher frequencies in our study could be due to both of these factors. We also observed shallower lateralization functions in the post- compared to the pretest in experiment 1. The same pattern was observed in Klingel et al. (2021) and attributed to a training-induced response compression, due to a response mapping to the visually reinforced range. This compression was not observed in experiment 2, possibly because responses were already compressed in the pretest. Importantly, any differences in the slope of the lateralization functions do not influence the estimated binaural-cue weights, as they affect k_{ITD} and k_{ILD} equally and therefore cancel out in the final weight estimation.

Limitations and Future Directions

A general limitation of the impact of training-induced changes in binaural-cue weights is their strong dependence on the auditory stimulus. In addition to requiring either sufficiently high or sufficiently low frequencies, reweighting seems to be restricted to conditions where low-frequency temporal-fine-structure ITDs do not contribute, which limits the ecological importance of the phenomenon in the normal auditory system where such cues are often available.

However, the results may be relevant for hearing-impaired or cochlear-implant (CI) listeners. Lacher-Fougère and Demany's (2005) results suggest that listeners with sensorineural hearing loss may not have access to fine-structure ITD cues, while retaining some sensitivity to envelope ITD cues. Also, CI listeners seem to have access to envelope ITD cues only. Many CI stimulation strategies encode ITDs only via the envelope of the stimulus waveform and, even when encoding ITDs via the pulse timing, CI listeners' sensitivity pattern resembles that for envelope ITDs in acoustic hearing (Bernstein & Trahiotis, 2002; Laback et al., 2007). In fact, binaural-cue reweighting has recently been observed in CI listeners when ITDs were encoded via the pulse timing of low-rate pulse trains (Klingel & Laback, 2021).

We presented auditory stimuli via headphones without HRTF filtering. This ensured that participants did not have access to monaural spectral localization cues, which might provide information about the stimulus azimuth and, thus, might prevent purely binaural-cue reweighting. Kumpik et al. (2010), for example, observed an increased weighting of monaural cues for azimuthal sound localization but no adaptation to changed binaural cues after training, when binaural cues were modified but monaural spectral cues were preserved at one ear. But open questions remain. For example, it is unclear whether the resulting lack of externalization in our study affected the weighting. Kumpik et al. (2019) used (non-individualized) HRTFs as well as reverberation to promote externalization and observed slightly (but significantly) ILD-dominant weights for their broadband stimuli. However, given that Macpherson and Middlebrooks (2002) also included HRTFs, the higher ILD weights in Kumpik et al. (2019) compared to Macpherson & Middlebrooks' wideband and our multiband stimuli more likely resulted from the added reverberation, which makes ITDs less reliable (Rakerd & Hartmann, 2010), than from the HRTFs or from externalization. Comparing our low- and high-frequency band as well as the multiband stimuli to Macpherson and Middlebrooks' low-pass, high-pass, and wideband stimuli results in very similar weight estimates. Additionally, it would be an interesting topic for future studies to clarify under which conditions binaural-cue reweighting and binaural-to-monaural-cue reweighting occurs for azimuthal sound localization. Particularly interesting are situations where listeners have access to spectral localization cues but access to low-frequency fine-structure ITD cues is prevented, e.g., due to traffic noise.

Summary and Conclusions

The present study is, to our knowledge, the first to show gradually increasing ILD weights (or decreasing ITD weights)

with increasing frequency (i.e., spectral region) of narrow-band stimuli. It therefore extends our knowledge on the duplex theory of sound localization, particularly on the transition from ITD- to ILD-dominant frequency regions. The across-frequency pattern of weights appears to be consistent with the corresponding patterns of ITD and ILD sensitivity from the literature as well as with the acoustic "goodness" of the cue. We further showed that binaural-cue reweighting is frequency dependent: To induce an increase in ILD weighting, the stimulus frequency needs to be sufficiently high and to induce an increase in ITD weighting, the stimulus frequency needs to be sufficiently low without including the low-frequency region providing fine-structure ITD cues. However, the observed increase in ITD weighting was so small that it reached significance only at specific azimuths. Which frequency band is used for training does not appear to influence the results systematically, suggesting some across-frequency generalization of binaural-cue reweighting. Such reweighting likely plays a role when listeners adapt to acoustic environments with altered robustness of binaural cues (e.g., when low-frequency ITD cues are masked by traffic noise) and has potential applications, such as supervised training after introducing a previously impeded cue to hearing devices such as cochlear implants.

ORCID iDs

Maike Klingel https://orcid.org/0000-0002-0019-1850
Bernhard Laback https://orcid.org/0000-0003-0929-6787

Research Data Availability Statement

The data is directly available for download from http://amtoolbox.org/amt-1.1.1-dev/auxdata/klingel2022/ and part of the Auditory Modeling Toolbox (AMT) development repository (i.e., they will be part of the next release). In the meantime, they can be accessed via the AMT functions "data_klingel2022" (raw data), and "exp_klingel2022" (ILD weights) as a download from sourceforge: https://sourceforge.net/p/amtoolbox/code/ci/master/tree/data/data_klingel2022.m and https://sourceforge.net/p/amtoolbox/code/ci/master/tree/experiments/exp_klingel2022.m, respectively.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Notes

1. https://assetstore.unity.com/packages/tools/integration/oculus-integration-82022
2. The package is available for download at https://sourceforge.net/p/expsuite/code/HEAD/tree/FrameWork/Unity/ViWoMI/.
3. https://www.oeaw.ac.at/isf/das-institut/software/hrtf-database

References

Ahrens A., Joshi S. N., & Epp B. (2020). Perceptual weighting of binaural lateralization cues across frequency bands. *Journal of the Association for Research in Otolaryngology*, *21*(6), 485–496. https://doi.org/10.1007/s10162-020-00770-3

Bernstein L. R., & Trahiotis C. (2002). Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli". *Journal of the Acoustical Society of America*, *112*(3), 1026–1036. https://doi.org/10.1121/1.1497620

Bernstein L. R., & Trahiotis C. (2011). Lateralization produced by interaural intensive disparities appears to be larger for high-vs low-frequency stimuli. *Journal of the Acoustical Society of America*, *129*(1), EL15–EL20. https://doi.org/10.1121/1.3528756

Brown A. D., Benichoux V., Jones H. G., Anbuhl K. L., & Tollin D. J. (2018). Spatial variation in signal and sensory precision both constrain auditory acuity at high frequencies. *Hearing Research*, *370*, 65–73. https://doi.org/10.1016/j.heares.2018.10.002

Buchholz J. M., Le Goff N., & Dau T. (2018). Localization of broadband sounds carrying interaural time differences: Effects of frequency, reference location, and interaural coherence. *Journal of the Acoustical Society of America*, *144*(4), 2225–2237. https://doi.org/10.1121/1.5058776

Carlile S. (2014). The plastic ear and perceptual relearning in auditory spatial perception. *Frontiers in Neuroscience*, *8*(237), 1–13. https://doi.org/10.3389/fnins.2014.00237

Dahmen J. C., Keating P., Nodal F. R., Schulz A. L., & King A. J. (2010). Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron*, *66*(6), 937–948. https://doi.org/10.1016/j.neuron.2010.05.018

David E. E.Jr, Guttman N., & van Bergeijk W. A. (1959). Binaural interaction of high–frequency complex stimuli. *Journal of the Acoustical Society of America*, *31*(6), 774–782. https://doi.org/10.1121/1.1907784

Deatherage B. H., & Hirsh I. J. (1959). Auditory localization of clicks. *Journal of the Acoustical Society of America*, *31*(4), 486–492. https://doi.org/10.1121/1.1907740

Gabriel K. J., Koehnke J., & Colburn H. S. (1992). Frequency dependence of binaural performance in listeners with impaired binaural hearing. *Journal of the Acoustical Society of America*, *91*(1), 336–347. https://doi.org/10.1121/1.402776

Gaik W. (1993). Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling. *Journal of the Acoustical Society of America*, *94*(1), 98–110. https://doi.org/10.1121/1.406947

Goupell M. J., & Stakhovskaya O. A. (2018). Across-channel interaural-level-difference processing demonstrates frequency

dependence. *Journal of the Acoustical Society of America*, *143*(2), 645–658. https://doi.org/10.1121/1.5021552

Grothe B., Pecka M., & McAlpine D. (2010). Mechanisms of sound localization in mammals. *Physiological Reviews*, *90*(3), 983–1012. https://doi.org/10.1152/physrev.00026.2009

Hebrank J., & Wright D. (1974). Spectral cues used in the localization of sound sources on the median plane. *Journal of the Acoustical Society of America*, *56*(6), 1829–1834. https://doi.org/10.1121/1.1903520

Henning B. (1980). Some observations on the lateralization of complex waveforms. *Journal of the Acoustical Society of America*, *68*(2), 446–454. https://doi.org/10.1121/1.384756

Keating P., Dahmen J. C., & King A. J. (2013). Context-specific reweighting of auditory spatial cues following altered experience during development. *Current Biology*, *23*(14), 1291–1299. https://doi.org/10.1016/j.cub.2013.05.045

Klingel M., Kopčo N., & Laback B. (2021). Reweighting of binaural localization cues induced by lateralization training. *Journal of the Association for Research in Otolaryngology*, *22*(5), 551–566. https://doi.org/10.1007/s10162-021-00800-8

Klingel M., & Laback B. (2021). Reweighting of binaural localization cues in bilateral cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology*, *23*(1), 119–136. https://doi.org/10.1007/s10162-021-00821-3

Klingel M., Spišák O., Seitz A., & Kopčo N. (2020). Reweighting of binaural localization cues induced by discrimination training. *Proceedings of the Forum Acusticum 2020*, pp. 925–929 *Lyon*. https://doi.org/10.48465/fa.2020.0108 https://hal.archives-ouvertes.fr/hal-03234181/document.

Klumpp R. G., & Eady H. R. (1956). Some measurements of interaural time difference thresholds. *Journal of the Acoustical Society of America*, *28*(5), 859–860. https://doi.org/10.1121/1.1908493

Kumpik D. P., Campbell C., Schnupp J. W., & King A. J. (2019). Re-weighting of sound localization cues by audiovisual training. *Frontiers in Neuroscience*, *13*(1164), 1–22. https://doi.org/10.3389/fnins.2019.01164

Kumpik D. P., Kacelnik O., & King A. J. (2010). Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. *Journal of Neuroscience*, *30*(14), 4883–4894. https://doi.org/10.1523/JNEUROSCI.5488-09.2010

Laback B., Majdak P., & Baumgartner W. D. (2007). Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing. *Journal of the Acoustical Society of America*, *121*(4), 2182–2191. https://doi.org/10.1121/1.2642280

Lacher-Fougère S., & Demany L. (2005). Consequences of cochlear damage for the detection of interaural phase differences. *Journal of the Acoustical Society of America*, *118*(4), 2519–2526. https://doi.org/10.1121/1.2032747

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

Middlebrooks J. C., & Green D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, *42*(1), 135–159. https://doi.org/10.1146/annurev.ps.42.020191.001031

Rakerd B., & Hartmann W. M. (2010). Localization of sound in rooms. V. Binaural coherence and human sensitivity to interaural time differences in noise. *Journal of the Acoustical Society of America*, *128*(5), 3052–3063. https://doi.org/10.1121/1.3493447

Shinn-Cunningham B. G., Durlach N. I., & Held R. M. (1998). Adapting to supernormal auditory localization cues. II. Constraints on adaptation of mean response. *Journal of the Acoustical Society of America*, *103*(6), 3667–3676. https://doi.org/10.1121/1.423107

Slattery W. H.III, & Middlebrooks J. C. (1994). Monaural sound localization: Acute versus chronic unilateral impairment. *Hearing Research*, *75*(1-2), 38–46. https://doi.org/10.1016/0378-5955(94)90053-1

Smith R. C., & Price S. R. (2014). Modelling of human low frequency sound localization acuity demonstrates dominance of spatial variation of interaural time difference and suggests uniform just-noticeable differences in interaural time difference. *PloS One*, *9*(2), e89033. https://doi.org/10.1371/journal.pone.0089033

Smoski W. J., & Trahiotis C. (1986). Discrimination of interaural temporal disparities by normal–hearing listeners and listeners with high–frequency sensorineural hearing loss. *Journal of the Acoustical Society of America*, *79*(5), 1541–1547. https://doi.org/10.1121/1.393680

Stecker G. C. (2010). Trading of interaural differences in high-rate gabor click trains. *Hearing Research*, *268*(1), 202–212. https://doi.org/10.1016/j.heares.2010.06.002

Stecker G. C., & Gallun F. J. (2012). Binaural hearing, sound localization, and spatial hearing. In K. L. Tremblay, & R. F. Burkard (Eds.), *Translational perspectives in auditory neuroscience: normal aspects of hearing* (pp. 383–433). Plural Publishing.

Strutt J. W. (1907). On our perception of sound direction. *Philosophical Magazine*, *13*(6), 214–232. https://doi.org/10.1080/14786440709463595.

Trahiotis C., & Bernstein L. R. (1990). Detectability of interaural delays over select spectral regions: Effects of flanking noise. *Journal of the Acoustical Society of America*, *87*(2), 810–813. https://doi.org/10.1121/1.398892

Van Wanrooij M. M., & van Opstal A. J. (2007). Sound localization under perturbed binaural hearing. *Journal of Neurophysiology*, *97*(1), 715–726. https://doi.org/10.1152/jn.00260.2006

Wightman F. L., & Kistler D. J. (1992). The dominant role of low–frequency interaural time differences in sound localization. *Journal of the Acoustical Society of America*, *91*(3), 1648–1661. https://doi.org/10.1121/1.402445

Wright B. A., & Fitzgerald M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proceedings of the National Academy of Sciences*, *98*(21), 12307–12312. https://doi.org/10.1073/pnas.211220498

Xie B. (2013). *Head-related transfer function and virtual auditory display*. J. Ross.