REWEIGHTING OF BINAURAL LOCALIZATION CUES INDUCED BY DISCRIMINATION TRAINING

Maike Ferber[1,2] **Ondrej Spišák**[3] **Aaron Seitz**[4] **Norbert Kopčo**[3]

[1] Department of Cognition, Emotion, and Methods in Psychology, University of Vienna, Austria
[2] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria
[3] Institute of Computer Science, P. J. Šafárik University in Košice, Slovakia
[4] Department of Psychology, University of California, Riverside, CA, USA
maike.ferber@univie.ac.at

ABSTRACT

Normal-hearing listeners weight binaural localization cues depending on the sound's frequency content. Interaural time differences (ITDs) dominate at low frequencies and interaural level difference (ILDs) dominate at high frequencies. The contribution of each cue to an azimuthal localization percept also depends on environmental factors such as room acoustics. Furthermore, ITD/ILD weighting was shown to change for stable sound conditions through a lateralization training in virtual reality (VR) using visual feedback. This study aims to induce similar reweighting using a simple left/right discrimination task without VR.

Participants were divided into two groups, a group trained to increase their ILD weighting and a no-training control group. Both groups completed an identical pre and post assessment involving a relative discrimination task without feedback, using various combinations of spatially inconsistent ITD and ILD. The training group additionally completed three sessions of an adaptive relative discrimination training including feedback (correct/incorrect) always consistent with the ILD azimuth. After each incorrect response, the auditory stimulus was repeated with the correct response shown on screen.

Responses followed the ILD azimuth significantly more often in the posttest than in the pretest for the training group but not for the control group, suggesting that binaural cue reweighting can be achieved by a simple discrimination task.

1. INTRODUCTION

Normal-hearing listeners primarily rely on two binaural cues for sound localization in the horizontal plane, the interaural time difference (ITD) and the interaural level difference (ILD) [1]. They apply frequency-dependent weights when combining these cues to determine the perceived azimuth of a sound source. While ITDs dominate at low frequencies (up to approximately 1.5 kHz), ILDs are dominant for high-frequency stimuli, which is known as the duplex theory of sound localization [2-3]. However, binaural cue weighting is not only determined by the sound's frequency content. For instance, it also depends on the overall level of the sound [4], whether the listener actively manipulates one of the cues [5], and on room acoustics [6].

Furthermore, listeners adapt to manipulated sound localization cues by changing the perceptual weights given to these cues. Several studies report a stronger weighting of monaural spectral cues compared to binaural localization cues for horizontal sound localization after wearing unilateral earplugs [7-9]. In another study, listeners were observed to increase their ILD weighting when ITDs were randomized, even though the auditory stimuli were irrelevant to the task [10]. Finally, normal-hearing listeners were shown to increase the relative weight given to either ITD or ILD, depending on which cue was visually reinforced during a lateralization training in a virtual audio-visual environment [11].

While the latter study [11] successfully induced reweighting, it was rather complex and required sophisticated equipment (e.g., virtual reality glasses). In addition, a compression of the lateralization responses from pre- to posttest was observed that might complicate the interpretation of the results. Therefore, we now seek to develop a simpler method to induce reweighting of the two binaural cues, which is not susceptible to response compression and which does not require sophisticated equipment. A left/right discrimination task matches these criteria. However, an early study attempting to reweight ITD and ILD through left/right discrimination training was not successful [12]. The stimuli used in that study were noise bands centered at 500 Hz, the binaural cues were close to the binaural cue threshold, and feedback was given after participants had responded. Therefore, potential reasons for the null result are that (1) the stimulus was in a frequency range where only ITDs but not ILDs arise naturally, (2) the binaural cues were not sufficiently salient, and (3) no corrective measures were taken if the participant responded "incorrectly". Based on the methods used in [11], we address these points in the present study by (1) using a noise band centered at 2.8 kHz to ensure that neither of the two cues are weighted predominantly strongly, (2) using a two-interval, relative discrimination task instead of a one-interval task, which allows us to test many spatial configurations as we are not restricted to azimuths close to the midline, and (3), in addition to giving feedback (correct/incorrect) after each response, repeating the auditory stimuli after each "incorrect" response with the "correct" response shown on the screen.

2. METHODS

2.1 Participants

21 participants completed the experiment, 10 of which were assigned to a training group and trained to increase their ILD weighting. The remaining 11 participants were assigned to a no-training control group. All but two participants had normal hearing (≤ 20 dB HL threshold at frequencies between 250 and 8000 Hz) as tested with an audiometer. The remaining two participants had thresholds ≤ 30 dB HL. The participants received monetary compensation for their participation. All procedures were approved by the UPJS Ethics Committee.

2.2 Apparatus and Stimuli

The experiment was controlled by Psychtoolbox-3 for MATLAB (The MathWorks, Natick, MA), running on a personal computer. Auditory stimuli were generated by an external sound card (RME Fireface 400) and presented via headphones (Sennheiser HD 800 S). Participants were seated in front of a monitor inside a soundproof booth and responded with the left and right arrow keys of a keyboard.

The same auditory stimuli as in [11] were used, namely 500-ms (including 50-ms raised-cosine on/off ramps) narrow-band white noise bursts with a bandwidth of one octave, geometrically centered at 2.8 kHz. Combinations of ITDs (ranging from -662 to +662 µs) and ILDs (ranging from -19.4 to +19.4 dB) were imposed on these stimuli. We used a set of ITDs and ILDs corresponding to azimuths between -70.2° (left) and +70.2° (right) (right) with a spacing of 3.6° between azimuths based on the calculations of Xie [13], who used the head related transfer functions (HRTFs) of the KEMAR dummy head with DB-61 small pinna at a source distance of 1.4 m. The stimulus choice ensured that both ITD and ILD increased monotonically with increasing azimuth. To discourage listeners from using absolute levels rather than ILD cues for determining the azimuth of a stimulus, the presentation level of each noise burst was independently roved by ±2.5 dB.

2.3 Procedure

The experiment was conducted on three consecutive days. On the first day, all participants completed a pre-training to get accustomed to the task, the first assessment (pretest), and, if they were assigned to the training group, the first training session. On the second day, only participants assigned to the training group came in and completed their second training session. On the third day, participants assigned to the training group performed the third training session and all participants completed the second assessment (posttest).

All tasks were two-interval, two-alternative-forced-choice tasks. The inter-stimulus-interval was 0 ms, however, due to the 50-ms on/off ramps a small gap was perceived between the two stimuli.

2.3.1 Pre-Training

The procedure of the pre-training was the same as of the training (see below), except for the auditory stimuli used. In the pre-training, only consistent-cue stimuli were used, meaning that the ITD and ILD of a stimulus always corresponded to the same azimuth. On each trial, one azimuth between ±59.4° with a 3.6° spacing between azimuths was chosen randomly and both the ITD and ILD of the first stimulus corresponded to that azimuth. The ITD and ILD of the second stimulus then corresponded to the azimuth either 10.8° to the left or 10.8° to the right (also chosen randomly) of the first stimulus' azimuth. Participants completed one block of 50 trials. We then checked their accuracy and if it was below 0.75, the pre-training was repeated until the criterion was reached or the participant was asked not to continue his/her participation in the study.

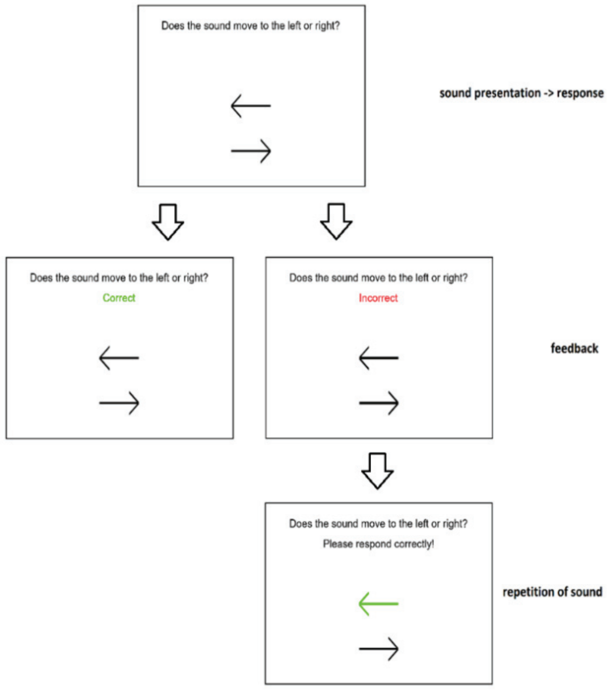

Figure 1. Experimental procedure. During the assessments, only the first screen was shown and no feedback was given. During the pre-training and training, feedback was given after each response and if the response was "incorrect" the stimulus pair was repeated with the "correct" response shown on screen (green arrow), requiring the subject to respond "correctly".

2.3.2 Assessment

During the assessments, participants did not receive feedback. They only listened to the two consecutive stimuli and responded on what side of the first stimulus they perceived the second stimulus (top panel of Figure 1).

The same inconsistent ITD/ILD combinations as in [11] were used. Each stimulus pair used two azimuths, az1 and az2, such that the first stimulus contained an ITD corresponding to az1 and an ILD corresponding to az2 and the second stimulus contained an ILD corresponding to az1 and an ITD corresponding to az2, or vice versa. The az1 was located between ±45° (uniformly distributed, 3.6° spacing between azimuths) and the az2 was located ±25.2° re. az1 (also with a uniform distribution and 3.6° spacing between azimuths). Therefore, one of the cues always moved to the right while the other cue moved to the left by the same amount. It was assumed that the perceived direction of motion is indicative of which cue contributed more to the azimuthal percept. There were 4 repetitions of each az1/az2 combination. We additionally included 52 catch trials that consisted of consistent-cue combinations to check, if participants are doing the task correctly. In the catch trials, the first stimulus corresponded to an azimuth between ±45° (uniformly distributed, 3.6° spacing) and the second stimulus corresponded to an azimuth shifted by 10.8° either to the left or right re. the first stimulus. That is in the catch trails, both ITD and ILD moved either to the left or to the right.

Each assessment consisted of 892 trials in a randomized order with a short break after each 100 trials. It took about 50 min to complete.

2.3.3 Training

During training, participants received feedback after each response (Figure 1). The feedback was always consistent with the ILD. That is, if the second stimulus ILD was smaller than the first stimulus ILD and the participant responded left, they would receive the positive feedback "correct", otherwise the response was "incorrect". After each "incorrect" response, the stimulus pair was repeated with the "correct" response shown on the screen and participants had to press the "correct" button to move on.

A two-down-one-up adaptive procedure was used for training. For the first stimulus of each trial, one azimuth between ±30.6° with a 3.6° spacing between azimuths was chosen randomly and both the ITD and ILD corresponded to that azimuth. The ILD azimuth of the second stimulus was shifted either to the left or right (chosen randomly) of the first stimulus' azimuth. The amount of this shift depended on the adaptive staircase: At the beginning of each session, the ILD azimuth was shifted by 32.4° and with each staircase adaptation, this shift was reduced by 3.6°. The ITD azimuth of the second stimulus was shifted from the ILD azimuth of the second stimulus in the direction towards the fist stimulus' azimuth by either 25.2°, 21.6°, or 18° (chosen randomly). Three staircases were adapted in parallel, one each for the three cue disparities (25.2°, 21.6°, and 18°). This ensured that at the beginning of each training session, both cues of the second stimulus were shifted to the same side of the first stimulus, making the task solvable independent of the binaural cue weights. For a cue disparity of 25.2°, the ITD azimuth was on the opposite side of the first stimulus' azimuth after the third staircase adjustment, for a cue disparity of 21.6° after the fourth staircase adjustment and for a cue disparity of 18° after the fifth staircase adjustment.

In each training session, participants completed 500 trials with a short break after each 100 trials. Each training session took about 30 min to complete.

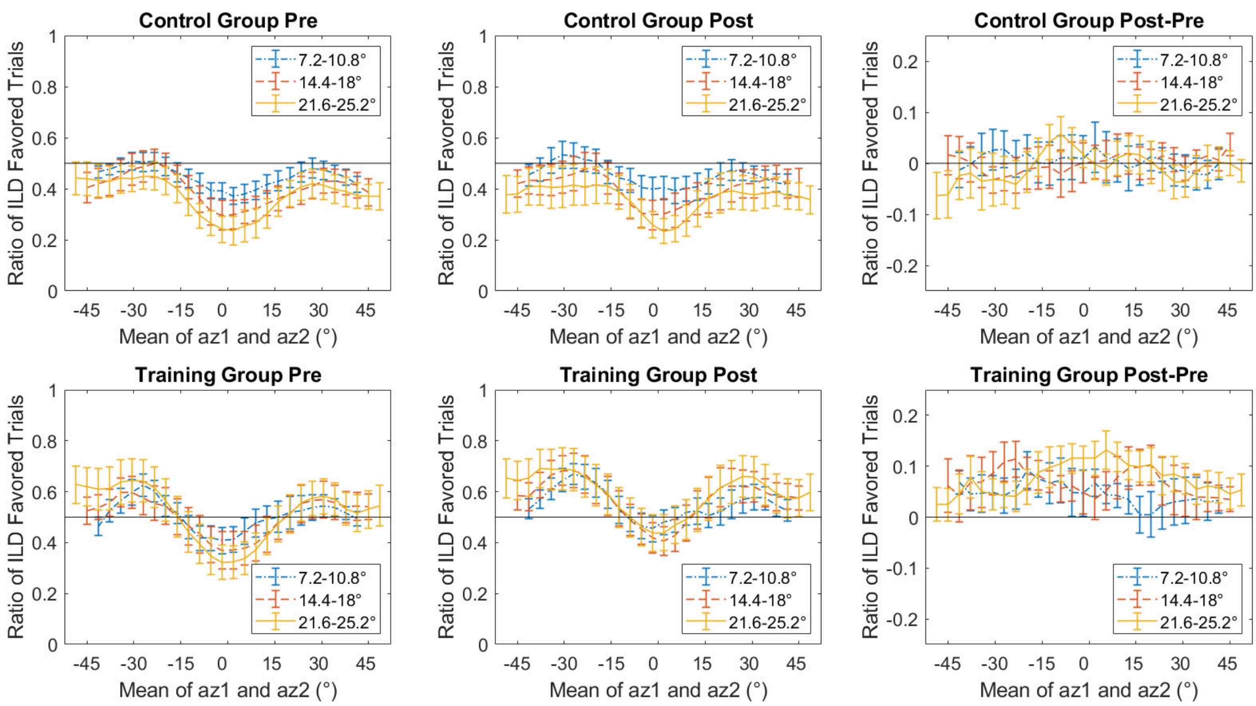

Figure 2. Ratio of trials in which responses followed the ILD azimuth as a function of the mean of az1 and az2 plotted separately for different az1-az2 cue disparities. Data are smoothed across cue disparity pairs (separate lines; disparity 3.6° is omitted) and azimuths (smoothing window with width of ±7.2°). The error bars show the standard error of the mean.

2.4 Analysis

The data were analyzed using MATLAB R2018b (The MathWorks, Natick, MA). For the assessments, we calculated the ratio of inconsistent-cue trials in which responses followed the ILD azimuth. Statistical analyses were performed using SPSS Statistics 20 (IBM, Armonk, NY).

3. RESULTS

Figure 2 shows the ratio of inconsistent-cue trials in which responses followed the ILD as a function of the az1/az2 mean and separately for different az1-az2 disparities. The top row shows the no-training control group, the bottom row shows the ILD training group. The first column shows the pretest data, the second column the posttest data, and the third column shows the posttest-pretest difference. The plotted data were smoothed across azimuths by using a rectangular smoothing window with a width of ±7.2° and averaged across neighboring az1-az2 disparities (the smallest disparity of 3.6° is omitted).

We ran a mixed-design ANOVA with the data shown in Figure 2. To ensure independent data points despite smoothing, only azimuths -32.4°, -16.2°, 0°, 16.2°, and 32.4° were included, leading to a 2 (*pre- vs. posttest*) x 3 (*mean azimuth disparities of 9°, 16.2°, and 23.4°*) x 5 (*azimuth*) x 2 (*control group vs. training group*) ANOVA. The ANOVA yielded significant main effects of *testing time* ($F(1,19) = 6.80$, $p = .017$, $\eta_p^2 = .263$) and *azimuth* ($F(2.28,43.39) = 17.56$, $p < .000$, $\eta_p^2 = .480$, Greenhouse-Geisser corrected) as well as several significant interactions. A significant interaction of *testing time* x *group* ($F(1,19) = 6.61$, $p = .019$, $\eta_p^2 = .258$) confirmed a significant increase in responses following the ILD for the training group while there was no significant change between the two assessments for the control group. A significant *azimuth disparity* x *group* interaction ($F(2,38) = 4.72$, $p = .015$, $\eta_p^2 = .199$) suggests that there were no significant differences between the azimuth disparities for the training group while the ratio of ILD-following responses decreased with increasing disparity for the control group. Finally, a significant *azimuth disparity* x *azimuth* interaction ($F(8,152) = 4.25$, $p < .000$, $\eta_p^2 = .183$) indicates that the cue weighting at different azimuth disparities varies with stimulus azimuth. Specifically, the ratio of ILD-following responses decreased with increasing disparity for stimuli near the midline and it increased at the lateral azimuths, especially for larger disparities. Confirming this observation, follow-up ANOVAs showed no significant difference between azimuths and cue disparities when 0° azimuth was excluded while at 0° azimuth, responses followed the ILD azimuth less often, particularly the larger the cue disparity.

Figure 3 shows the change in the ratio of inconsistent-cue trials in which responses followed the ILD azimuth in the posttest vs pretest, separately for each participant. It shows that the individual participants' ILD weight (defined here as the ratio of ILD-following responses) varied across a wide range of values. For 9 out of 10 training group participants, responses followed the ILD more often in the post- than in the pretest, while the ratio of ILD-following responses increased or decreased for approximately the same number of participants in the control group.

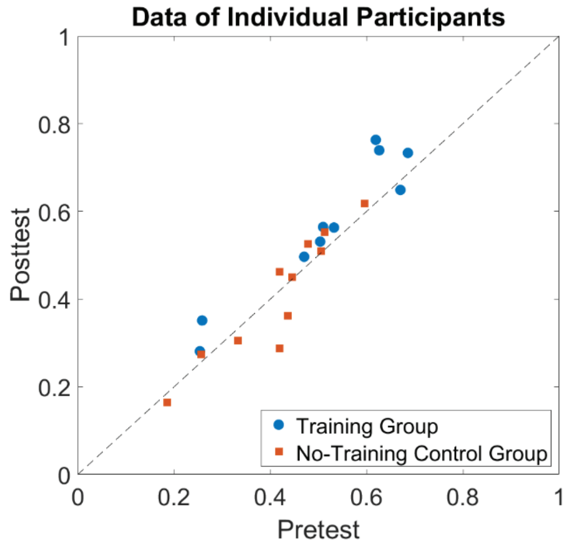

Figure 3. Posttest-to-pretest scatterplot of the ratio of inconsistent-cue trials in which responses followed the ILD azimuth for each participant.

4. DISCUSSION

The present results suggest that binaural cue reweighting can be achieved not only through lateralization training but also by a simple left/right discrimination task.

It has previously been shown that both ITD and ILD weighting can be increased through a lateralization training in a virtual audio-visual environment, depending on which of the two cues is visually reinforced [11]. However, the methods used in that study were rather complex and required a VR setup. Furthermore, a compression of responses from pre- to posttest was observed, potentially complicating the interpretation of the results. Another recent study observed an increase in ILD weighting when ITDs were randomized during a simple visual oddball task, even though the auditory stimuli accompanying the visual stimuli were not needed to perform the training task [10]. However, in that latter study, no increase in ITD weighting was observed when ILDs were randomized, and an increase in ILD weighting was also found for a condition in which ITDs were not randomized, making it difficult to draw strong conclusions. In comparison, the present study used a simple left/right discrimination task which does not require VR equipment and is not susceptible to response compression. Also, it included a no-training control group which did not show binaural cue reweighting, confirming that the difference in ILD weighting observed in the training group were indeed induced by the training.

Although an early study attempting to reweight the binaural cues with a left/right discrimination training did not yield significant results [12], in the present study, training-group responses followed the ILD azimuth significantly more often in the post- compared to the pretest. These inconsistent results might be explained by differences in the experimental design. The current study used auditory stimuli for which neither of the two cues is weighted predominantly strongly, it involved a variety of spatial configurations instead of only using cues close to the binaural cue threshold, and it provided multi-modal feedback while requiring a corrective response after “incorrect” responses.

We observed the ILD weights to be lower for central azimuths. This is consistent with the previous lateralization study [11], which also showed a trend for this pattern, although it did not reach significance. Furthermore, we observed the ILD weights at the midline to be lower the larger the azimuth disparity. This makes sense given the experimental design: For an ILD weight $\neq 0.5$, the amount of perceived movement is expected to grow with the azimuth disparity, as the difference between the ILD azimuths of the first and second stimulus equals the azimuth disparity of each stimulus.

The current study also introduced a new method for measuring binaural cue weights. Traditionally, ITD/ILD trading ratios were measured by fixing one of the cues and letting the participant adjust the other cue until the auditory image is centered [4]. However, this leads to a stronger weighting of the to-be-adjusted cue, probably because attention is shifted towards it [5]. Estimating binaural cue weights based on the lateralization of stimuli with spatially inconsistent ITD and ILD [2,11] is not susceptible to this bias, but requires sophisticated equipment (e.g., head tracking, if responses are given with the head pointing method). The method used here is also not susceptible to an attentional bias as no cue is actively manipulated and it only requires a simple left/right response, therefore not needing a head tracking device.

Importantly, the current study only included one type of stimulus. Therefore, future studies need to examine how effective the training will be for stimuli with different spectral content, for ITD training, and how well the training generalizes to different frequencies and stimuli.

In conclusion, binaural cue reweighting can be induced with a simple left/right discrimination task, which might make a training more easily accessible for a wide range of listeners, e.g. after introducing a previously impeded cue to hearing devices such as cochlear implants.

5. ACKNOLEDGEMENTS

The study was supported by the uni:docs fellowship program for doctoral candidates of the University of Vienna, the WTZ grant ASH (Adaptability in Spatial Hearing), #MULT 07/2020, APVV DS-FR-19-0025, and the EU Marie Curie grant H2020-MSCA-RISE-2015 #691229.

6. REFERENCES

[1] J.C. Middlebrooks, and D.M. Green: “Sound Localization by Human Listeners,” *Annual Review of Psychology*, Vol. 42, No 1, pp. 135-159, 1991.

[2] E.A. Macpherson, and J.C. Middlebrooks: “Listener Weighting of Cues for Lateral Angle: The Duplex Theory of Sound Localization Revisited,” *Journal of the Acoustical Society of America*, Vol. 111, No. 5, pp. 2219-2236, 2002.

[3] J.W. Strutt: “On Our Perception of Sound Direction,” *Philosophical Magazine*, Vol. 13, No. 6, pp. 214-232, 1907.

[4] B.H. Deatherage, and I.J. Hirsh: “Auditory Localization of Clicks,” *Journal of the Acoustical Society of America*, Vol. 31, No. 4, pp. 486-492, 1959.

[5] A.G. Lang, and A. Buchner: “Relative Influence of Interaural Time and Intensity Differences on Lateralization is Modulated by Attention to One or the Other Cue,” *Journal of the Acoustical Society of America*, Vol. 124, No. 5, pp. 3120-3131, 2008.

[6] B. Rakerd, and W.M. Hartmann: “Localization of Sound in Rooms. V. Binaural Coherence and Human Sensitivity to Interaural Time Differences in Noise,” *Journal of the Acoustical Society of America*, Vol. 128, No. 5, pp. 3052-3063, 2010.

[7] P. Keating, J.C. Dahmen, and A.J. King: “Context-Specific Reweighting of Auditory Spatial Cues Following Altered Experience During Development,” *Current Biology*, Vol. 23, No. 14, pp. 1291-1299, 2013.

[8] D.P. Kumpik, O. Kacelnik, and A.J. King: “Adaptive Reweighting of Auditory Localization Cues in Response to Chronic Unilateral Earplugging in Humans,” *Journal of Neuroscience*, Vol. 30, No. 14, pp. 4883-4894, 2010.

[9] M.M. van Wanrooij, and A.J. van Opstal: “Sound Localization under Perturbed Binaural Hearing,” *Journal of Neurophysiology*, Vol. 97, pp. 715-726, 2007.

[10] D.P. Kumpik, C. Campbell, J. Schnupp, and A.J. King: “Re-weighting of Sound Localization Cues by Audiovisual Training,” *Frontiers in Neuroscience*, Vol. 13, No. 1164, pp. 1-22, 2019.

[11] M. Ferber, B. Laback, and N. Kopco: “Vision-Induced Reweighting of Binaural Localization Cues,” *Journal of the Acoustical Society of America*, Vol. 143, No. 3, pp. 1813-1813, 2018.

[12] L.A. Jeffress, and D. McFadden: “Differences of Interaural Phase and Level in Detection and Lateralization,” *Journal of the Acoustical Society of America*, Vol. 49, No. 4B, pp. 1169-1179, 1971.

[13] B. Xie: *Head-related transfer function and virtual auditory display*, J. Ross, Plantation, FL, 2013.