Rate and persistence of adaptation remains consistent across multiple sets of spectral cues for sound localisation

Paul Friedrich
Department of Biology, Leipzig University

February 2023

1 Introduction

Accurate localization of sounds is required for navigation, communication, predation and escape. It is therefore not surprising that neural and structural specializations have evolved to perform this complex task. The human cortex for example continuously integrates sensory input across modalities to estimate the relative direction and distance of objects (things) in the environment. Whereas the topography of visual and somatosensory space is represented by a point-by-point mapping of primary receptors, the auditory scene must first be computed from incoming sound waves. In this scenario, the direction-dependent filtering profile of the pinnae and upper body result in spectral cues that aid in the inference of sound source location.As these filters underlie lifelong changes, beyond the developmental period, the auditory system must maintain its ability to recalibrate the mapping of spectral cues to locations in space. [source]

2 Methods and Materials

Participants
X participants (x males and y females, mean age a +- b years) took part in the experiment. They were informed about the relevant experimental procedures before providing their consent. Participants had no history of hearing disorder and had normal or corrected vision. The experimental procedures were approved by the ethics committee at Leipzig University. Monetary compensation was provided to the participants based on the time they spend at the experiment.

Earmolds
To alter participants' perception of sound source elevation, their pinnae were modified by silicone molds. As a result, spectral cues derived from the shape of the external ears were changed sufficiently to diminish participants ability to locate sounds on the vertical axis. Earmolds were created by applying fast curing, skin safe silicone (SkinTite, Smooth-On, Macungie, USA) to the cymba conchae, cavum conchae and the antihelix while keeping the ear canal unobstructed [Fig. X]. Volume and shape of the earmolds varied across individuals to achieve a similar degree of disruption of sound localization ability [Fig. X]. – opted for behavioral effect

Experimental procedure
The experiment was designed to adapt listeners to multiple sets of spectral cues for sound localization: participants wore two distinct pairs of earmolds consecutively (further denoted as M1 and M2), each over the course of a five-day adaptation period. The acoustical and behavioral impact of the molds were measured and the trajectories of participants' adaptation throughout the following days were recorded to capture the occurrence of metaplasticity. To test whether learning a new set of spectral cues interferes with a previously learned mapping, the persistence of adaptation to the initial earmolds (M1) was measured after participants adapted to the second pair of molds (M2) for five days. This measurement was repeated for M2 after another five days without earmolds [Fig X]. During their first visit, participants were familiarized with the environment, equipment and procedures of the free-field localization task. To minimize procedural learning during the experiment, participants completed at least one localization run and were free to continue practicing until they felt comfortable with the task. No feedback was given. After the initial familiarization, participants performed one localization task to measure their baseline localization accuracy. Once this task was completed, participants' HRTFs were acquired. Participants' ears were then modified by fitting the first pair of earmolds (M1), before immediately repeating the localization task. To capture changes to spectral cues induced by the earmolds, another set of HRTFs was acquired. From that day on, the silicone molds were worn by the participants for five consecutive days. Throughout this adaptation period, participants underwent a daily routine of training sessions. Each training session was followed by a free-field localization task. On the final day of the first adaptation period (day 5 in Fig. X), participants completed a short training session and a localization test. As a control, a subset of x participants performed an additional localization task presenting stimuli with varying spectral content in each trial. The earmolds were then removed, and participants' localization accuracy was immediately measured again to test for aftereffects. The second pair of earmolds (M2) was then fitted to the participants' ears, their initial localization accuracy was measured and HRTFs were acquired. Over the next 5 days, the procedure was repeated as described for the first adaptation period, including daily training sessions and localization tasks. After mold removal at the end of the second adaptation period (day 10 in Fig. X), the initial earmolds (M1) were briefly re-inserted and participants completed a localization run. To compare the persistence of M1 and M2 adaptation, participants returned to the lab after 5 days for a final localization test with the second molds (M2) re-inserted. The localization tests, binaural recordings and training sessions were conducted in a hemi-anechoic room (a x b x c m). Participants were seated in a comfortable chair in front of a spheric array of 45 loudspeakers (Mod1, Orb Audio, New York, USA) with a diameter of 1.4 m, covering the frontal hemisphere. Loudspeakers were hidden by an acoustically transparent black curtain to avoid visual cues of the sound source positions. Optionally, a small light emitting diode that was visible through the curtain indicating the central location of the frontal hemifield (0° Azimuth, 0° Elevation). During the localization tasks and training sessions, participants wore a headband with a laser pointer and an electromagnetic motion sensor (METAMOTIONRL, MBIENTLAB INC, San Francisco, USA) attached. The laser light was reflected by the curtain and provided visual feedback for the participants to indicate perceived sound source directions. Real time head orientation and position captured by the motion sensor were used to calculate azimuth and elevation of the participants' response.

Stimuli
The stimuli used in the free-field localization task were 225 ms long sequences of pulsed pink noise, each composed of five equally spaced bursts of 25 ms duration. In the additional localization task, stimuli consisted of 225 ms long mixtures of environmental sounds. Each stimulus was composed of 6 randomly arranged excerpts of sounds drawn from a list of 42 recordings and had a unique spectrum. The Stimuli were re-generated on each trial and controlled by a custom python script using the slab toolbox [Schönwiesner et al., (2021)]. Overall sound pressure level (SPL) of the stimuli at the position of the participants' ears was 42 dB. Stimuli were processed digitally and amplified via TDT System 3 hardware (Tucker Davis Technologies, Alachua, USA). For every loudspeaker, the transfer function was measured by a probe microphone (Brüel Kjær, Nærum, Denmark) positioned at equal distance and orientation. A bank of inverse finite impulse response (FIR) filters was designed for each speaker to reduce differences in amplitude and frequency response across the loudspeakers. This was done to minimize spectral localization cues independent of participant's HRTFs.

Localization task
44 Loudspeakers were used for the localization task, covering 102° in azimuth (-52.5° to 52.5°) and 75° in elevation (-37.5° to 37-5°) of the hemisphere in front of the participant. Loudspeaker positions are described in an interaural-polar coordinate system. The loudspeakers were arranged on a spherical grid, formed by 7 loudspeakers on the horizontal and 7 loudspeakers on the vertical

axis. Loudspeakers were distributed on the sphere with an angular distance of 17.5° in azimuth and 12.5° in elevation between neighboring speakers [Fig. X]. The 4 outermost loudspeakers (at +/- 52.5° azimuth and +/- 37.5° elevation) were excluded. At the beginning of each trial, participants were instructed to aim the head mounted laser at a centrally presented LED while pressing the button on a handheld box. When the button was pressed, participant's initial head position was recorded and the stimulus was presented at a pseudorandom direction. Participants were instructed to indicate the perceived direction by turning their head towards the sound source and to confirm their response by pressing the button again. The horizontal and vertical angular displacement from the initial to the indicating head orientation was used to measure participant's responses. No feedback was given. Each direction was presented three times during a localization run, with an angular distance of at least 35° between sound locations of two consecutive trials to reduce adaptation and assimilation (Cross, 1973).

Training

The training task was designed to accelerate the adaptation to new spectral cues and was introduced to participants as a game-like scenario. Participants were instructed to find the location of a pulsed pink noise played from one of the 44 speakers. Proprioceptive feedback was provided by varying pulse duration and the delay between the pulses depending on participant's head orientation. Duration and delay of consecutive pulses decreased logarithmically with the angular distance between sound direction and the participant's head direction, from up to 500 ms at 65° angular distance. (the maximal distance (52.5° in azimuth and 37.5° in elevation). The pulse train gradually merged into a continuous sound when participants directed the head mounted laser within 3° angular distance to the sound source. If participants remained oriented at the target area for at least 500 ms, the sound source was considered found and a popular video game sound was played as a reward signal. The target sound location was then switched at least 45° away from the previous one. Additionally, participants scored points for every found location and were rewarded more points if they located the sounds faster. Participants tried to score as many points as possible within 90 seconds. The final score was displayed on a screen after each round and a leaderboard encouraged competition amongst participants. Throughout the adaptation periods, participants underwent a daily routine of three 10-minute training sessions, intermitted by 5-minute breaks.

Directional transfer functions

Binaural recordings were conducted to extract directional transfer functions (DTFs) of participant's pinnae with and without silicone molds. PUM-3046L-R miniature microphones (PUIaudio, Fairborn, USA) were inserted in the ear canal to measure the sound pressure level at the ear eardrum. To minimize non-directional contributions by standing-wave pattern in the canal, the microphones were placed 2 mm into the entrance of the blocked ear canal. Participants were seated in front of the loudspeaker array and were asked to remain stationary during the measurement while aiming the head mounted laser at the central LED. Frequency modulated sweeps of 100 ms duration were presented 30 times from each of the 7 loudspeakers on the vertical midline (0° azimuth, -37.5° to 37.5° elevation). Repetitions were averaged for every location to increase signal to noise ratio. Recordings were digitized via TDT system 3 hardware at a sampling rate of 97 kHz. DTFs were extracted from the time-averaged recordings by taking the ratio of the Fourier transform of the acquired signal to the Fourier transform of the input signal. DTFs were smoothed in the cepstral domain by a 1500 Hz low-pass FIR filter.

Statistical analysis

Sound localization accuracy was quantified by the root mean square of the distance (RMSE) between the physical target and the perceived response locations for azimuth and elevation respectively. Horizontal and vertical variance of participant's responses was quantified by taking the grand mean across standard deviations (SD) of the response coordinates for each sound location. The participant's ability to perceive sound source elevation was additionally quantified by the elevation gain (EG), as the slope of the linear regression lines between target and response elevations (Hofman et al., 1998).

3 Results

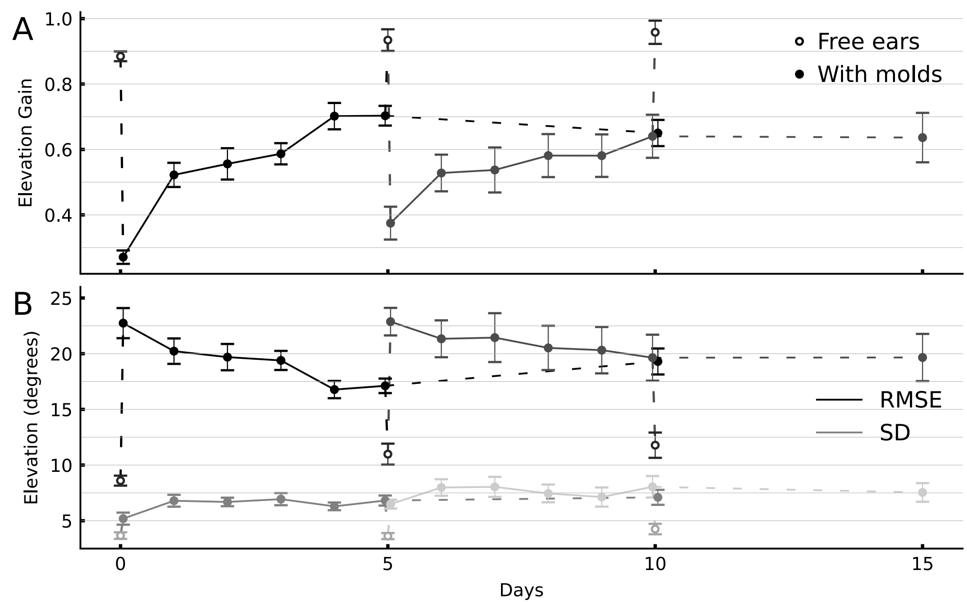
Earmolds reduced vertical localization performance

Insertion of earmolds degraded vertical localization performance (see Fig. 1, days 0 and 5; one-tailed Wilcoxon signed rank test of localization performance; ears free vs earmolds 1; EG: p = 3x10-5, RMSE: p = 3x10-5, SD: p = 6x10-3, ears free vs earmolds 2; EG: p = 5x10-4, RMSE: p = 5x10-4, SD: p = 5x10-4). On the horizontal plane, only the standard deviation of responses increased (one-tailed Wilcoxon signed rank test of performance; ears free vs earmolds 1; SD: p = 0.04, ears free vs earmolds 2: SD: p = 5x10-3). Both sets of earmolds caused a similar decrease of vertical localization performance (two-tailed Wilcoxon signed rank test, differences between free ears and earmolds 1 on day 0 vs difference between ears free and earmolds 2 on day 5; EG: p = 0.465, RMSE: p = 0.700, SD: p = 0.123).

Participants consecutively adapted to two novel sets of DTFs

Participants wore two different sets of earmolds during two consecutive 6-day adaptation periods. Adaptation was driven by multisensory experience while wearing the molds throughout the day, accompanied by five sessions of daily sensory-motor training at the lab. Vertical sound localization performance improved significantly for both sets of earmolds except for response variability (SD), which increased throughout the adaptation period (Fig 1, one-tailed Wilcoxon signed rank tests of first vs last day of molds; Earmolds 1: EG: p = 3x10-5, RMSE: p = 6x10-5, SD: p = 0.008; Earmolds 2: EG: p = 3x10-4, RMSE: p = 0.032, SD: p = 0.004). As expected, horizontal localization was not affected by adaptation to the earmolds (Fig 3 B, one-tailed Wilcoxon signed rank tests of first vs last day of molds; Earmolds 1; RMSE: p = 0.555, SD: p = 0.467; Earmolds 2; RMSE: p = 0.485, SD: p = 0.515).

Figure 1: A test figure with its caption side by side



No difference between rates of adaptation to first and second earmolds

To investigate possible effects of metaplasticity (i.e., an effect of the previous adaptation to earmolds 1 on the subsequent adaptation to earmolds 2) the rates of adaptation to the first and second earmolds were compared. To assess the rate of adaptation across participants independent of initial acoustical disruption caused by the earmolds, the increase in vertical localization accuracy during adaptation (performance on day 0 vs day 6) was divided by the initial decrease caused by the molds. Individual adaptation rates varied continuously and did not fall into discernable groups. No difference was found between the two sets of earmolds (two-tailed Wilcoxon signed rank test of Molds 1 vs Molds 2, Reduction of vertical RMSE from day 0 to day 6 divided by initial increase; Earmolds 1: 0.66 +/- 0.06 vs Earmolds 2: 0.74 +/- 0.09, p = 0.413). Individual adaptation performance with the first set of earmolds was positively related to adaptation performance with the second set of molds, although not significant (R = 0.47, p = 0.142). No correlation was found between vertical spectral information of the earmolds in the 3.7 – 12.9 kHz band and adaptation performance.

Earmolds reduced spectral information available for vertical sound localization

Application of silicone molds to the Pinnae altered spectral cues across the measured 4-16 kHz band (Fig. 2 A - C). Both sets of earmolds had a similar effect in reducing the prominent spectral notch situated in the 5-12 kHz band and the spectral peaks to its flanks. Similar effects have been observed in previous studies using earmolds to modify spectral cues (Trapeau, Schönwiesner 2015, Carlile 2014, Hofman 1998). Comparing the VSI of modified ears to participants' free ears showed a reduction of vertical spectral information (VSI) available for elevation discrimination in the 5.3-11.7 kHz octave band (one-tailed Wilcoxon signed rank test; VSI free ears: 0.61 +/- 0.04, VSI earmolds 1: 0.45 +/- 0.05, p = 0.003, VSI earmolds 2: 0.39 +/- 0.04, p = 0.004). As expected, the VSI of the left and right ear without molds were correlated in this band (Spearman correlation of left ear and right ear VSI; R = 0.55, p = 0.035). This correlation persisted after the earmolds were applied but was not significant for the second set of molds (earmolds 1: R = 0.56, p = 0.035, earmolds 2: R = 0.53, p = 0.064). Figure 4 shows the VSI of all participants with free ears and both sets of earmolds.

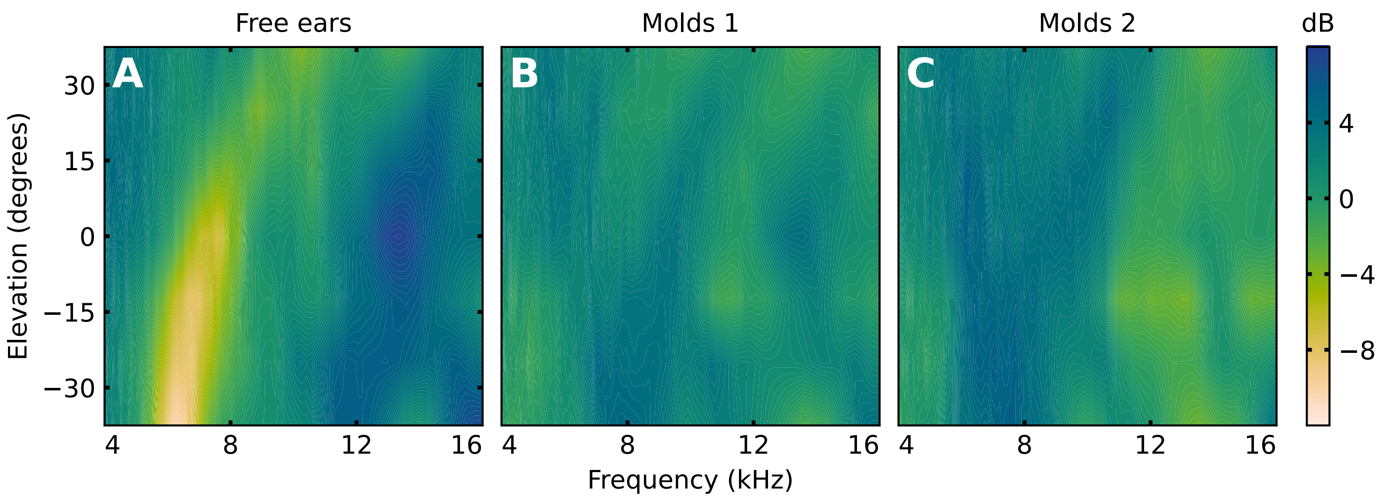


Figure 2: my caption of the figure, and here is a link **??**

Free ears localization accuracy was not explained by VSI or spectral strength
Vertical spectral information and spectral strength was computed in 5 octave bands between 4 and 16 kHz (4–8 kHz, 4.8–9.5 kHz, 5.7–11.3 kHz, 6.7–13.5kHz, 8–16kHz) from sets of DTFs obtained with and without molds. As previously reported by **?**, VSIs of participants' free ears varied among frequency bands (Kruskal-Wallis test, p = 8x10-3) and peaked in the 5.7 – 11.3 kHz band [Fig 5]. When comparing VSIs in this band with participants' initial vertical localization accuracy, only vertical SD was significantly correlated (Spearman correlation of free ears VSI and vertical SD: R = - 0.55, p = 0.031). No correlation was found between spectral strength and behavioral metrics in this band. **?** reported a systematic variation in the frequencies of spectral features among individuals' DTFs within the 3.7 – 12.9 kHz band. To account for the inter-individual differences in the distribution of spectral information across octave bands in the present data, behavioral measures were additionally compared to spectral features in this range of frequencies. Here the relation between participants' free ears VSI and vertical localization accuracy was in the predicted direction, although not significant (Fig 5 A, Spearman correlation of free ears VSI and vertical RMSE: R = - 0.39, p = 0.151).

Modified pinna shapes were within the physiological range
To confirm that spectral changes induced by the earmolds were physiologically plausible, the VSI dissimilarities in the 5.7-11.3 kHz band between the free and modified ears of each participant were compared to VSI dissimilarities between all possible pairs of participants' free ears [Fig 4 B]. The overlap of distributions shows that spectral changes induced by both sets of molds were comparable in magnitude to the natural spectrum of differences between individuals' ears.

Differences between first and second mold VSI dissimilarity to free ears
Silicone molds were fitted with the aim of achieving consistent reduction of vertical localization performance across subjects and earmolds. To achieve this, stronger alterations in morphological features of the outer ears became necessary for the second set of molds compared to the first set. Consequently, differences between DTFs with molds and participants' free ears were larger for the second set of earmolds (Wilcoxon one tailed signed rank test, VSI dissimilarity in the 3.7 – 12.9 kHz band, ears free and earmolds 1: 0.56 +/- 0.06 vs ears free and earmolds 2: 0.72 +/- 0.07, p = 0.002).

Spectral changes occurred at similar frequencies across participants and were different for both sets of earmolds
To confirm whether spectral changes occurred at similar frequencies across participants yet were at different frequencies for the two consecutive sets of earmolds, the probability of spectral change between free ears and earmolds was mapped for each frequency bin and elevation (Fig. 3 A-C). Spectral changes were defined as the absolute differences between DTFs measured before and after mold insertion above a given threshold. This threshold was a participant-specific measure of spectral difference across DTFs with free ears and was defined by the mean RMS difference across all combinations of DTFs (in dB) at each elevation (average across participants: 4.89 dB +/-0.15). Based on these thresholds, binary maps of spectral changes were created for each set of earmolds and participant (above-threshold changes were set to 1, all other values were set to 0). The average of these maps across participants shows the proportion of participants for which earmolds induced spectral changes above the threshold at each frequency bin and elevation.

Effects of acoustic dissimilarity between free ears and ear-

molds on localization accuracy

To investigate whether acoustic and behavioral effects of the earmolds were related, the VSI dissimilarity between DTFs in the 3.7 – 12.9 kHz band with and without molds was compared to the decrease in participant's localization performance after insertion of the earmolds. A trend of increasing vertical RMSE for larger acoustic differences was found for the first set of earmolds (Fig 5 B, Spearman correlation of vertical RMSE in the first test with molds compared to free ears and VSI dissimilarity: $R = 0.38$, $p = 0.175$). The relation between acoustic differences and behavior was still visible on the last day of the adaptation period (Fig 5 C, Spearman correlation of vertical RMSE in the last test with molds compared to free ears baseline and VSI dissimilarity: $R = 0.38$, $p = 0.185$). No such trend was found for vertical localization and VSI dissimilarity between free ears and the second set of molds (Spearman correlation of vertical RMSE in the first test with molds compared to free ears baseline and VSI dissimilarity: $R = - 0.07$, $p = 0.832$). Because initial vertical localization accuracy with the second earmolds could additionally depend on acoustic similarities to the previously learned set, differences in localization performance between the final test with earmolds 1 and the initial test with the earmolds 2 were compared to the VSI dissimilarity between both molds. A trend was found of increasing vertical error with greater acoustic dissimilarity between the first and second set of earmolds (Fig 5 D, Spearman correlation of vertical RMSE in the first test with second molds compared to last test with first molds and VSI dissimilarity: $R = 0.29$, $p = 0.385$).

Adaptation was generalizable

To rule out the possibility of participants memorizing location-specific spectral features of the training stimuli in the localization test, the test was repeated with a subset of six participants on the last day of each adaptation period using stimuli of random spectral content (USOs). The effect of USO stimuli on vertical localization error did not differ between adapted earmolds and free ears indicating that generalizable perceptual learning had taken place (Friedman test; differences in vertical RMSE between pink noise and USO localization across conditions; Ears free: -0.38 +/- 1.82, Earmolds 1: 2.5 +/- 0.46, Earmolds 2: -0.15 +/- 0.45, $p = 0.135$).

No aftereffect on free ears localization performance after mold removal

Previous studies reported the absence of an aftereffect on localization performance with free ears after adaptation to new spectral cues for sound localization (Hofman 1998, Trapeu and Schönwiesner 2015). To confirm these findings, free ears localization accuracy was measured immediately after mold removal at the end of each adaptation period. No aftereffect was present for elevation gain but. An increasing impact on participants' vertical localization accuracy with their native ears was observed after each adaptation period (see Fig. 1 B, one-tailed Wilcoxon signed rank test, vertical RMSE; free ears baseline vs free ears day 5: $p = 0.002$, free ears baseline vs free ears day 10: $p = 6x10-4$).

4 References

References

P. M. Hofman. Relearning sound localization with new ears. *nature neuroscience*, 1(5), 1998.

J. C. Middlebrooks. Individual differences in external-ear transfer functions reduced by scaling in frequency. *The Journal of the Acoustical Society of America*, 106(3):1480–1492, Sept. 1999. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.427176. URL https://pubs.aip.org/jasa/article/106/3/1480/553209/Individual-differences-in-external-ear-transfer.

R. Trapeau, V. Aubrais, and M. Schönwiesner. Fast and persistent adaptation to new spectral cues for sound localization suggests a many-to-one mapping mechanism. *The Journal of the Acoustical Society of America*, 140(2):879–890, Aug. 2016. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.4960568. URL https://pubs.aip.org/jasa/article/140/2/879/663968/Fast-and-persistent-adaptation-to-new-spectral.