

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

Paul Friedrich  
Department of Biology, Leipzig University

February 2023

## I. 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 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 (Wightman and Kistler (1989), Blauert (1996)). As these filters underlie lifelong changes beyond the developmental period (Otte et al. (2013), Clifton et al. (1988)), the auditory system must retain its ability to recalibrate the mapping of spectral cues to locations in space. Hofman (1998) demonstrated such a recalibration experimentally by inserting molds in the concha of adult listeners. Localization accuracy was initially impaired by the modified spectral cues and recovered within a few weeks as participants adapted to the molds. Interestingly, listeners localization accuracy with their native ears remains unchanged once the molds are removed, indicating that accommodation to new spectral cues does not affect the representation of the original cues. It has thus been suggested that adapting to a new pinna shape resembles the acquisition of a new language, in that different sets of pinna filters are stored in parallel and without much interference. In this view, both learned representations are always activated by sounds and the auditory system selects the correct spectral-to-spatial mapping based on the current pinna filter. However, Trapeau et al. (2016) showed that after adaptation, listeners were able to accurately localize sounds with their native ears at the first trial, without previous knowledge of the stimulus spectrum, tactile information about the absence of the molds or visual feedback. A lack of even short-lived aftereffects could be explained by a different mapping mechanism, in which several spectral profiles become associated with one spatial location and no selection between discrete maps would be required (Trapeau et al., 2016). One hope in pursuing these questions is that a deeper understanding of the neural mechanisms underlying this plasticity could uncover principles which can be applied more generally. So far, studies investigating the adaptation process have focused on the effects of learning one additional set of spectral cues (see Carlile (2014) for a comprehensive overview). However, simultaneous adaptation to multiple pinna shapes may reveal limitations in the auditory systems' ability to encode distinct spectral-to-spatial maps without interference and thus help to determine the underlying mechanism. In the present study participants learned to localize sounds with two different sets of earmolds within two consecutive five-day adaptation periods. To establish both representations in quick succession, participants underwent daily sessions of sensory-motor training, which has been shown to accelerate the adaptation to modified pinnae (Carlile and Blackman (2014), Parsehian and Katz (2012), Trapeau et al. (2016)). This paradigm was chosen to test (1) whether the second adaptation benefits from the neural plasticity induced during previous cue relearning (i.e. occurrence of metaplasticity) and (2) if the previously observed absence of aftereffects on listeners original mappings could be confirmed for a recently learned set.

## II. METHODS

### A. Participants

15 participants (11- females and 4 males, mean age  $23 \pm 3$  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.

## B. Earmolds

To alter participants' perception of sound source elevation, their pinnae were consecutively modified by two sets of 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 cyma conchae, cavum conchae and the antihelix while keeping the ear canal unobstructed [Fig. X]. Shape and volume of the earmolds varied between individuals to achieve similar degrees of disruption of vertical sound localization across participants. Two participants did not complete the experiment (first earmolds: n = 14, second earmolds: n = 13) and were removed from the analysis for these molds.

## C. 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 molds 1 and molds 2), 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 was measured after participants adapted to the second pair of molds for five days. This measurement was repeated for the second set of molds after five days without earmolds (see Fig. 5). 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, 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. 5), participants completed a short training session and a localization test. As a control, a subset of 6 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 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. 5), the first set of earmolds was briefly re-inserted and participants completed a localization run. To compare the persistence of adaptation to both sets of molds, participants returned to the lab after 5 days for a final localization test with the second molds 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.

## D. 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 and Bialas, 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). To minimize spectral localization cues independent of participant's HRTFs, transfer functions for every loudspeaker were 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.

## E. 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 ([Ward, 1979](#)).

## F. 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 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  $45^\circ$  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 a sportive competition. Throughout the adaptation periods, participants underwent a daily routine of three 10-minute training sessions, intermitted by 5-minute breaks.

## G. 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 ( $0^\circ$  azimuth,  $-37.5^\circ$  to  $37.5^\circ$  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. To remove non-directional portions of the transfer functions (such as ear canal and microphone transfer functions) each measured transfer function was divided by the grand average across the DTFs of all participants with and without molds ( $n = 588$ ). To avoid over-representing higher frequencies, DTFs were processed with a bank of octave spaced cosine band-pass filters as proposed by [Middlebrooks \(1999\)](#). Center frequencies were equally spaced at 0.0286 octaves between 4 and 16 kHz resulting in 2% frequency difference between each of the 83 filters. The DTFs obtained from the left and right ear of one participant are shown in Fig. 1. Spectral information in participants' DTFs that could potentially guide vertical sound localization was quantified by measures introduced in previous studies. Vertical spectral information (VSI), VSI dissimilarity ([Trapeau et al., 2016](#)) and spectral strength ([Andéol et al., 2013](#)) were computed from DTFs obtained from the left and right ear. For comparison with behavioral measures left and right ear values were averaged.

## H. 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, 1998](#)). Throughout

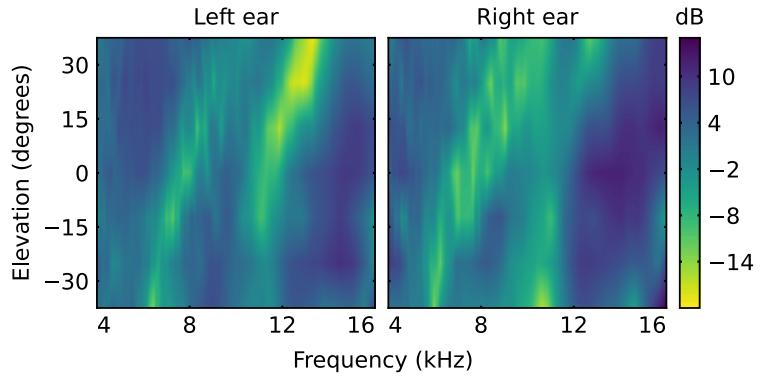


Figure 1: DTFs across 7 elevations on the vertical midline obtained from one participant. Color maps show the spectral amplitude computed at 83 frequency bins between 4 and 16 kHz. The spectral profile was linearly interpolated between the measured DTFs for display.

the analysis, paired comparisons were statistically assessed using Wilcoxon signed-rank tests. Relations between measures were determined using Spearman correlation coefficients, p-values smaller than 0.05 were regarded as significant.

### III. RESULTS

#### A. VSI and localization performance with free ears were related

Participants' directional transfer functions (DTFs) were recorded across the vertical midline with and without earmolds. To quantify spectral information available for vertical sound localization in the individual sets of DTFs, VSI (Trapeau et al., 2016) and spectral strength (Andéol et al., 2013) were computed in 5 octave bands between 4 and 16 kHz (4–8 kHz, 4.8–9.5 kHz, 5.7–11.3 kHz, 6.7–13.5 kHz, 8–16 kHz). Spectral information of participants' free ears varied among frequency bands (Kruskal-Wallis test, VSI:  $p = 0.0087$ , spectral strength:  $p = 10^{-7}$ ). The dissimilarity between DTFs across elevations (VSI) peaked in the 5.7–11.3 kHz band as previously reported by Trapeau et al. (2016) (Fig. 2 A). The VSIs of all participants in this band are shown in Fig. 2 B. Participants' left and right ears had similar VSI (Spearman correlation between the VSI of left and right ears in the 5.7–11.3 kHz band:  $r = 0.45, p = 0.0897$ ). Spectral strength, i.e., the variance within individual DTFs, indicated the highest spectral detail in the neighboring 6.7–13.5 kHz band. When joining these two bands for the analysis, VSI correlated with vertical localization accuracy (Fig. 2 C) and vertical SD (free ears VSI in the 5.7–13.5 kHz band compared to vertical localization performance: RMSE:  $R = -0.58, p = 0.0249$ , EG:  $R = 0.15, p = 0.6025$ , SD:  $R = -0.58, p = 0.0249$ ). No correlation was found between behavioral metrics and VSI in the other frequency bands. Spectral strength did not correlate with behavior in the tested frequency bands.

#### B. Earmolds reduced spectral information in the 5.7–11.3 kHz band

Application of silicone molds to the Pinnae altered spectral cues in the 4–16 kHz band. The consecutively applied sets of earmolds (further denoted as molds 1 and molds 2) attenuated the spectral notch

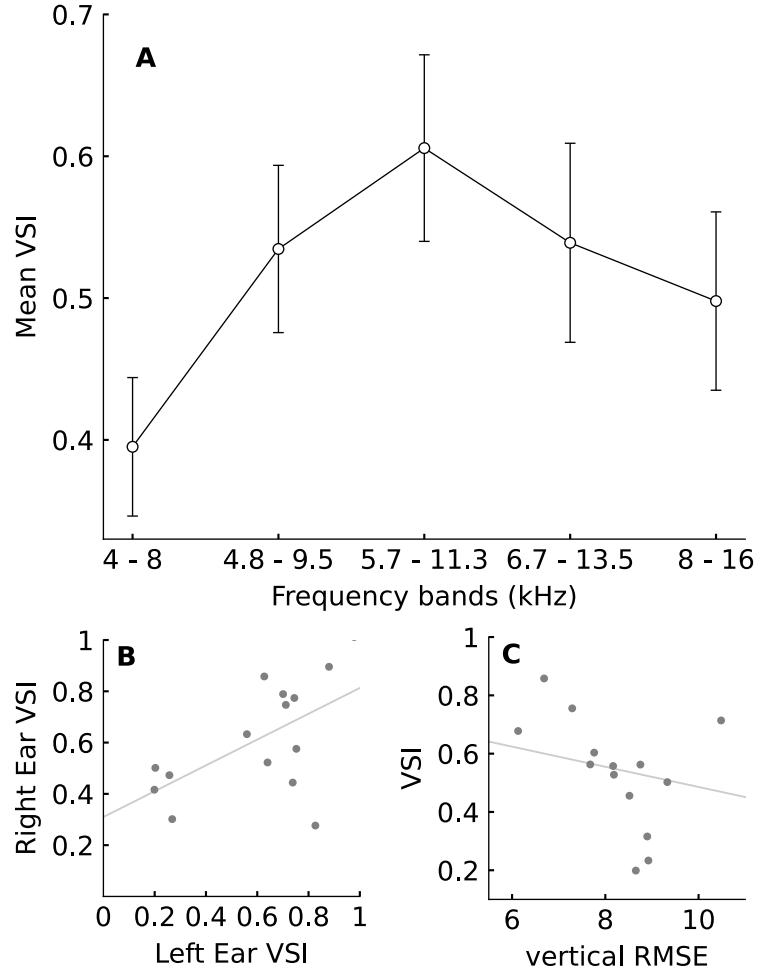


Figure 2: Free ears VSI and localization performance. **A**, Variation of VSI across five octave bands between 4 and 16 kHz. VSI was averaged across sets of DTFs obtained from participants' left and right ears without earmolds. VSI varies significantly and is highest in the 5.7–11.3 kHz band. **B**, Left and right ear VSI of each participant in the 5.7–11.3 kHz band. VSIs of the left and right ears were similar. **C**, Relation between VSI in the 5.7 – 13.5 kHz band and vertical localization accuracy.

situated in the 5–11 kHz band and the neighbouring spectral peak in the 11–14 kHz band (Fig. 3, A-C). Similar effects have been observed in previous studies using pinna modifications to alter spectral cues for sound localization (Trapeau et al. (2016), Wanrooij and Opstal (2005), Hofman (1998)). Comparing VSIs of participants’ modified and free ears in the 5.7–11.3 kHz band showed that earmolds reduced the amount of spectral information available for elevation discrimination in this band (Fig. 4 B, differences between VSI of free and modified ears; free ears:  $0.61 \pm 0.04$ , vs molds 1:  $0.45 \pm 0.05$ ,  $p = 0.0022$ , vs molds 2:  $0.39 \pm 0.04$ ,  $p = 0.0036$ ). The two sets of earmolds led to similar levels of VSI reduction compared to participants’ free ears (free ears VSI reduction caused by molds 1:  $0.18 \pm 0.06$  vs molds 2:  $0.18 \pm 0.06$ ,  $p = 0.1973$ ). The relation between VSIs of participants’ left and right ears persisted after mold insertion but was not significant for the second pair of earmolds (molds 1:  $r = 0.56$ ,  $p = 0.0336$ , molds 2:  $r = 0.53$ ,  $p = 0.064$ ). Fig. 4 A shows the VSI of all participants with modified and free ears. Wanrooij and Opstal (2005) reported the formation of new acoustic cues at higher frequencies after the insertion of silicone molds. To test whether volume reduction of pinna cavities caused by the silicone molds increased the frequencies of spectral features, spectral information was additionally compared in the 11.3–13.5 kHz band. The first set of earmolds increased VSI in this band compared to free ears (differences between VSI of free and modified ears in the 11.3–13.5 kHz band; free ears:  $0.5 \pm 0.06$ , vs molds 1:  $0.64 \pm 0.07$ ,  $p = 0.0387$ ).

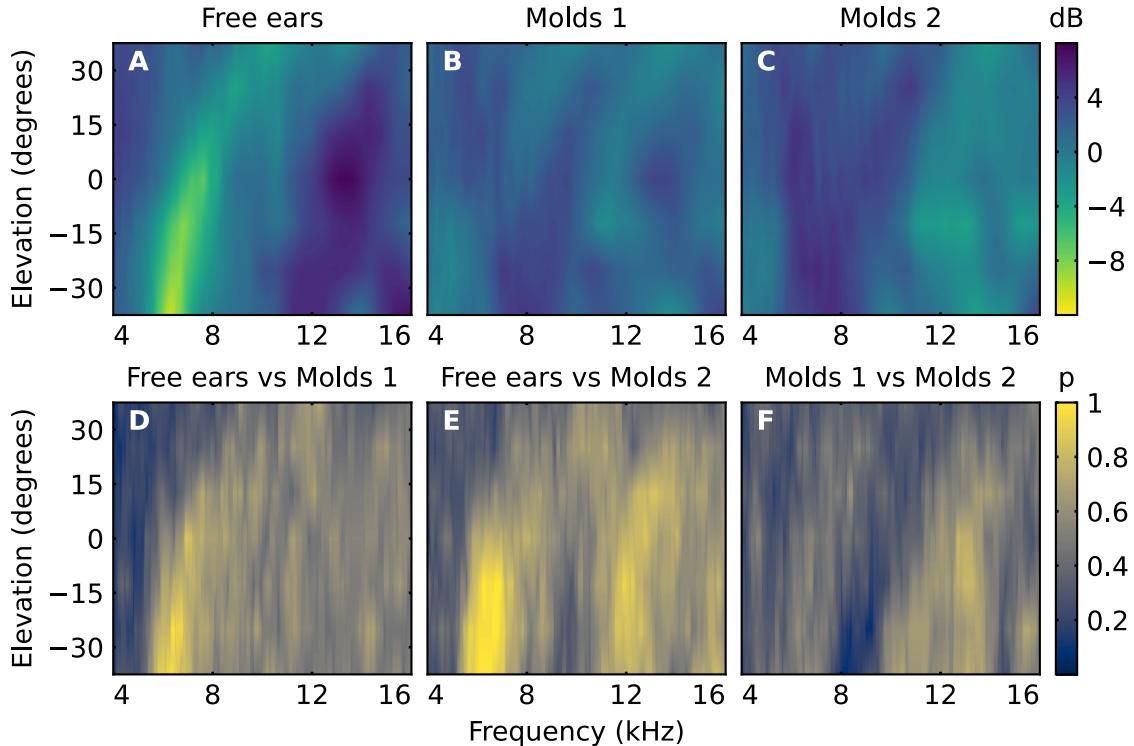


Figure 3: Acoustic effect of the earmolds. **A**, Mean across participants’ DTFs with free ears. **B-C**, Mean across participants’ DTFs with modified ears. Earmolds attenuated spectral notches and peaks in the 5–14 kHz frequency range. **D-E**, Probability maps showing the proportion of participants for which the molds induced marked changes in spectral amplitude at each elevation and frequency bin. Acoustic changes induced by each set of molds occurred at similar frequencies across participants. **F**, Probability map showing the proportion of marked changes between the first and second pair of earmolds. Both sets induced acoustic changes at different frequencies.

### C. Mold induced acoustic changes occurred at similar frequencies

To confirm whether spectral changes induced by each set of earmolds were at similar frequencies across participants, the probability of spectral change caused by pinna modifications was mapped for

each frequency bin and elevation (Fig. 3 D-F). 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:  $4.89 \text{ dB} \pm 0.15$ ). Based on these thresholds, binary maps of spectral changes were created for each set of molds 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 molds induced spectral changes above the threshold at each frequency bin and elevation. The frequencies of marked changes were similar across participants for each set of earmolds, and differed between the two sets.

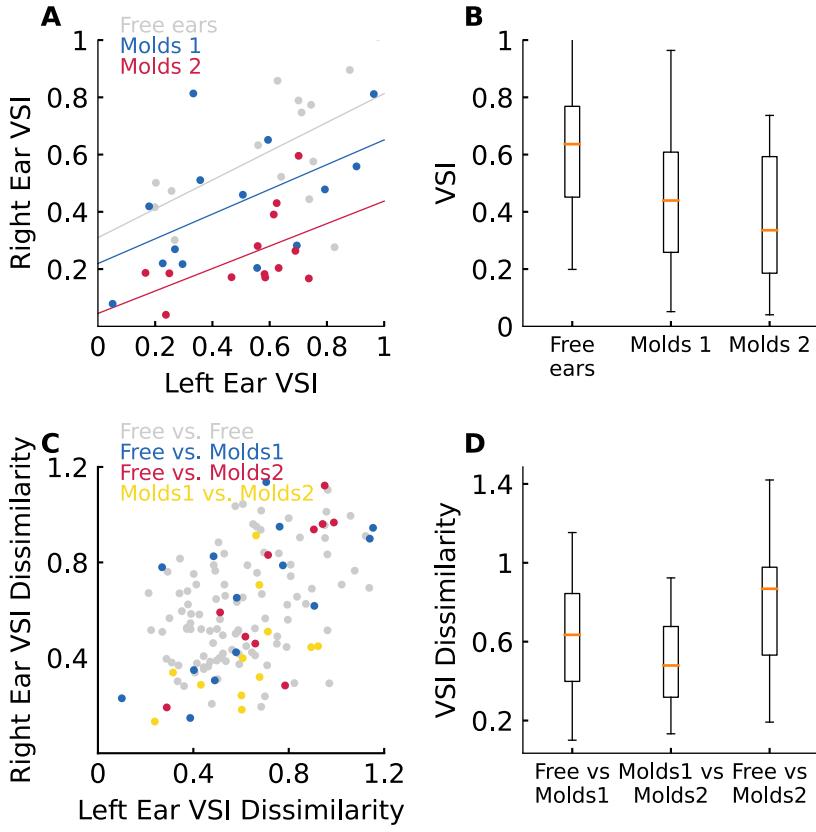


Figure 4: VSI and VSI dissimilarity across pinna shapes. **A**, Left and right ear VSI in the 5.7–11.3 kHz band. Each dot represents one participant. Grey dots show VSI values of free ears, blue and red dots represent VSI values with earmolds 1 and 2, respectively. Relations between left and right ear VSI are shown by regression lines. The similarity between left and right ear VSI persisted after mold insertion. **B**, VSI with and without molds in the 5.7–11.3 kHz band. Both sets of molds reduced VSI significantly. **C**, VSI dissimilarities across participants in the 5.7–13.5 kHz band. Grey dots show the VSI dissimilarity of free left and right ears between all combinations of two participants. Each colored dot shows the VSI dissimilarity between pinna shapes of one participant. The outline of the grey dot cloud largely includes the colored dots, indicating that acoustic differences induced by the molds did not exceed acoustic differences between participants' natural ears. **D**, VSI dissimilarities between the three pinna shapes in the 5.7–13.5 kHz band. VSI dissimilarity between free ears and the first earmolds was similar to VSI dissimilarity between the first and second set of molds.

## D. Proportional acoustic dissimilarities across consecutive pinna shapes

Each pair of earmolds modified the previously adapted set of DTFs. To test whether spectral differences that were consecutively induced by the molds were in the same range, VSI dissimilarities between the three different pinna shapes (ears free, earmolds 1 and earmolds 2) were computed in the 5.7–13.5 kHz frequency band. The dissimilarity between participants' free ears and the first set of earmolds was comparable to the dissimilarity between the first and the second set of molds (Fig. 4 D, VSI dissimilarity in the 5.7–13.5 kHz band, free ears and molds 1:  $0.64 \pm 0.06$  vs molds 1 and molds 2:  $0.51 \pm 0.05$ ;  $p = 0.2544$ ). The second set of earmolds induced larger spectral differences to participants' free ears than the first set (VSI dissimilarity free ears and molds 1:  $0.64 \pm 0.06$  vs free ears and molds 2:  $0.8 \pm 0.06$ ;  $p = 0.0005$ ). To confirm that spectral changes induced by the earmolds were physiologically plausible, VSI dissimilarities between free and modified ears of each participant were compared to VSI dissimilarities between all possible pairs of participants' free ears (Fig. 4 C). 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.

## E. Earmolds reduced vertical localization performance

Insertion of earmolds degraded vertical localization performance (see Fig. 5, days 0 and 5; vertical localization performance; ears free vs molds 1; RMSE:  $p = 3 \times 10^{-5}$ , EG:  $p = 3 \times 10^{-5}$ , SD:  $p = 0.0062$ , ears free vs molds 2; RMSE:  $p = 0.0005$ , EG:  $p = 0.0005$ , SD:  $p = 0.0005$ ). On the horizontal axis, only response variance was affected by the earmolds (horizontal localization performance; ears free vs molds 1; SD:  $p = 0.04$ , ears free vs molds 2: SD:  $p = 0.005$ ). Both sets of earmolds caused a similar decrease of vertical localization performance (mold induced drop in vertical localization performance, two-tailed Wilcoxon signed rank test; molds 1 vs molds 2; EG:  $p = 0.465$ , RMSE:  $p = 0.700$ , SD:  $p = 0.123$ ).

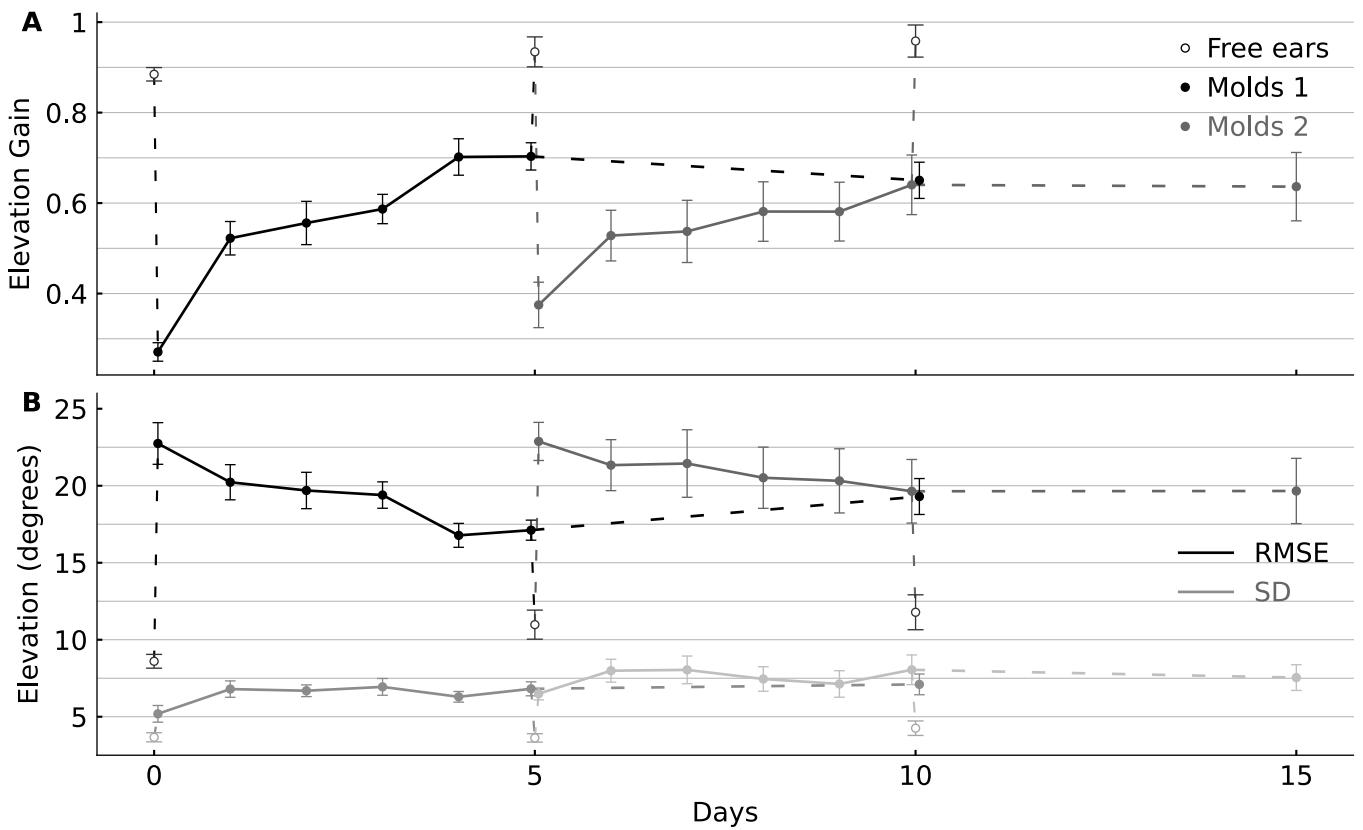


Figure 5: Time course of vertical localization performance. White dots indicate performance with free ears, black and grey dots indicate performance with molds 1 and molds 2, respectively. Data points represent the mean across participants (ears free:  $n = 15$ , molds 1:  $n = 14$ , molds 2:  $n = 13$ ). Error bars show the standard error. Slim dotted lines plot changes in localization performance after mold insertion and removal. Thick dotted lines show differences between the last test of the adaptation period and the final adaptation persistence test with each set of molds. **A**, Evolution of EG and **B**, localization error (RMSE) and response variability (SD) on the vertical plane throughout the two consecutive adaptation periods. Participants EG was reduced while RMSE and SD were increased after insertion of earmolds. EG and RMSE recovered within each of the two five-day adaptation periods, while SD further increased. Participants' EG recovered faster with the first set of molds than with the second set. After each mold removal, EG returned to baseline levels whereas RMSE increased compared to the first test with free ears (white dots on day 5 and 10). Although EG with molds 1 increased during adaptation to molds 2, adaptation persistence did not differ between the two sets of earmolds.

## F. Participants adapted to both pinna modifications

Participants wore the two different sets of earmolds during two consecutive five-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. 5, vertical localization performance on day 0 vs day 5 with earmolds; molds 1: EG:  $p = 3 \times 10^{-5}$ , RMSE:  $p = 6 \times 10^{-5}$ , SD:  $p = 0.0042$ ; molds 2: EG:  $p = 0.0002$ , RMSE:  $p = 0.032$ , SD:  $p = 0.0081$ ). As expected, horizontal localization was not affected by adaptation to the earmolds (horizontal localization on day 0 vs day 5 with earmolds; molds 1: RMSE:  $p = 0.555$ , SD:  $p = 0.467$ , molds 2: RMSE:  $p = 0.485$ , SD:  $p = 0.515$ ). 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; free ears:  $-0.38 \pm 1.82$ , molds 1:  $2.5 \pm 0.46$ , molds 2:  $-0.15 \pm 0.45$ ,  $p = 0.135$ ).

## G. Adaptation rate did not increase with the second set of earmolds

To investigate effects of metaplasticity, i.e., a faster relearning of sound localization with new pinna shapes after previously adapting to a different set of pinna modifications, adaptation was compared between the first and the second set of molds. To quantify the rate of adaptation across participants and earmolds independent of initial acoustic disruption caused by the molds, the gain in localization performance during adaptation (performance on day 0 vs day 5) was divided by the initial performance drop on day 0. Individual rates of adaptation for each set of molds varied continuously and did not fall into discernable groups. Participants' adaptation rates in EG were higher with the first set of molds than with the second set. No differences between earmold adaptation rates were found for vertical RMSE (performance gain from day 0 to day 5 divided by initial performance drop; molds 1 vs molds 2; RMSE:  $0.34 \pm 0.06$  vs  $0.26 \pm 0.09$ ,  $p = 0.2065$ , EG:  $0.7 \pm 0.04$  vs  $0.45 \pm 0.07$ ,  $p = 0.0122$ ). Individual adaptation rates with the first and second set of molds were positively related, although not significant (Spearman correlation coefficient:  $r = 0.47$ ,  $p = 0.142$ ).

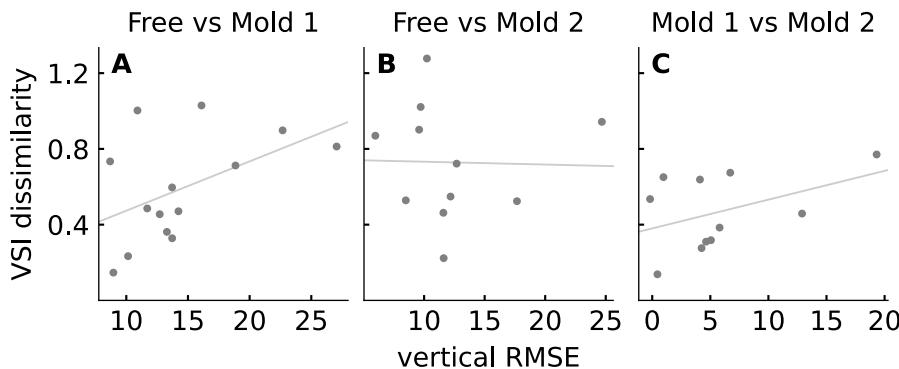


Figure 6: Acoustic dissimilarity and localization performance. Relation between VSI dissimilarity and vertical RMSE after mold insertion; **A**, with molds 1 compared to free ears, **B**, with molds 2 compared to free ears, **C**, with molds 2 compared to molds 1. Behavior with the second set of earmolds was better explained by acoustic dissimilarity to the previously adapted earmolds than to participants' native ears.

found for the first set of earmolds (Fig. 6 A, Spearman correlation of VSI dissimilarity between free ears and molds 1 and increase in vertical RMSE after mold 1 insertion:  $r = 0.43$ ,  $p = 0.122$ ). No such

## H. Effects of VSI dissimilarity on the disruption of localization performance

To test whether acoustic and behavioral effects of the earmolds were related, VSI dissimilarity between DTFs in the 5.7 – 13.5 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

trend was found for vertical localization performance and VSI dissimilarity between free ears and the second set of molds (Fig. 6 B, Spearman correlation of vertical RMSE in the first test with molds 1 compared to free ears baseline and VSI dissimilarity:  $r = -0.18, p = 0.5926$ ). Because initial vertical localization accuracy with the second earmolds could depend on acoustic similarities to the previously learned set, differences in localization performance on the last day of adaptation to earmolds 1 and the initial test with the earmolds 2 were compared to the VSI dissimilarity between both sets of molds. A trend of increasing vertical error with greater acoustic dissimilarity between the first and second set of earmolds was found (Fig. 6 C, Spearman correlation of increase in vertical RMSE from the last test with molds 1 to the first test with molds 2 and VSI dissimilarity between consecutive earmolds:  $r = 0.35, p = 0.2847$ ).

## I. Free ears localization accuracy decreased during adaptation

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), Trapeau et al. (2016), Carlile and Blackman (2014), Wanrooij and Opstal (2005)). To confirm these findings, free ears localization accuracy was measured immediately after earmolds were removed at the end of each adaptation period. No aftereffect was found for EG but an increasing impact on participants' vertical localization accuracy with their native ears was observed after each adaptation period (Fig. 5 B, vertical RMSE; free ears baseline vs free ears day 5:  $p = 0.002$ , free ears baseline vs free ears day 10:  $p = 0.0006$ ).

## J. Vertical localization accuracy with the first earmolds decreased during adaptation to the second molds

To investigate whether learning a new set of spectral cues for sound localization interfered with a previously learned set, earmolds were reinserted five days after mold removal and localization was re-measured. During the five days after the first earmolds were removed, participants adapted to another set of pinna modifications. The second set of earmolds served as a control because participants were not exposed to new spectral cues in the five days following mold 2 removal (see Fig. 5). Participants vertical localization accuracy with molds 1 was significantly decreased after five days of adaptation to the second molds (RMSE with molds 1 on day 5:  $17.12 \pm 0.65$  and day 10:  $19.3 \pm 1.17, p = 0.015$ ). Other metrics of localization performance did not differ significantly (performance with molds 1 on day 5 vs day 10; EG:  $p = 0.1514$ , SD:  $p = 0.2997$ ). Localization performance with the second molds remained unchanged after 5 days with free ears (performance with molds 2 on day 10 vs day 15; RMSE:  $19.64 \pm 0.91$  vs  $19.66 \pm 1.08$ , EG:  $0.64 \pm 0.03$  vs  $0.64 \pm 0.06$ , SD:  $8.05 \pm 0.79$  vs  $7.54 \pm 0.54$ ). Adaptation persistence was defined as the difference between the last test with molds in the adaptation period and the final persistence tests with molds 1 and 2 on days 10 and 15, respectively. No difference in adaptation persistence was found between the first and second set of molds (adaptation persistence with molds 1 vs molds 2; RMSE:  $p = 0.2119$ , EG:  $p = 0.2349$ , SD:  $p = 0.1902$ ). Fig. 7 shows the evolution of response pattern with both sets of earmolds averaged across participants throughout the time course of the experiment.

Adaptation to modified cues affected Localization accuracy with previous cues set decreased during adaptation

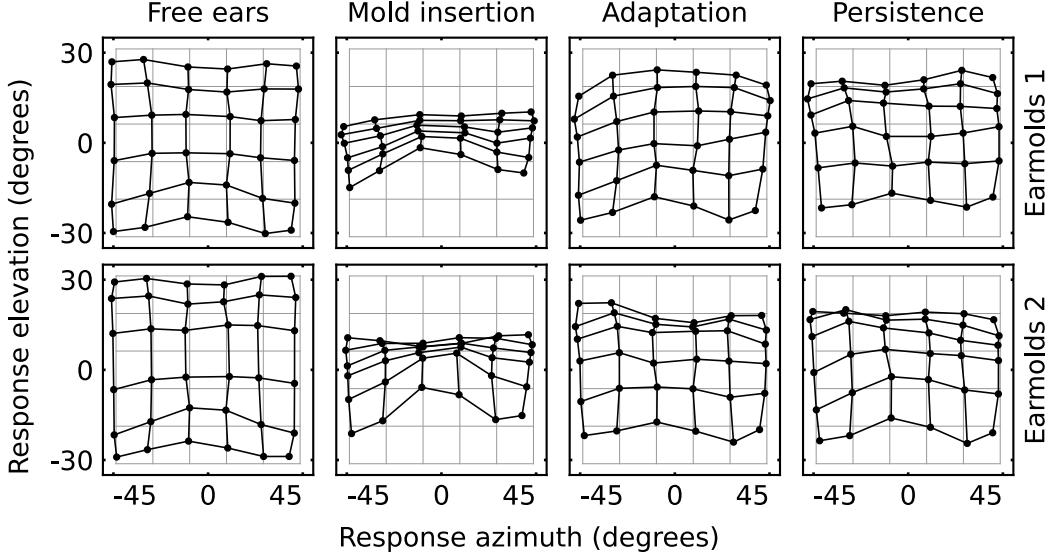


Figure 7: Change of response pattern throughout the experiment. Grey panels show sound target locations, averaged for targets belonging to similar directions by dividing the target space into twenty-five half-overlapping sectors (Hofman, 1998). Black dots represent sound localization responses for each target averaged across all participants. The first and second row show behavioral impact, subsequent adaptation and adaptation persistence with molds 1 and 2, respectively.

## IV. DISCUSSION

Listeners were successively adapted to distinct sets of spectral cues for sound localization to investigate effects of metaplasticity and interference between the learned spectral mappings. Modified pinna shapes were sufficiently different to cause repeated disruption of vertical localization. Participants relearned sound localization with both sets of silicone molds within consecutive five-day adaptation periods. Individual adaptation performance did not improve with the second set, suggesting an absence of meta-plasticity. Adaptation persistence did not differ between earmolds, showing that learning a second set of modified spectral cues does not degrade the previously learned representation. In contrast to previous findings, adaptation to new spectral cues reduced participants vertical accuracy with free ears.

### A. Acoustic comparison of pinna modifications

Application of silicone molds modified participants DTFs in a wide range of frequencies (see Fig. 3) and resulted in large behavioral effects. This study aimed to compare the effects of learning multiple sets of spectral cues on adaptation rate and persistence. To enable this comparison, the magnitude of acoustic changes caused by the two sets of earmolds should be similar. It has been shown that larger acoustic differences between free and modified ears caused a stronger decrease in vertical localization performance (Wanrooij and Opstal, 2005) and reduced the rate of subsequent adaptation to the new spectral cues (Trapeau et al., 2016). To quantify behaviorally relevant changes induced by the consecutive pinna modifications, the frequencies of salient spectral features were initially identified. Previous findings show that spectral cues which may guide vertical sound localization are not equally distributed across frequencies. Instead, distinct features such as notches and peaks that vary with elevation tend to be situated in specific frequency bands (Langendijk and Bronkhorst (2002), Trapeau et al. (2016)). In the cat, these features are extracted by edge detecting neurons of the DCN which are thought to be associated with spectral cue processing (Davis et al. (2003a), Reiss and Young (2005)). Two previously introduced measures, VSI and spectral strength, were used to compare the amount of spectral information in participants' free ears across frequency bands. Both measures varied significantly between the five octave bands and were highest in the 5.7–11.3 kHz (VSI) and 6.7–13.5 kHz band (spectral strength),

reflecting the contribution of prominent acoustic features located in these bands (the spectral notch and its neighbouring peak between 5 and 14 kHz in figure Fig. 3 A). Middlebrooks (1999) reported a systematic variation in the frequencies of spectral features among individuals, which is indicated by the broad peak and large error bars of VSI across frequency bands in Fig. 2 A. VSI between 5.7 and 13.5 kHz varied between individuals and was positively correlated with individual localization accuracy on the vertical axis, suggesting that acoustic features in this band largely contributed to the estimation of sound elevation. In accordance with previous studies, silicone molds attenuated the spectral notch between 5.7 and 11.3 kHz (Hofman (1998), Wanrooij and Opstal (2005), Trapeau et al. (2016)), indicated by a reduction of VSI in this band after mold insertion. This reduction did not differ significantly between the first and the second set of earmolds (see Fig. 4 B). The volumes of individual silicone molds were adjusted to participants concha volumes aiming to induce consistent disruption of elevation perception (EG) across individuals. Compared to the first earmolds, the second set required noticeably larger volumes to achieve similar levels of EG reduction, indicated by the increased VSI dissimilarity between participants' unmodified ears and the second molds in Fig. 4 D. Despite these variations in volume, each set caused acoustic changes at similar frequencies across individuals (see Fig. 3 D, E). Intuitively, smaller concha volumes interact with shorter wavelengths, which is reflected by an upward shift in the frequencies of marked changes induced by the second set. This effect might also result in the formation of new spectral cues at higher frequencies (Wanrooij and Opstal, 2005), and could be shown for the first set of earmolds, which increased VSI between 11.3 and 13.5 kHz compared to unmodified ears. The experiment was designed to answer two questions: (1) whether learning a set of spectral cues facilitates subsequent adaptation to a second set and (2), if learning a second set interferes with a previously learned representation. VSI dissimilarity was negatively related to adaptation performance in a previous study using modified pinnae (Trapeau et al., 2016), and was therefore taken into account when interpreting behavioral results. In this view, acoustic similarities between both sets of molds could facilitate adaptation to the second set, because little cue reweighing would be required to accommodate to the new spectral mapping. Moreover, learning a second set of DTFs with very similar spectral cues could enhance the persistence of the previously learned spectral mapping and counter possible effects of interference due to the repeated cue relearning. However, VSI dissimilarities did not differ between successive pinna shapes (Fig. 4 D), suggesting that hypothesis testing was unlikely to be confounded by disproportionate acoustic differences.

## B. Adaptation to successive earmolds

Insertion of earmolds significantly reduced vertical localization performance. This effect varied across individuals and between successive molds and was not fully explained by acoustic dissimilarities between pinna shapes (Fig. 6). Previous studies suggested that individual performance might also depend on non-acoustic factors such as attention, perceptual abilities and neural processes underlying spectral feature analysis. Differences in initial localization with the first and second molds were more pronounced for sounds located at lower elevations (Fig. 7). Apart from the outer ear, the head and torso also interact with incoming sound waves and provide coarse elevation cues for sources below the horizontal midline. Spectral cues created by the torso are situated at frequencies below 3 kHz (Asano et al. (1990), Algazi et al. (2001)), and were unlikely to be affected by the earmolds. Participants might have learned to partially rely on these cues during the first adaptation, causing the observed differences in performance reduction between both sets. Localization performance improved to various degrees throughout the five day adaptation periods. Individual differences varied continuously from minor adaptation to full recovery. Previous studies reported strong inter-individual variation in adaptation to spectral cues (Hofman (1998), Wanrooij and Opstal (2005), Trapeau et al. (2016)). Elevation perception recovered more quickly with the first set of earmolds than with the second set, likely due to the larger acoustic changes these molds induced to participants native ears. Adaptation rate did not increase with repeated cue

relearning within the time span of this experiment. Longer periods of cue relearning might be necessary to induce the physiological changes associated with metaplasticity. Earlier studies found no aftereffects on localization with free ears when the adapted earmolds were removed (Carlile and Blackman (2014), Hofman (1998), Trapeau et al. (2016), Wanrooij and Opstal (2005)), indicating that learned spectral-to-spatial mappings do not override pre-existing maps. In agreement with these results, adaptation persistence did not differ significantly between the successive earmolds, suggesting that adaptation to the second set of molds did not interfere with the previously learned representation. Vertical accuracy (RMSE) with the first molds was decreased after learning a second set of spectral cues, and a similar decrease in accuracy was observed for individuals' performance with free ears. However, free ears accuracy simultaneously decreased on the horizontal plane. A loss of general localization accuracy throughout the experiment could be explained by procedural factors such as participants trying to quickly complete the localization task. A model proposed by Hofman and Van Opstal (1998) posits that sound source elevation is estimated by comparing the spectrum of incoming sounds with spectral templates associated to different locations. After the second adaptation period, participants were able to access three different sets of spectral cues for sound localization, which required the underlying neural populations to accommodate two additional spectral mappings in quick succession and with minimal interference. This raises the question at which stage of the auditory pathway the learned spectral templates are represented. The required neural plasticity might arise in subcortical structures. DCN neurons and their projections to the central nucleus of the ICC have been shown to encode spectral features in the cat (Davis et al., 2003a) and evidence exists for a topographical representation of auditory space in the SCC of the ferret (King and Hutchings, 1987). However, the ability to learn multiple sets of spectral cues without interference indicate that spectral maps might be represented at the level of the auditory cortex. Recent studies found tuning curves in low-level AC encoding sound elevation, which were flattened after pinna modification and recovered their original shape as participants adapted to modified cues and regained elevation perception (Trapeau and Schönwiesner, 2018). Learning rates and adaptation persistence exceeded those observed in other studies using similar training paradigms (Parseihian and Katz (2012), Trapeau et al. (2016)), which may result from the gamified nature and increased duration of the daily training sessions. Studies on ferrets could show that the tuning of spectrotemporal receptive fields in the auditory cortex was driven by stimuli with behavioral relevance and that training induced changes could be preferentially selected depending on the behavioral context. To further investigate mechanisms underlying spectral cue relearning, future research could use acoustically controlled virtual environments. Training in virtual reality can significantly increase listeners sound localization performance with previously unknown DTFs, without the necessity of permanent exposure to physical pinna modifications (Honda et al. (2007), Steadman et al. (2019a)). Importantly, virtual DTFs allow spectral cues at amplitudes and frequencies that are not feasible when using earmolds and enable the precise manipulation of acoustic features.

## References

- V. R. Algazi, C. Avendano, and R. O. Duda. Elevation localization and head-related transfer function analysis at low frequencies. *The Journal of the Acoustical Society of America*, 109(3):1110–1122, Mar. 2001. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.1349185. URL <https://pubs.aip.org/jasa/article/109/3/1110/553520/Elevation-localization-and-head-related-transfer>.
- G. Andéol, E. A. Macpherson, and A. T. Sabin. Sound localization in noise and sensitivity to spectral shape. *Hearing Research*, 304:20–27, Oct. 2013. ISSN 03785955. doi: 10.1016/j.heares.2013.06.001. URL <https://linkinghub.elsevier.com/retrieve/pii/S0378595513001445>.

- F. Asano, Y. Suzuki, and T. Sone. Role of spectral cues in median plane localization. *The Journal of the Acoustical Society of America*, 88(1):159–168, July 1990. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.399963. URL <https://pubs.aip.org/jasa/article/88/1/159/626776/Role-of-spectral-cues-in-median-plane>.
- J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. The MIT Press, Oct. 1996. ISBN 978-0-262-26868-4. doi: 10.7551/mitpress/6391.001.0001. URL <https://doi.org/10.7551/mitpress/6391.001.0001>.
- S. Carlile. The plastic ear and perceptual relearning in auditory spatial perception. *Frontiers in Neuroscience*, 8, Aug. 2014. ISSN 1662-453X. doi: 10.3389/fnins.2014.00237. URL <http://journal.frontiersin.org/article/10.3389/fnins.2014.00237/abstract>.
- S. Carlile and T. Blackman. Relearning Auditory Spectral Cues for Locations Inside and Outside the Visual Field. *Journal of the Association for Research in Otolaryngology*, 15(2):249–263, Apr. 2014. ISSN 1525-3961, 1438-7573. doi: 10.1007/s10162-013-0429-5. URL <http://link.springer.com/10.1007/s10162-013-0429-5>.
- R. K. Clifton, M. G. Clarkson, J. Gwiazda, and J. A. Bauer. Growth in Head Size During Infancy: Implications for Sound Localization. *Developmental Psychology*, 24(4):477483, 1988. doi: 0012-1649/88/\$00.75.
- K. A. Davis, R. Ramachandran, and B. J. May. Auditory Processing of Spectral Cues for Sound Localization in the Inferior Colliculus. *JARO - Journal of the Association for Research in Otolaryngology*, 4(2):148–163, June 2003a. ISSN 1525-3961, 1438-7573. doi: 10.1007/s10162-002-2002-5. URL <http://link.springer.com/10.1007/s10162-002-2002-5>.
- K. A. Davis, R. Ramachandran, and B. J. May. Auditory Processing of Spectral Cues for Sound Localization in the Inferior Colliculus. *JARO - Journal of the Association for Research in Otolaryngology*, 4(2):148–163, June 2003b. ISSN 1525-3961, 1438-7573. doi: 10.1007/s10162-002-2002-5. URL <http://link.springer.com/10.1007/s10162-002-2002-5>.
- P. M. Hofman. Relearning sound localization with new ears. *nature neuroscience*, 1(5), 1998.
- P. M. Hofman and A. J. Van Opstal. Spectro-temporal factors in two-dimensional human sound localization. *The Journal of the Acoustical Society of America*, 103(5):2634–2648, May 1998. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.422784. URL <https://pubs.aip.org/jasa/article/103/5/2634/557650/Spectro-temporal-factors-in-two-dimensional-human>.
- A. Honda, H. Shibata, J. Gyoba, K. Saitou, Y. Iwaya, and Y. Suzuki. Transfer effects on sound localization performances from playing a virtual three-dimensional auditory game. *Applied Acoustics*, 68(8):885–896, Aug. 2007. ISSN 0003682X. doi: 10.1016/j.apacoust.2006.08.007. URL <https://linkinghub.elsevier.com/retrieve/pii/S0003682X06001824>.
- A. J. King and M. E. Hutchings. Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret: a map of auditory space. *Journal of Neurophysiology*, 57(2):596–624, Feb. 1987. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.1987.57.2.596. URL <https://www.physiology.org/doi/10.1152/jn.1987.57.2.596>.
- E. H. A. Langendijk and A. W. Bronkhorst. Contribution of spectral cues to human sound localization. *The Journal of the Acoustical Society of America*, 112(4):1583–1596, Oct. 2002. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.1501901. URL <https://pubs.aip.org/jasa/article/112/4/1583/550090/Contribution-of-spectral-cues-to-human-sound>.

- 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. J. Otte, M. J. H. Agterberg, M. M. Van Wanrooij, A. F. M. Snik, and A. J. Van Opstal. Age-related Hearing Loss and Ear Morphology Affect Vertical but not Horizontal Sound-Localization Performance. *Journal of the Association for Research in Otolaryngology*, 14(2):261–273, Apr. 2013. ISSN 1525-3961, 1438-7573. doi: 10.1007/s10162-012-0367-7. URL <http://link.springer.com/10.1007/s10162-012-0367-7>.
- G. Parseihian and B. F. G. Katz. Rapid head-related transfer function adaptation using a virtual auditory environment. *The Journal of the Acoustical Society of America*, 131(4):2948–2957, Apr. 2012. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.3687448. URL <https://pubs.aip.org/jasa/article/131/4/2948/831116/Rapid-head-related-transfer-function-adaptation>.
- L. A. J. Reiss and E. D. Young. Spectral Edge Sensitivity in Neural Circuits of the Dorsal Cochlear Nucleus. *The Journal of Neuroscience*, 25(14):3680–3691, Apr. 2005. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.4963-04.2005. URL <https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.4963-04.2005>.
- M. Schönwiesner and O. Bialas. s(ound)lab: An easy to learn Python package for designing and running psychoacoustic experiments. *Journal of Open Source Software*, 6(62):3284, June 2021. ISSN 2475-9066. doi: 10.21105/joss.03284. URL <https://joss.theoj.org/papers/10.21105/joss.03284>.
- M. A. Steadman, C. Kim, J.-H. Lestang, D. F. M. Goodman, and L. Picinali. Short-term effects of sound localization training in virtual reality. *Scientific Reports*, 9(1):18284, Dec. 2019a. ISSN 2045-2322. doi: 10.1038/s41598-019-54811-w. URL <https://www.nature.com/articles/s41598-019-54811-w>.
- M. A. Steadman, C. Kim, J.-H. Lestang, D. F. M. Goodman, and L. Picinali. Short-term effects of sound localization training in virtual reality. *Scientific Reports*, 9(1):18284, Dec. 2019b. ISSN 2045-2322. doi: 10.1038/s41598-019-54811-w. URL <https://www.nature.com/articles/s41598-019-54811-w>.
- R. Trapeau and M. Schönwiesner. The Encoding of Sound Source Elevation in the Human Auditory Cortex. *The Journal of Neuroscience*, 38(13):3252–3264, Mar. 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2530-17.2018. URL <https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2530-17.2018>.
- 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>.
- M. M. V. Wanrooij and A. J. V. Opstal. Relearning Sound Localization with a New Ear. *Journal of Neuroscience*, 25(22):5413–5424, June 2005. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0850-05.2005. URL <https://www.jneurosci.org/content/25/22/5413>. Publisher: Society for Neuroscience Section: Behavioral/Systems/Cognitive.
- L. M. Ward. Stimulus Information and Sequential Dependencies in Magnitude Estimation and Cross-Modality Matching. *Journal of Experimental Psychology: Human Perception and Performance*, 5(3): 444–459, 1979.

F. L. Wightman and D. J. Kistler. Headphone simulation of free-field listening. II: Psychophysical validation. *The Journal of the Acoustical Society of America*, 85(2):868–878, Feb. 1989. ISSN 0001-4966, 1520-8524. doi: 10.1121/1.397558. URL <https://pubs.aip.org/jasa/article/85/2/868/807017/Headphone-simulation-of-free-field-listening-II>.