Supporting Information

Parise et al. 10.1073/pnas.1322705111

SI Text

Comparison Between the Localization Biases and the Statistics of Natural Sounds and Filtering Properties of the Outer Ear

Strictly speaking, the perceptual biases measured in the psychophysical task do not represent the internal mappings between frequency and elevation but the outcome of such mappings which we estimated with a Bayesian model. Nevertheless, to further prove the link between the measured statistics and observed behavior without relying on a model (which is necessarily based on a set of assumptions, which might as such be wrong), we also directly measured the correlation between the perceptual biases and the frequency–elevation mapping (FEM) in the environment and in the filtering properties of the outer ear. To do so, we used the frequency-dependent bias observed when participants were tilted by 90° (Fig. 1*B*, red lines), that is, when the head- and world-centered FEMs were made orthogonal, so that the elevation bias in head-centered coordinates should reflect the contribution of the head-centered FEM, whereas the azimuth bias (again in head-centered coordinates) should reflect the contribution of the world-centered FEM.

As in the previous section, the similarity between the shapes of the FEM measured from the psychophysical task (Fig. 1*B*, red lines) and from the statistics of the stimulus (Fig. 1*C*) was measured in term of Pearson correlation. The mean and confidence interval of the correlation were calculated by an iterative resampling procedure whereby the estimated priors were correlated to the FEM measured from the mean of a subset of one-fifth of the all recordings (n = 9,962), one-fifth of the head-related transfer functions (HRTFs) (n = 9), and one-fifth of the observers (n = 2). The procedure was repeated 1,000 times.

The correlation between the azimuth bias and the statistics of the environment was 0.76 [95% confidence interval (c.i.) = 0.56–0.87], the correlation between the azimuth bias and the FEM in the proximal stimulus was 0.89 (95% c.i. = 0.72–0.95). The correlation between the elevation bias and the statistics of the environment was 0.90 (95% c.i. = 0.83–0.96), and the correlation between the elevation bias and the FEM in the proximal stimulus was 0.78 (95% c.i. = 0.65–0.86). Notably, this pattern of correlation by and large confirms the findings based on the priors estimated using the Bayesian model and provides further converging evidence supporting the conclusion that the perceptual FEMs reflect the statistics of the proximal and distal stimuli.

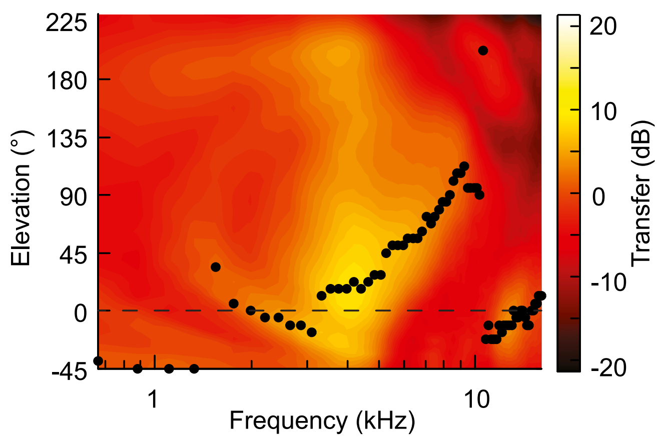

Fig. S1. Average HRTF, obtained by averaging all 45 HRTFs of the CIPIC database. The black dots, representing the elevation with maximum transfer for each frequency, show a clear mapping between frequency and elevation. The FEM reported in Fig. 1*C* (*Lower*) was obtained by calculating for each individual HRTF the elevation with maximum transfer for each frequency (i.e., the black dots here), and then averaging the results across the 45 HRTFs.

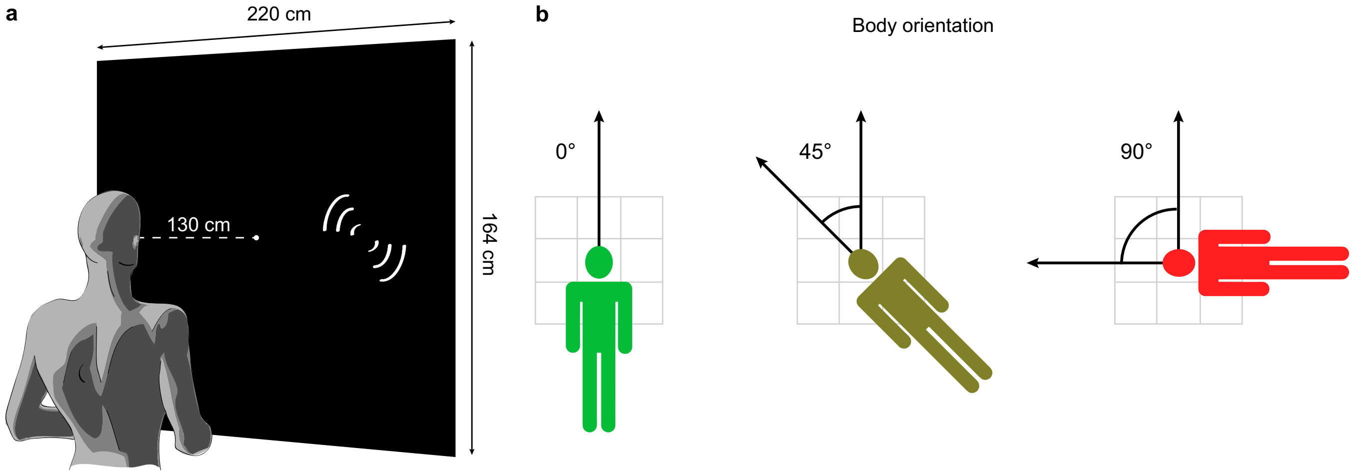

Fig. S2. (*A*) Schematic representation of the experimental setup. (*B*) Representation of the three different body orientations. The vertical arrows represent the world-centered elevation; the tilted arrows represent the head-centered elevation. When the body of the participant is not tilted (*Left*), head- and world-centered elevation overlap, whereas when the body is tilted by 90° (*Right*), the head- and world-centered elevations are orthogonal. The gray grids represent the physical position of the speakers.

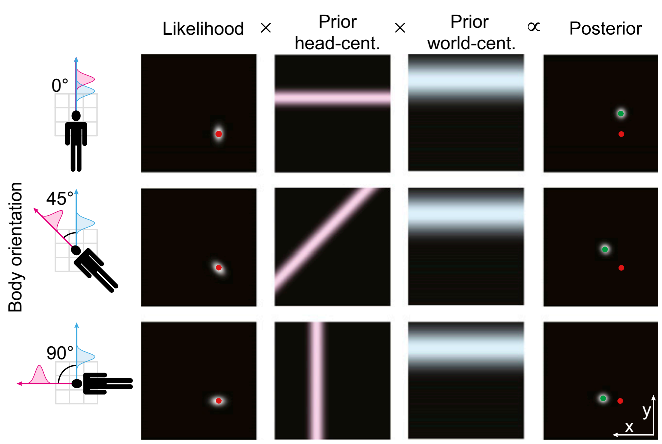

Fig. S3. Schematic illustration of the Bayesian model. The icons on the left represent the different orientations of the observers (in rows). The left column represents the likelihood function (the sensory information). The red dots represent the physical position of the stimulus $s = (s_x, s_y)$. The second and the third columns represent the frequency-dependent priors on elevation in head- and world-centered coordinates, respectively. The last column on the right represents the posterior distribution; the red dot represents the physical position of the stimuli, whereas the green dot represents the maximum a posteriori, that is, the perceived position of the stimuli. Note how the perceived position is shifted away from the actual position as a function of both the frequency-dependent priors and body orientation. Colors indicate the reference frame of the priors (magenta = head-centered; cyan = world-centered). This figure represents the case of the localization of a 4.5–8-kHz band-pass auditory stimulus coming from the bottom-right speaker (s_{wc} = [15, −15]°). Frequency-independent distortions of perceived auditory space (*Modeling*) are not represented.

Movie S1. Auditory stimuli used in the sound localization experiment. To better appreciate how perceived spatial elevation changes as a function of the spectra of the stimuli, we recommend playing the sounds using loudspeakers (not headphones), and listening with the eyes closed.

Movie S1