

# 1 Evoked Responses to Localized 2 Sounds Suggest Linear 3 Representation of Elevation in 4 Human Auditory Cortex

5 Ole Bialas<sup>1 2 3</sup>✉, Burkhard Maess<sup>3</sup>, Marc Schönwiesner<sup>1 4</sup>

6 <sup>1</sup>Faculty for Life Sciences, Leipzig University, Germany; <sup>2</sup>Department of Biomedical  
7 Engineering, University of Rochester, NY, USA; <sup>3</sup>Max Planck Institute for Human  
8 Cognitive and Brain Sciences, Leipzig, Germany; <sup>4</sup>Département de Psychologie,  
9 Université de Montréal, Montreal, Canada

10

## 11 Abstract

12 The auditory system computes the position of a sound along each of the three spatial axes,  
13 azimuth, elevation and distance, from very different acoustical cues. The extraction of sound  
14 azimuth from binaural cues (differences in arrival time and intensity between the ears) is well  
15 understood, as is the representation of these binaural cues in the auditory cortex of different  
16 species. Sound elevation is computed from monaural spectral cues arising from  
17 direction-dependent filtering of the pinnae, head, and upper body. The cortical representation of  
18 these cues in humans is still debated. We have shown that the fMRI blood-oxygen  
19 level-dependent activity in small parts of auditory cortex relates monotonically to perceived  
20 sound elevation and tracks listeners internal adaptation to new spectral cues. Here we confirm  
21 the previously suggested cortical code with a different method that reflects neural activity rather  
22 than blood oxygenation (electroencephalography), show that elevation is represented relatively  
23 late in the cortex, with related activity peaking at about 400 ms after sound onset, and show that  
24 differences in sound elevation can be decoded from the electroencephalogram of listeners,  
25 particularly from those who can distinguish elevations well. We used an adaptation design to  
26 isolate elevation-specific brain responses from those to other features of the stimuli. These  
27 responses gradually increased with decreasing sound elevation, consistent with our previous  
28 fMRI findings and population rate code for sound elevation. The long latency as well as the  
29 topographical distribution of the elevation-specific brain response indicates the involvement of  
30 higher-level cognitive processes not present for binaural cue representation. The differences  
31 between brain responses to sounds at different elevations predicted the listeners sound  
32 localization accuracy, suggesting that these responses reflect perceived elevation. This is, to our  
33 knowledge, the first study that demonstrates the cortical encoding of sound elevation in humans  
34 with high-temporal resolution. Our results agree with previous findings from functional magnetic  
35 resonance imaging, providing strong support for the hypothesis that elevation is represented in a  
36 population-rate code. This represents a critical advance in our understanding of spatial auditory  
37 processing along a dimension that is still poorly understood.

38

✉ **For correspondence:**  
ole.bialas@posteo.de (OB)

**Present address:** Center for  
Advanced Brain Imaging,  
Department of Biomedical  
Engineering, University of  
Rochester, NY, USA

**Data and Code availability:**  
Data is hosted on OpenNeuro  
and can be obtained, together  
with the code via the paper's  
accompanying  
[GitHub-repository](#).

**Funding:** OB was supported  
by funding from the  
Foundation of German  
Business. The funders had no  
role in the experimental  
design or decision to publish.

**Competing interests:** The  
author declare no competing  
interests.

**Acknowledgment:** This  
preprint was created using the  
[LaPreprint template](#) by Mikkel  
Roald-Arbøl.

## 39 Introduction

40 The auditory system does not encode the position of a stimulus on the sensory epithelium in the  
41 cochlea. Instead, it must observe the movement of the eardrums, caused by the superpositioned  
42 sound waves impinging on the ear, and infer source locations from that movement. To this end,  
43 the auditory system processes interaural and spectral localization cues and maps them to locations  
44 in space. Interaural cues are differences in the sound's intensity and timing between both ears  
45 (Rayleigh, 1907; Sandel et al., 1955). They are computed in the superior olivary complex of the  
46 auditory brainstem (McAlpine, Jiang, and Palmer, 2001; Yin, 2002) and determine our perception  
47 of a sound's azimuth. Spectral cues are patterns of peaks and notches in the sound's spectrum  
48 resulting from directional filtering through head and pinnae (Batteau, 1967; Wightman and Kistler,  
49 1989a; Hofman, Van Riswick, and Van Opstal, 1998). The positions of peaks, notches and edges in  
50 the spectrum are computed in the dorsal and posterior ventral cochlear nucleus and indicate the  
51 sound's elevation (Nelken and Young, 1994; Reiss and Young, 2005).

52 Interaural and spectral cues are integrated in the midbrain where neurons responding to sounds  
53 in a narrow region form a topographic representation of auditory space (Middlebrooks and Knud-  
54 sen, 1984; King and Hutchings, 1987; Gaese and Johnen, 2000). However, most cortical neurons do  
55 not respond selectively, but rather modulate their response across a large section of space, prefer-  
56 entially within the contralateral hemifield (Imig, Irons, and Samson, 1990; Middlebrooks et al., 1994;  
57 Brugge, Reale, and Hind, 1996; McAlpine, Jiang, and Palmer, 2001). Thus, sound location could be  
58 represented in an opponent-channel code where two neural populations are broadly tuned to the  
59 contralateral hemifield, encoding location in their relative levels of activity (Salminen et al., 2010;  
60 Magezi and Krumbholz, 2010). This model predicts the greatest change in the neural response,  
61 and hence highest perceptual resolution, around the midline where tuning curves of both popula-  
62 tions intersect. While this may be a plausible mechanism for the perception of azimuth, which is  
63 most accurate around the midline, it does not explain the human perception of sound elevation  
64 which has a relatively constant accuracy throughout the frontal field (Wightman and Kistler, 1989b;  
65 Middlebrooks and Green, 1991).

66 Instead, elevation could be represented by the overall level of activity within a single neural  
67 population. Evidence from fMRI suggests that all elevation-sensitive voxels respond with a similar  
68 linear decrease in activity to increasing elevation and that this tuning co-varies with the effects of  
69 behavioral manipulations (Trapeau and Schönwiesner, 2018). However because fMRI measures  
70 the hemodynamic response which is delayed with respect to neural activity, the latency of cortical  
71 elevation processing remains unclear. Also, lying down during scanning might change the neural  
72 tuning which is, at least in part, allocentric (Schechtman, Shrem, and Deouell, 2012; Town, Brim-  
73 ijoin, and Bizley, 2017). Thus we used EEG for measuring neural responses with high temporal  
74 resolution in a more natural listening situation to test the hypothesis that the cortex represents  
75 sound elevation in a linear population rate code.

76 Unfortunately, EEG, unlike fMRI, can not measure the elevation-sensitive populations in isola-  
77 tion but instead picks up a mixture of all instantaneous neural activity. Our previous fMRI results  
78 suggest that only a fraction of auditory cortex encodes sound elevation (Trapeau and Schönwies-  
79 ner, 2018). Thus elevation-specific components might be obscured by responses to sound onsets.  
80 Indeed, a study that attempted to decode sound location from EEG found that the decoding accu-  
81 racy for elevation was barely above chance (Bednar, Boland, and Lalor, 2017). To overcome this  
82 limitation, we used neural adaptation which is the decay in neural activity following repeated or  
83 continuous stimulation due to a variety of physiological mechanisms (Benda, 2021). By playing  
84 a longer trail of noise (adapter) before each short stimulus (probe) we cause adaptation of the  
85 sound-responsive neurons. Since adapter and probe are cross-faded and identical except for their  
86 elevation, the response to the probe is driven by the change in elevation, rather than the over-all  
87 sound onset response. We used two different variations of this paradigm, one where the adapter  
88 is presented from a loudspeaker and one where it is presented from fully open "hovering"-style

89 headphones. The latter design has the advantage that the adapter is not affected by the directional  
90 ear filter and thus does not cause adaptation of elevation-sensitive neurons (Møller et al., 1995).

## 91 **Methods & Materials**

### 92 **Subjects**

93 Twenty-three subjects (15 female) participated in the first and thirty subjects (15 female) in the  
94 second experiment. They were between 20 and 33 years old and had no history of neurological  
95 or hearing disorders. All subjects gave informed consent and were monetarily compensated at an  
96 hourly rate. All procedures were approved by the ethics committee of the medical faculty at the  
97 University of Leipzig (reference number 248/17-ek).

### 98 **Apparatus**

99 Subjects sat in the center of a custom-built spherical array of loudspeakers (model Mod1, Sherman  
100 Oaks, CA, USA) inside a 40 m<sup>2</sup> hemi-anechoic chamber (Industrial Acoustics Company, Niederkrüchten,  
101 Germany). Two additional speakers were mounted hovering next to the subject's ears, pointed at  
102 the ear canal, and used as fully open headphones in the second experiment. Because of their  
103 proximity and orientation, sounds from those loudspeakers were unaffected by the listeners di-  
104 rectional transfer function (Møller et al., 1995). This allowed us to simultaneously present sounds  
105 from loudspeakers at different elevations in the array as well as non-spatial headphone sounds.  
106 We equalized the transfer functions of each loudspeaker by applying an inverse filter to the stim-  
107 uli upon presentation. Two digital signal processors and six 8-channel amplifiers (models RX8.1  
108 and SA8, Tucker-Davis Technologies, Alachua, FL, USA) drove the loudspeakers at 50 kHz sampling  
109 rate. The processors' digital ports controlled the LEDs attached to each loudspeaker, obtained re-  
110 sponses from a custom-built button box, and sent event triggers to the EEG. We used a 64-channel  
111 system (model BrainAmpMR, Brain Products, Gilching Germany) to record EEG at a sampling rate  
112 of 500 Hz. The active silver/silver chloride electrodes were fixed on the subject's head with an  
113 elastic cap (EasyCap, Germany) according to the international 10/20 system with FCz as reference  
114 electrode. Electrode impedance was kept below 2 k $\Omega$ . Two cameras (model Firefly S, Teledyne FLIR,  
115 OR, USA), positioned between loudspeakers, were used to monitor the subjects' head pose.

### 116 **Software**

117 We programmed the digital signal processors using Real-time Processor Visual Design Studio (Tucker-  
118 Davis-Technologies, Alachua, FL, USA) and controlled the cameras using the Spinnaker SDK (Tele-  
119 dyne FLIR, OR, USA). The software of both devices provides an API that we integrated in a custom  
120 Python module that controlled the experimental apparatus (Bialas, 2022). We estimated subjects'  
121 head pose by capturing an image of their head, localizing facial landmarks with a deep neural  
122 network, implemented in Pytorch (Paszke et al., 2017), and mapping these points to a generic  
123 3D model using functions from the OpenCV library (Bradski, 2000). Stimuli and trial-sequences  
124 were generated using the slab Python module for psychoacoustics (Schönwiesner and Bialas, 2021).  
125 We recorded EEG signals using Brain Vision Recorder (Brain Products, Gilching Germany), and im-  
126 ported them into MNE-Python for analysis (Gramfort et al., 2013). A full list of the software envi-  
127 ronment can be found in the accompanying [online repository](#).

### 128 **Preprocessing**

129 A de-noising algorithm, combining filtering and source separation, removed power line artifacts  
130 while minimizing temporal distortions due to filtering (Cheveigné, 2020). We then applied a causal,  
131 minimum- phase, high-pass filter with a hamming window and a 1 Hz cutoff frequency. After  
132 epoching (without applying a baseline), channels in which the correlation between the actual signal  
133 and the signal predicted by the neighboring channels was less than 0.75 for more than 40 percent  
134 of the time (Bigdely-Shamlo et al., 2015) were replaced with interpolated data from the surrounding

135 channels using spherical splines. We then subtracted the average signal across all channels from  
136 each channel ("average reference"). Next, independent component analysis unmixed the signal,  
137 and an algorithm removed components corresponding to eye-blinks based on their topographical  
138 correlation with a previously determined template (Viola et al., 2009; Plöchl, Ossandón, and König,  
139 2012). Finally, an algorithm determined channel-specific rejection thresholds and repaired or re-  
140 moved epochs where the thresholds were exceeded (Jas et al., 2017). Except for the initial selection  
141 of a reference for blink removal, the entire preprocessing pipeline was automated. Afterwards, the  
142 data were inspected by eye to assess the effect of preprocessing.

### 143 **Experiment I**

144 Subjects sat comfortably on a height-adjustable chair in an anechoic chamber. Target loudspeak-  
145 ers stood at a distance of 3.2 m at elevations of 37.5°, 12.5°, -12.5° and -37.5° with respect to  
146 the subject's interaural plane. Because perception of sound source elevation is slightly more ac-  
147 curate for lateral targets (Makous and Middlebrooks, 1990), all target speakers were positioned  
148 at an azimuth of 10° to the subject's right. Initially, we tested subjects' ability to localize sounds  
149 and familiarized them with the setup. To avoid that subjects explicitly learn the target speakers'  
150 directional transfer functions during training, we used a different set of loudspeakers located at  
151 elevations of 50°, 25°, 0°, -25° and -50°. In each of the 200 test trials, subjects heard a 150 ms  
152 burst of noise with a 5 ms on- and offset ramp from one speaker. Subjects localized each sound  
153 by pressing one of four buttons on a custom-built box. There was no time limit for responding,  
154 and the subsequent trial started automatically after the subject had responded to the previous  
155 stimulus. The order of speakers was randomized without direct repetition of the same speaker.  
156 Subjects were instructed to keep their head and gaze aligned with the fixation cross at 0° azimuth  
157 and elevation. After completion of the test, we prepared the EEG electrodes. During recordings,  
158 each trial started with 600 ms of noise (adapter) played from the speakers at either 37.5° or -37.5°.  
159 Then, a 150 ms burst of noise played from one of the other speakers, resulting in six different  
160 adapter-probe pairs. Adapter and probe had overlapping 5 ms on- and offset ramps so that the  
161 sound intensity remained constant during the transition. The adapter's initial position was chosen  
162 randomly and changed every 30 trials. The probe's location was chosen randomly without direct  
163 repetitions of the same speaker. Every probe was followed by a 350 ms silent inter-stimulus inter-  
164 val. In five percent of all trials, the probe did not come from one of the target positions but from a  
165 random speaker within the frontal field. Subjects had to respond to these deviant trials by pressing  
166 a random button as fast as possible. If they managed to respond within one second after sound  
167 onset, the trial was considered a success. After one second, the trial was considered failed, and  
168 stimulation resumed. The experiment was divided into four blocks, each of which consisted of 504  
169 trials and lasted 35 min in total. We instructed subjects to keep their head and gaze aligned with  
170 the fixation cross throughout the recording but we did not check whether they complied.

### 171 **Experiment II**

172 Again, subjects completed an initial test in which they had to localize sounds coming from loud-  
173 speakers at 50°, 25°, 0°, -25° and -50°. In each of the 30 trials, subjects heard a 100 ms from one  
174 of the speakers which they localized by pointing their head in the direction of the speaker and  
175 pressing a button. This triggered the cameras to acquire images from which the head-pose was  
176 estimated. After localizing the sound, subjects had to return to the central fixation cross and press  
177 the button again to start the next trial. If their head pose was not aligned with the fixation cross  
178 a warning tone prompted them to adjust their position. The first 15 stimuli were accompanied  
179 by a visual cue (i.e. a flashing LED at the loudspeakers' location) and familiarized subjects with  
180 the procedure. The stimulus protocol during recording was similar to the experiment I except for  
181 the duration of adapter (1000 ms), probe (100 ms) and inter-stimulus interval (900 ms). Also, the  
182 adapter was played from headphones instead of one of the target speakers. In five percent of  
183 trials, subjects heard a tone after the inter-stimulus interval which prompted them to localize the

184 probe they had just heard. Thus, they had to pay attention to each probe's location since they did  
185 not know whether they would be asked to localize any given sound. The experiment was divided  
186 into six blocks, each consisting of 240 trials, roughly lasting 55 min in total.

### 187 **Elevation Gain**

188 To quantify the accuracy of the subjects' perception of elevation, we used the elevation gain (EG)  
189 which is the slope of the regression between the target and response elevation. Thus, the EG  
190 measures how strongly subjects modulate their response with changes in the target's elevation  
191 which 1 indicating perfect localization and 0 indicating random responses. We chose EG because it  
192 robust to outliers and has been shown to capture changes in behavior and neural tuning following  
193 manipulation (Hofman, Van Riswick, and Van Opstal, 1998; Trapeau and Schönwiesner, 2018). We  
194 computed the EG for the localization tests in both experiments as well the task in experiment II.  
195 In the latter, the EG for two subjects, which was slightly negative, was set to zero for subsequent  
196 analyses.

### 197 **Permutation Testing**

198 When investigating a phenomenon with unknown latency and locus it is desirable to test for dif-  
199 ferences between conditions across all points in time and space while controlling the false alarm  
200 rate. This is achieved by a permutation-based cluster test which operates under the null hypothesis  
201 that all conditions are exchangeable and finds spatio-temporal clusters that violate this assump-  
202 tion (Maris and Oostenveld, 2007). Using this algorithm, we calculated F-scores, selected samples  
203 where the F-score exceeded a threshold corresponding to an uncorrected p-value of 0.05 and clus-  
204 tered them for temporal and spatial adjacency. The size of an observed cluster is defined by the  
205 sum of all F-scores. By repeating this procedure on randomly permuted data, we obtain a distri-  
206 bution of cluster sizes that occur by chance. The p-value of the actually observed effects is given  
207 by the probability of observing a similar or larger effect under the permutation distribution. Thus,  
208 the cluster test is sensitive while avoiding the multiple comparisons problem because statistical  
209 inference is constrained to the distribution of cluster sizes. This however means that the spatial  
210 and temporal extent of the clusters is purely descriptive and not statistically controlled (Sassen-  
211 hagen and Draschkow, 2019). We used the cluster test implemented in MNE-Python with 10000  
212 permutations. We conducted one test per subject, comparing all conditions simultaneously. If the  
213 test returned more than one significant cluster for a subject we ignored all but the largest.

### 214 **Decoding**

215 While significant clusters indicate differences between conditions, they do not indicate which con-  
216 ditions differ in what way. Thus, to complement permutation testing, we used multivariate logistic  
217 regression to decode sound elevation from brain responses. We chose logistic regression because  
218 it is simple, well-understood, and robust to overfitting (Dreiseitl and Ohno-Machado, 2002). The  
219 multivariate logistic function was fit to the data sample-by-sample using ridge-regression, and de-  
220 notes the probability of a given observation (i.e. the instantaneous distribution of voltage across  
221 channels) belonging to either one of two classes (i.e. elevations). To avoid picking an arbitrary deci-  
222 sion threshold, we used the receiver operating characteristic (ROC) curve, which denotes the ratio  
223 between true and false positives across all possible thresholds. Thus, the area under the ROC curve  
224 is an unbiased measure of accuracy with 1 indicating perfect and 0.5 indicating random classifica-  
225 tion. To avoid overfitting we used 100-fold leave-one-out cross-validation, splitting the data into  
226 100 segments, fitting the logistic function to 99 of them and testing it on the left out segment. Each  
227 segment was used for testing once, and the accuracy was averaged across all segments. Finally,  
228 we used the bootstrap method to resample the subject-specific decoding accuracies 10000 times  
229 (Efron, 1992). The resampled data's mean is an estimate of the group's average decoding accuracy  
230 across time, and its standard deviation indicates the uncertainty of the mean estimate. We used



231 the logistic regression implementation from the scikit-learn package (Pedregosa et al., 2011) and  
232 the sliding estimator function from MNE-Python, which applied the classifier sample-by-sample.

### 233 **Elevation Tuning**

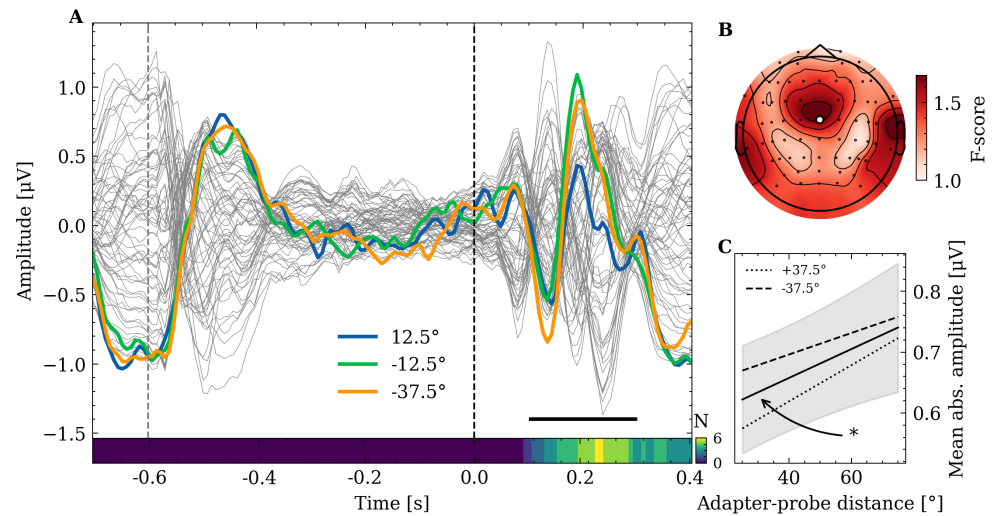
234 Our main objective was to estimate how brain responses to sounds change with elevation. To  
235 this end, we selected the channels where the difference between conditions (i.e. the F-score) was  
236 largest in the time interval after subjects heard the probe. Then, we estimated the relationship  
237 between average ERP amplitude and sound elevation for all subjects using ordinary least squares  
238 linear regression. To quantify the uncertainty in the estimated relationship we resampled the data  
239 10000 times using the bootstrap method, applied linear regression to each resampled set, and  
240 computed the standard deviation. Since the bootstrapped sampling distribution is normal, over  
241 95 percent of all estimated linear models lie within two standard deviations of the mean. However,  
242 since EEG-sensors measure Voltage relative to the reference, an observed change at any channel  
243 could also reflect a change in the reference (i.e. the average of all channels) of opposite sign. To  
244 overcome this limitation we computed the current source density (CSD) using the implementation  
245 from MNE-Python, which is based on spherical spline surface Laplacians. The CSD is a reference-  
246 free estimate of radial current flow at the scalp where positive values represent outward (i.e. from  
247 the brain to the scalp) and negative values inward current flow (Kayser and Tenke, 2015). Addi-  
248 tionally, it acts as a spatial high-pass filter suppressing the effects of volume conductance and  
249 improving the spatial resolution (Cohen, 2014). This effectively simplifies the variance structure  
250 of the data making CSD a useful preprocessing step for principle component analysis (Kayser and  
251 Tenke, 2006). PCA transforms the data to a n-dimensional orthogonal space so that each dimen-  
252 sion accounts for the largest possible share of variance. Thus, the principle components are pat-  
253 terns of covariance reflective of underlying neural generators that generated the observed data  
254 (Cohen, 2014). This follows the rationale that channels co-vary because they are affected by the  
255 same generator which might be a localized source or a distributed network. PCA is a simple and  
256 well-understood method for data-driven detection of relevant components and has been shown  
257 to produce meaningful summaries of average evoked responses (Donchin, 1966; Chapman and  
258 McCrary, 1995; Kayser and Tenke, 2006).

### 259 **Lowest sounds cause the largest responses**

260 Participants were able to localize sounds accurately as shown by an average EG of 0.74 ( $SD = 0.19$ ) in  
261 the initial localization tests, which is similar to what was reported in previous studies (Hofman, Van  
262 Riswick, and Van Opstal, 1998; Trapeau and Schönwiesner, 2018). During the continuous sound  
263 presentation in experiment 2, uncertainty about target trials and added memory load reduced  
264 the average elevation gain significantly compared to the performance in the localization test just  
265 before (0.78 and 0.45,  $t(29) = 6.05$ ,  $p < 0.001$ ).

266 Neural responses generally increased with decreasing sound elevation in both experiments  
267 although the effects differed in size and latency. In the first experiment, elevation-specific event-  
268 related potentials (ERPs) differed mostly in the time interval from 100 ms to 300 ms after probe  
269 onset (colored bar in 1A). This was confirmed by a permutation-based cluster test, which compared  
270 the responses to all six adapter-probe combinations (adapter at 37.5° and probes at 12.5°, -12.5°,  
271 -37.5° and adapter at -37.5° and probes at -12.5°, 12.5°, 37.5°). The test found significant clusters,  
272 comprising most of the electrodes, for 9 of the 23 subjects. Examination of the F-scores, averaged  
273 across time between 100 ms and 300 ms (Fig.1B) revealed that the difference between elevation-  
274 specific ERPs was largest at the central electrode Cz. Thus, we selected this electrode to investigate  
275 how the ERP changed with elevation more closely.

276 The average ERP at Cz showed two deflections, of opposite polarity, which increased in ampli-  
277 tude with the separation of adapter and probe (Fig.1A). To quantify this trend, we computed the  
278 average absolute ERP amplitude in the time interval between 100 ms and 300 ms after probe onset  
279 for each adapter-probe combination and regressed it against the distance between adapter and

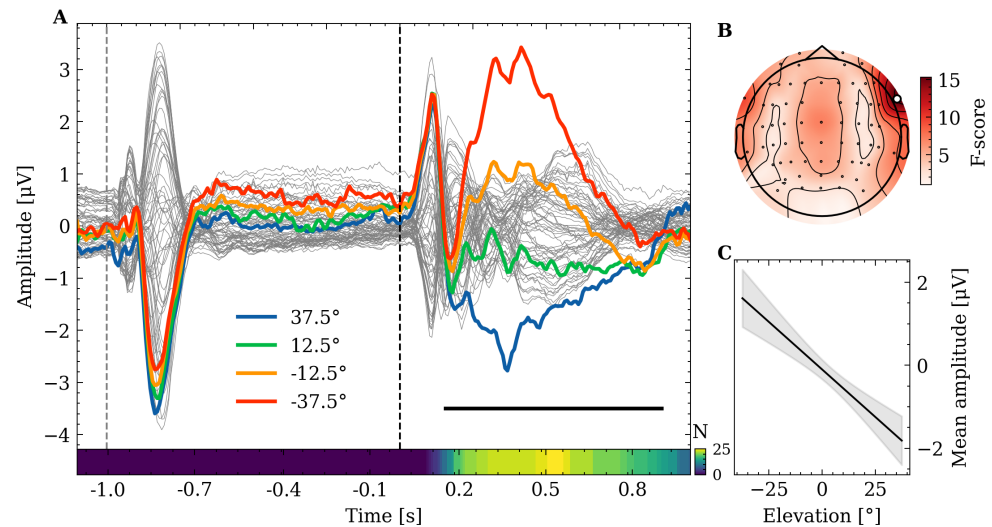


**Figure 1. A:** Grand average evoked response from the first experiment. Gray lines show the voltage at each channel averaged across sound sources and colored lines show the response to each probe at the channel Cz where the difference between conditions was largest (white dot in **B**), when the adapter was played from 37.5°. The gray dashed line marks the onset of the adapter and the black dashed line the transition to the probe. The horizontal color bar indicates the number of subjects for whom the cluster-test showed a significant differences between conditions at any point in time. **B:** topographical distribution of F-scores. Color represents the average F-score at each channel in the time interval marked by the horizontal black line in **A**. The white dot indicates the channel with the largest score (Cz). **C:** relationship between adapter-probe separation and average absolute ERP amplitude. The solid gray line shows the change in average absolute amplitude across all sound sources and the dotted and dashed lines show the change across all sources when the adapter located at -37.5° and 37.5° respectively. The shaded interval shows  $\pm 2SD$ , estimated via bootstrapping.

280 probe (Fig.2. The average absolute amplitude significantly increased with separation of the adapter  
 281 and probe ( $R = 0.18, p = 0.033$ ). This increase was steeper when the adapter was located at 37.5°  
 282 ( $\beta = 29.8 * 10^{-4}$ ) compared to when the adapter was located at -37.5° ( $\beta = 17.7 * 10^{-4}$ ). Even though  
 283 the trend was not significant when considering only those trials where the adapter was played at  
 284 37.5° or -37.5° ( $R = 0.22, p = 0.06$  and  $R = 0.14, p = 0.26$ ), this observation is compatible with a mono-  
 285 tonic population-rate code which predicts that elevation-sensitive neurons respond maximally to  
 286 the lowest sound (i.e. the adapter at -37.5°), rendering them insensitive to subsequent sounds  
 287 from higher elevations.

288 The second experiment addressed the confound between adapter and probe position by using  
 289 a non-spatial adapter. Consequently, the results offered clearer support for the population-rate  
 290 code hypothesis. The permutation test which compared responses to the different probes at 37.5°,  
 291 12.5°, -12.5° and -37.5° elevation found significant clusters for most subjects between 200 ms and  
 292 800 ms with a peak around 500 ms after probe onset (colored bar in Fig.2A). The scalp-distribution  
 293 of F-scores averaged between 150 ms and 900 ms after probe onset (Fig.2B) revealed that the dif-  
 294 ference between elevation-specific ERPs was largest at fronto-temporal electrodes with a peak at  
 295 FT10. Again, we chose this electrode for a closer inspection of elevation-specific changes in the  
 296 ERP.

297 The average ERP at FT10 exhibited a sustained deflection that was strongly modulated in ampli-  
 298 tude by the probe's elevation (colored lines in Fig.2A). The amplitude was largest when the probe  
 299 was located at -37.5° and smallest when it was located at 37.5°. Linear regression found that the re-  
 300 lationship between the probe's elevation and the average response amplitude in the time interval  
 301 from 150 ms to 900 ms after onset was highly significant ( $R = -0.66, p < 0.001$ ).



**Figure 2.** **A:** Grand average evoked response from the second experiment. Gray lines show the voltage at each channel averaged across conditions and colored lines show the response to each probe at the channel FT10 where the difference between conditions was largest (white dot in **B**). The gray dashed line marks the onset of the adapter and the black dashed line the transition to the probe. The horizontal color bar indicates the number of subjects for whom the cluster-test showed a significant differences between conditions at any point in time. **B:** topographical distribution of F-scores. Color represents the average F-score at each channel in the time interval marked by the horizontal black line in **A**. The white dot indicates the channel with the largest score (FT10). **C:** relationship between probe elevation and average ERP amplitude. The shaded interval shows  $\pm 2SD$ , estimated via bootstrapping.

### 302 Elevation decoding accuracy predicts task performance

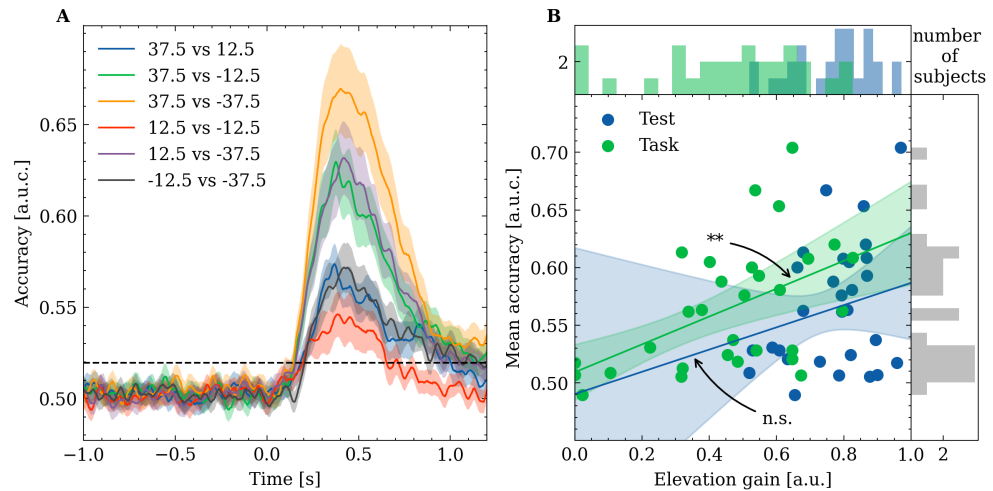
303 Sounds at different elevations evoked visibly different ERPs. Thus, a logistic regression classifier  
304 which decoded sound elevation from EEG performed above chance for all sound source pairs  
305 (Fig.3A). Decoding accuracy increased with distance between adapter and probe and was not de-  
306 pendent on the sounds' absolute elevation (e.g. the curves for 37.5° vs 12.5° and -37.5° vs. -12.5°  
307 were virtually identical). Decoding accuracy followed a similar time course in all conditions: it re-  
308 mained at chance level while the adapter was presented, started to increase around 200 ms and  
309 peaked around 400 ms after probe onset. Participants who were better at distinguishing sounds at  
310 different elevations also had more distinctive ERPs in response to sounds from these different ele-  
311 vations, allowing better decoding accuracy. Subjects' average decoding accuracy between 150 ms  
312 and 900 ms after probe onset correlated with their elevation gain during the experimental task  
313 ( $R = 0.51, p = 0.004$ ). Notably, this relationship is mostly due to the upper right quadrant in Fig-  
314 ure3B being empty, meaning there were no subjects who's brain responses were decodable but  
315 who failed at the task. There were, however, several subjects for whom decoding failed but who  
316 still performed the task accurately.

317 We also regressed the average decoding accuracy against the EG during the initial localization  
318 test to investigate whether the relationship between decoding and performance was specific to  
319 the experimental task or applied to sound localization in general. While linear regression revealed  
320 a positive relationship between decoding accuracy and localization test EG, this trend was not sig-  
321 nificant ( $R = 0.23, p = 0.23$ ), likely due to the small number of trials and lack of variance across  
322 subjects in the localization test.

### 323 Auditory component activates with decreasing elevation

324 Because ERP morphology depends on the chosen reference it is unclear whether the observed  
325 changes in ERP amplitude with elevation reflect a decrease or a change in polarity with respect to  
326 the underlying current flow. To resolve this ambiguity we computed the current source density





**Figure 3. A:** accuracy across time for decoding the ERPs to all possible pairs of probes. The displayed data are averages across all cross-validation splits. **B:** relationship between decoding accuracy and localization behavior. The lines show the average relationship between decoding accuracy and EG during the localization test (blue) and experimental task (green). Shaded intervals show  $\pm 2SD$  estimated via bootstrapping. Bar plots on the right show the marginal distributions of EG during the test (blue), EG during the task (green) and average decoding accuracy (gray).

327 (CSD), a reference-free estimate of current flow, and performed a principle component analysis.  
328 The principle component accounting for most variance had a topography suggesting an auditory  
329 origin and showed a deflection which gradually increased in amplitude for decreasing elevation  
330 (Fig.4A). The second and third components might reflect activity of the motor and prefrontal cortices  
331 respectively (Fig.4B&C). These additional components were unaffected by elevation. Together, the  
332 three components accounted for 88 % of the variance in the average evoked response.

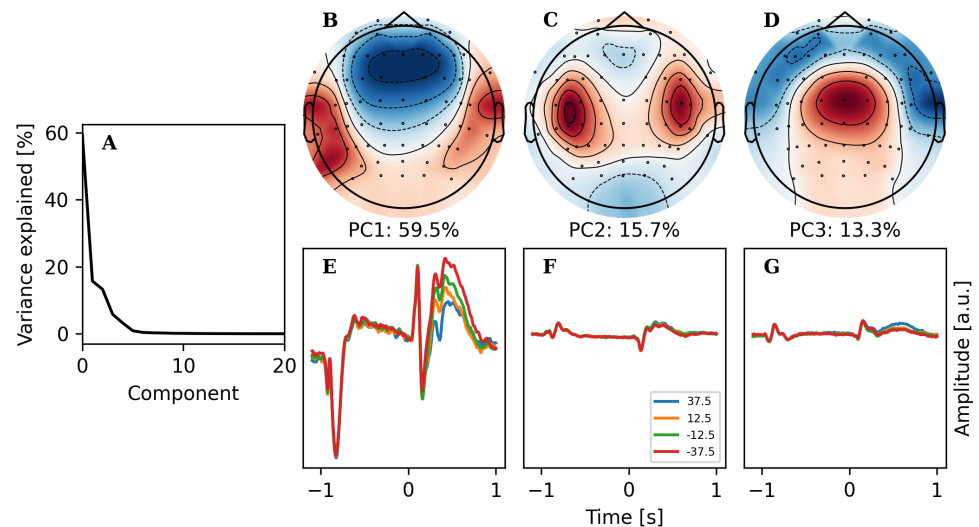
## 333 Discussion

### 334 Evoked responses encode sound elevation

335 Both experiments demonstrate that the cortical processing of sound elevation can be assessed  
336 with EEG. Using fMRI, we previously identified voxels in auditory cortex tuned to sound elevation  
337 (Trapeau and Schönwiesner, 2018), but the lower spatial resolution of EEG would not have allowed  
338 us to isolate the responses from these small patches of auditory cortex. Indeed, a previous EEG  
339 study attempted to decode sound elevation from EEG responses and found that decoding accu-  
340 racy exceeded chance level just barely for only some of the listeners (Bednar, Boland, and Lalor,  
341 2017). The adaptation design helped to circumvent this issue by separating neural activity related  
342 to sound onset and elevation in time. In the second experiment, we showed that changes in sound  
343 elevation evoke distinct cortical responses that can be decoded accurately, and that decoding accu-  
344 racy predicts individual localization performance. Decoding accuracy for all pair-wise elevation  
345 comparisons followed the same time course with a single peak at around 400 ms after the eleva-  
346 tion information became available. The time course and the fact that accuracy was predictive of  
347 task performance suggests that we decoded a physiological process distinctly related to elevation  
348 rather than processes associated with acoustical features that co-vary but are not causally related  
349 with sound elevation. Elevation-specific activity was well captured by a single principle component,  
350 suggesting that the observed encoding of elevation reflects a distinct physiological process.

### 351 A monotonic population-rate code for elevation

352 We hypothesized that the auditory cortex represents sound elevation in a population response that  
353 decreases monotonically with increasing elevation. Both experiments presented here offer sup-



**Figure 4.** Principle components of the average current source density. **A:** the percentage of variance accounted for by each component. **B-D:** Channel weights of the components as topographical distribution on the scalp. The percentage indicates the total variance accounted for by the respective component. **E-F** Component loading across time for each sound elevation.

354 port for this hypothesis. In experiment 1, ERP amplitude increased with separation of the adapter  
355 and probe when the adapter was played from the highest position. Under the assumption of a rate  
356 code with minimal response at high elevations, this adapter would evoke small responses and en-  
357 able increasingly larger responses for probes at lower elevations, with corresponds to the observed  
358 pattern. Consequently, the adapter played from the lowest elevation would lead to strong adap-  
359 tation independent of probe elevation. Thus, as expected, differences between elevation-specific  
360 ERPs were larger when the adapter was located at a high elevation. Unexpectedly however, probes  
361 tended to evoke larger responses when the adapter was located at a low elevation. This may be  
362 explained by co-varying acoustic factors contributing to the probe response and it does not con-  
363 tradict the representation we propose since absolute amplitude does not carry information in a  
364 rate code. To summarize, while experiment 1 showed that elevation is reflected in the ERP, its in-  
365 terpretation is limited by the fact that responses are affected by both the adapter and the probe's  
366 position.

367 The purpose of experiment 2 was to remove the interaction of adapter and probe position by  
368 using a non-spatial adapter (see Methods). Consequently, elevation-specific modulation of ERP am-  
369 plitude was about an order of magnitude larger compared to the first experiment, and this method  
370 allowed direct measurement of monotonic elevation tuning functions. Sounds from the lowest el-  
371 evations evoked the largest response, and response amplitude gradually declined with increasing  
372 elevation in a way that was well described by a linear model. The sign of the elevation-specific re-  
373 sponse amplitude changed around the listener's eye level. Because EEG-electrodes measure volt-  
374 age relative to a common reference, it is unclear whether this reflects a gradual decrease or change  
375 in polarity of the underlying neural current. We were able to decide between these alternatives by  
376 computing a reference-free estimates of neural current flow (current source density), which bene-  
377 fits topographical localization (Kayser and Tenke, 2015). One principle component accounted for  
378 more than half of the ERP variance and showed a topography that suggested it originated in the au-  
379 ditory cortex. This component showed a gradual increase in response amplitude with decreasing  
380 elevation, supporting the idea of a monotonic population rate code. Two additional components,  
381 which together with the first accounted for almost 90% of the variance, showed topographies sug-  
382 gesting they originated in motor and frontal cortices. Involvement of those regions plausible given

383 that the experimental task required the participants to translate their auditory perception into  
384 goal-directed head movements.

385 Thus, our results bear striking similarity to changes in the BOLD-response to elevated sounds  
386 (Trapeau and Schönwiesner, 2018). This is remarkable given the fact that many neural phenom-  
387 ena are not similarly reflected in EEG and fMRI. For example, increases in BOLD-activity may not  
388 be picked up by EEG if the sources are not synchronized or their geometric orientation prevents  
389 a summation of their local field potentials (Buzsáki, Anastassiou, and Koch, 2012), and increases  
390 in EEG power can reflect an overall decrease of the underlying neural activity (Musall et al., 2014).  
391 Thus, our combined findings constrain the cortical encoding of sound elevation to sources that af-  
392 fect EEG and fMRI signals in the same way. They also suggest that the cortical encoding of elevation  
393 is egocentric, because fMRI recordings in supine and EEG recordings in upright position revealed  
394 a similar trend.

### 395 **Latency of cortical elevation processing**

396 In the first experiment, we found the largest elevation-specific response between 200 ms and  
397 250 ms after probe onset. This latency is similar to other reports on the cortical processing of  
398 elevation. Using MEG, Fujiki and colleagues found that unexpected sounds, deviating in eleva-  
399 tion, caused mismatch responses between 150 ms and 250 ms (Fujiki et al., 2002). Bednar and col-  
400 leagues found that some participants' brain responses to two differently elevated sound sources  
401 could be decoded above chance between 200 ms and 400 ms after sound onset (Bednar, Boland,  
402 and Lalor, 2017). Notably both those studies also reported correlates of sound azimuth which oc-  
403 curred earlier in time reflecting the fact that binaural cues depend on precise temporal information  
404 while spectral cues result from a detailed analysis of spectral patterns.

405 In the second experiment, the elevation-specific component started at a similar time but lasted  
406 much longer, so that most subjects showed significant differences across ERPs between 200 ms  
407 and 800 ms. This is reflected in the time-course of decoding accuracy, which peaked around 400 ms  
408 after probe onset. Remarkably, elevation-specific ERPs could still be distinguished 1 s after probe  
409 onset, which means that the brain could access the relevant perceptual representation at the end  
410 of a trial when participants were informed whether they had to indicate the sound location in this  
411 trial. Thus, the increased duration of elevation-specific responses might be a result of the experi-  
412 mental task where maintaining the probe's perception facilitated performance. Interestingly, visual  
413 research identified late ERP components that scale with task-difficulty in latency and magnitude in  
414 a way that is predictive of individual performance (Philiastides and Sajda, 2006; Philiastides and  
415 Sajda, 2007). Such perceptual persistence could be implemented by feedback loops which rever-  
416 berate the neural representation after stimulus offset (VanRullen and Koch, 2003; Large, Aldcroft,  
417 and Vilis, 2005). The perceptual persistence of a visual stimulus is inversely related to its dura-  
418 tion and intensity (Coltheart, 1980). Thus, perceptual persistence could provide a mechanism of  
419 evidence accumulation for perceptual decision making under difficult conditions. However, the  
420 present study was not designed to investigate the effects of task-difficulty and further research is  
421 required to answer those questions.

### 422 **The cortical representation of sound direction**

423 It is thought that the auditory cortex represents sound azimuth as the difference between the  
424 rates of activity in two opponent neural channels, each tuned to the contralateral hemifield so that  
425 most azimuth sensitive neurons respond maximally to cues outside the range created by the head  
426 (McAlpine, Jiang, and Palmer, 2001). This may sound counter-intuitive but, in a rate code, accuracy  
427 is not related to the absolute magnitude but rather the change in neural response. Thus, placing  
428 the tuning curve's peak outside the physiological range places the slope in the center, where az-  
429 imuthal localization is most accurate (Harper and McAlpine, 2004). Similarly, the fact that sounds at  
430 low elevations cause larger responses than sounds at high elevations does not mean that they are

431 localized more accurately. Instead, the steady slope of the population response suggests that accu-  
432 racy remains constant across elevations, which lines up well with localization behavior (Wightman  
433 and Kistler, 1989b; Middlebrooks and Green, 1991). There is no cardinal reason why the maximum  
434 of the tuning is at low rather than high elevations. While it is tempting to argue that this may be  
435 due to a tendency for deeper spectral notches at low elevations, this explanation was disproved  
436 by our previous finding that the tuning function depends on perception rather than the acoustic  
437 nature of the cues: the tuning curve flattens when listeners are presented with spectral cues from  
438 other ears, and they re-emerge as the listener adapts to the new cues (Trapeau and Schönwiesner,  
439 2018).

440 If both elevation and azimuth were represented in a rate code, integration could be fast and re-  
441 quire little computational effort. In the simplest case, the normalized rates of both codes could be  
442 summed to obtain a joint representation of angular separation, in which the difference in activity  
443 elicited by a pair of sound sources is proportional to their angular separation. This unidimensional  
444 activity gradient would be insufficient to encode the joint two-dimensional direction of a sound  
445 source, but the uncertainty could be resolved by the difference in latency between azimuth and  
446 elevation processing (Fujiki et al., 2002; Bednar, Boland, and Lalor, 2017). However, it remains to  
447 be seen to which degree spectral cues and binaural cues are integrated in the auditory cortex, es-  
448 pecially in more naturalistic listening situations. The adapter-probe paradigm may help to resolve  
449 this question, because it allows attenuation of populations of auditory neurons without directly  
450 affecting spatial processing.

## 451 References

- 452 Batteau, Dwight W (1967). "The role of the pinna in human localization". In: *Proceedings of the Royal*  
453 *Society of London. Series B. Biological Sciences* 168.1011, pp. 158–180.
- 454 Bednar, Adam, Francis M Boland, and Edmund C Lalor (2017). "Different spatio-temporal electroen-  
455 cephalography features drive the successful decoding of binaural and monaural cues for sound  
456 localization". In: *European Journal of Neuroscience* 45.5, pp. 679–689.
- 457 Benda, Jan (2021). "Neural adaptation". In: *Current Biology* 31.3, R110–R116.
- 458 Bialas, Ole (2022). "freefield". In: DOI: [10.5281/zenodo.6027388](https://doi.org/10.5281/zenodo.6027388).
- 459 Bigdely-Shamlo, Nima et al. (2015). "The PREP pipeline: standardized preprocessing for large-scale  
460 EEG analysis". In: *Frontiers in neuroinformatics* 9, p. 16.
- 461 Bradski, G. (2000). "The OpenCV Library". In: *Dr. Dobb's Journal of Software Tools*.
- 462 Brugge, John F, Richard A Reale, and Joseph E Hind (1996). "The structure of spatial receptive fields  
463 of neurons in primary auditory cortex of the cat". In: *Journal of Neuroscience* 16.14, pp. 4420–  
464 4437.
- 465 Buzsáki, György, Costas A Anastassiou, and Christof Koch (2012). "The origin of extracellular fields  
466 and currents—EEG, ECoG, LFP and spikes". In: *Nature reviews neuroscience* 13.6, pp. 407–420.
- 467 Chapman, Robert M and John W McCrary (1995). "EP component identification and measurement  
468 by principal components-analysis". In: *Brain and cognition* 27.3, pp. 288–310.
- 469 Cheveigné, Alain de (2020). "ZapLine: A simple and effective method to remove power line artifacts".  
470 In: *NeuroImage* 207, p. 116356.
- 471 Cohen, Mike X (2014). *Analyzing neural time series data: theory and practice*. MIT press.
- 472 Coltheart, M (1980). "The persistences of vision". In: *Philosophical Transactions of the Royal Society*  
473 *of London. B, Biological Sciences* 290.1038, pp. 57–69.
- 474 Donchin, Emanuel (1966). "A multivariate approach to the analysis of average evoked potentials".  
475 In: *IEEE Transactions on Biomedical Engineering* 3, pp. 131–139.
- 476 Dreiseitl, Stephan and Lucila Ohno-Machado (2002). "Logistic regression and artificial neural net-  
477 work classification models: a methodology review". In: *Journal of biomedical informatics* 35.5-6,  
478 pp. 352–359.

- 479 Efron, Bradley (1992). "Bootstrap methods: another look at the jackknife". In: *Breakthroughs in statis-*  
480 *tics*. Springer, pp. 569–593.
- 481 Fujiki, Nobuya et al. (2002). "Human cortical representation of virtual auditory space: differences  
482 between sound azimuth and elevation". In: *European Journal of Neuroscience* 16.11, pp. 2207–  
483 2213.
- 484 Gaese, BH and A Johnen (2000). "Coding for auditory space in the superior colliculus of the rat". In:  
485 *European Journal of Neuroscience* 12.5, pp. 1739–1752.
- 486 Gramfort, Alexandre et al. (2013). "MEG and EEG Data Analysis with MNE-Python". In: *Frontiers in*  
487 *Neuroscience* 7.267, pp. 1–13. DOI: [10.3389/fnins.2013.00267](https://doi.org/10.3389/fnins.2013.00267).
- 488 Harper, Nicol S and David McAlpine (2004). "Optimal neural population coding of an auditory spatial  
489 cue". In: *Nature* 430.7000, pp. 682–686.
- 490 Hofman, Paul M, Jos GA Van Riswick, and A John Van Opstal (1998). "Relearning sound localization  
491 with new ears". In: *Nature neuroscience* 1.5, pp. 417–421.
- 492 Imig, Thomas J, W Andrew Irons, and Frank R Samson (1990). "Single-unit selectivity to azimuthal di-  
493 rection and sound pressure level of noise bursts in cat high-frequency primary auditory cortex".  
494 In: *Journal of neurophysiology* 63.6, pp. 1448–1466.
- 495 Jas, Mainak et al. (2017). "Autoreject: Automated artifact rejection for MEG and EEG data". In: *Neu-*  
496 *rolmage* 159, pp. 417–429.
- 497 Kayser, Jürgen and Craig E Tenke (2006). "Principal components analysis of Laplacian waveforms  
498 as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball  
499 tasks". In: *Clinical neurophysiology* 117.2, pp. 348–368.
- 500 — (2015). "On the benefits of using surface Laplacian (current source density) methodology in  
501 electrophysiology". In: *International journal of psychophysiology: official journal of the International*  
502 *Organization of Psychophysiology* 97.3, p. 171.
- 503 King, AJ and ME Hutchings (1987). "Spatial response properties of acoustically responsive neurons  
504 in the superior colliculus of the ferret: a map of auditory space". In: *Journal of neurophysiology*  
505 57.2, pp. 596–624.
- 506 Large, Mary-Ellen, Adrian Aldcroft, and Tutis Vilis (2005). "Perceptual continuity and the emergence  
507 of perceptual persistence in the ventral visual pathway". In: *Journal of neurophysiology* 93.6,  
508 pp. 3453–3462.
- 509 Magezi, David A and Katrin Krumbholz (2010). "Evidence for opponent-channel coding of interaural  
510 time differences in human auditory cortex". In: *Journal of neurophysiology* 104.4, pp. 1997–2007.
- 511 Makous, James C and John C Middlebrooks (1990). "Two-dimensional sound localization by human  
512 listeners". In: *The journal of the Acoustical Society of America* 87.5, pp. 2188–2200.
- 513 Maris, Eric and Robert Oostenveld (2007). "Nonparametric statistical testing of EEG-and MEG-data".  
514 In: *Journal of neuroscience methods* 164.1, pp. 177–190.
- 515 McAlpine, David, Dan Jiang, and Alan R Palmer (2001). "A neural code for low-frequency sound  
516 localization in mammals". In: *Nature neuroscience* 4.4, pp. 396–401.
- 517 Middlebrooks, John C and David M Green (1991). "Sound localization by human listeners". In: *Annual*  
518 *review of psychology* 42.1, pp. 135–159.
- 519 Middlebrooks, John C and Eric I Knudsen (1984). "A neural code for auditory space in the cat's  
520 superior colliculus". In: *Journal of Neuroscience* 4.10, pp. 2621–2634.
- 521 Middlebrooks, John C et al. (1994). "A panoramic code for sound location by cortical neurons". In:  
522 *Science* 264.5160, pp. 842–844.
- 523 Møller, Henrik et al. (1995). "Transfer characteristics of headphones measured on human ears". In:  
524 *Journal of the Audio Engineering Society* 43.4, pp. 203–217.
- 525 Musall, Simon et al. (2014). "Effects of neural synchrony on surface EEG". In: *Cerebral Cortex* 24.4,  
526 pp. 1045–1053.
- 527 Nelken, Israel and Eric D Young (1994). "Two separate inhibitory mechanisms shape the responses  
528 of dorsal cochlear nucleus type IV units to narrowband and wideband stimuli". In: *Journal of*  
529 *neurophysiology* 71.6, pp. 2446–2462.



- 530 Paszke, Adam et al. (2017). "Automatic differentiation in PyTorch". In.
- 531 Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning*  
532 *Research* 12, pp. 2825–2830.
- 533 Philiastides, Marios G and Paul Sajda (2006). "Temporal characterization of the neural correlates  
534 of perceptual decision making in the human brain". In: *Cerebral cortex* 16.4, pp. 509–518.
- 535 Philiastides, Marios G and Paul Sajda (2007). "EEG-informed fMRI reveals spatiotemporal charac-  
536 teristics of perceptual decision making". In: *Journal of Neuroscience* 27.48, pp. 13082–13091.
- 537 Plöchl, Michael, José P Ossandón, and Peter König (2012). "Combining EEG and eye tracking: identi-  
538 fication, characterization, and correction of eye movement artifacts in electroencephalographic  
539 data". In: *Frontiers in human neuroscience* 6, p. 278.
- 540 Rayleigh, Lord (1907). "XII. On our perception of sound direction". In: *The London, Edinburgh, and*  
541 *Dublin Philosophical Magazine and Journal of Science* 13.74, pp. 214–232.
- 542 Reiss, Lina AJ and Eric D Young (2005). "Spectral edge sensitivity in neural circuits of the dorsal  
543 cochlear nucleus". In: *Journal of Neuroscience* 25.14, pp. 3680–3691.
- 544 Salminen, Nelli H et al. (2010). "The neural code for interaural time difference in human auditory  
545 cortex". In: *The Journal of the Acoustical Society of America* 127.2, EL60–EL65.
- 546 Sandel, Thomas T et al. (1955). "Localization of sound from single and paired sources". In: *the Journal*  
547 *of the Acoustical Society of America* 27.5, pp. 842–852.
- 548 Sassenhagen, Jona and Dejan Draschkow (2019). "Cluster-based permutation tests of MEG/EEG  
549 data do not establish significance of effect latency or location". In: *Psychophysiology* 56.6, e13335.
- 550 Schechtman, Eitan, Talia Shrem, and Leon Y Deouell (2012). "Spatial localization of auditory stimuli  
551 in human auditory cortex is based on both head-independent and head-centered coordinate  
552 systems". In: *Journal of Neuroscience* 32.39, pp. 13501–13509.
- 553 Schönwiesner, Marc and Ole Bialas (2021). "s(ound)lab: An easy to learn Python package for design-  
554 ing and running psychoacoustic experiments." In: *Journal of Open Source Software* 6.62, p. 3284.
- 555 Town, Stephen M, W Owen Brimijoin, and Jennifer K Bizley (2017). "Egocentric and allocentric rep-  
556 resentations in auditory cortex". In: *PLoS Biology* 15.6, e2001878.
- 557 Trapeau, Régis and Marc Schönwiesner (2018). "The encoding of sound source elevation in the  
558 human auditory cortex". In: *Journal of Neuroscience* 38.13, pp. 3252–3264.
- 559 VanRullen, Rufin and Christof Koch (2003). "Visual selective behavior can be triggered by a feed-  
560 forward process". In: *Journal of Cognitive Neuroscience* 15.2, pp. 209–217.
- 561 Viola, Filipa Campos et al. (2009). "Semi-automatic identification of independent components rep-  
562 resenting EEG artifact". In: *Clinical Neurophysiology* 120.5, pp. 868–877.
- 563 Wightman, Frederic L and Doris J Kistler (1989a). "Headphone simulation of free-field listening. I:  
564 stimulus synthesis". In: *The Journal of the Acoustical Society of America* 85.2, pp. 858–867.
- 565 — (1989b). "Headphone simulation of free-field listening. II: Psychophysical validation". In: *The Jour-  
566 nal of the Acoustical Society of America* 85.2, pp. 868–878.
- 567 Yin, Tom CT (2002). "Neural mechanisms of encoding binaural localization cues in the auditory  
568 brainstem". In: *Integrative functions in the mammalian auditory pathway*. Springer, pp. 99–159.