NeuroImage 181 (2018) 683–691

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Neural tracking of auditory motion is reflected by delta phase and alpha power of EEG

Adam Bednar [a,*], Edmund C. Lalor [a,b]

[a] School of Engineering, Trinity Centre for Bioengineering, and Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland
[b] Department of Biomedical Engineering, Department of Neuroscience, and Del Monte Institute for Neuroscience, University of Rochester, Rochester, NY, USA

ARTICLE INFO

Keywords:
Auditory motion
Electroencephalography
Decoding
Sound localization
ITD
ILD

ABSTRACT

It is of increasing practical interest to be able to decode the spatial characteristics of an auditory scene from electrophysiological signals. However, the cortical representation of auditory space is not well characterized, and it is unclear how cortical activity reflects the time-varying location of a moving sound. Recently, we demonstrated that cortical response measures to discrete noise bursts can be decoded to determine their origin in space. Here we build on these findings to investigate the cortical representation of a continuously moving auditory stimulus using scalp recorded electroencephalography (EEG).

In a first experiment, subjects listened to pink noise over headphones which was spectro-temporally modified to be perceived as randomly moving on a semi-circular trajectory in the horizontal plane. While subjects listened to the stimuli, we recorded their EEG using a 128-channel acquisition system. The data were analysed by 1) building a linear regression model (decoder) mapping the relationship between the stimulus location and a training set of EEG data, and 2) using the decoder to reconstruct an estimate of the time-varying sound source azimuth from the EEG data. The results showed that we can decode sound trajectory with a reconstruction accuracy significantly above chance level. Specifically, we found that the phase of delta (<2 Hz) and power of alpha (8–12 Hz) EEG track the dynamics of a moving auditory object.

In a follow-up experiment, we replaced the noise with pulse train stimuli containing only interaural level and time differences (ILDs and ITDs respectively). This allowed us to investigate whether our trajectory decoding is sensitive to both acoustic cues. We found that the sound trajectory can be decoded for both ILD and ITD stimuli. Moreover, their neural signatures were similar and even allowed successful cross-cue classification. This supports the notion of integrated processing of ILD and ITD at the cortical level. These results are particularly relevant for application in devices such as cognitively controlled hearing aids and for the evaluation of virtual acoustic environments.

1. Introduction

We can effortlessly localize moving auditory objects in space, follow their changing position and direction of motion, as well as perceive the sound velocity. Nevertheless, how these sensitivities to moving sounds are represented along the auditory pathway is unclear.

It is known that spatial acoustic cues are already processed at the level of the brainstem. For example, it has been shown that neurons within the inferior colliculus (IC) are sensitive to interaural time and level differences (ITDs and ILDs; see review by Grothe et al. (2010). However, it is uncertain whether motion analysis occurs at this level. Spitzer and Semple (1993) showed that sound location sensitivity of IC neurons depends on the previous sound position, and that the neurons are sensitive to sound direction. One could interpret this finding as sensitivity of IC to auditory motion. Nevertheless, McAlpine et al. (2000) argued that these effects might be due to neuronal response adaptation and not to motion sensitivity per se.

At the level of cortex, neuroimaging studies have indicated that non-primary regions of auditory cortex, and in particular planum temporale and posterior STG are sensitive to moving sound stimuli (Baumgart et al., 1999; Lewis et al., 2000; Pavani et al., 2002; Warren et al., 2002; Krumbholz et al., 2005a). However, these structures often respond to spatial static sounds and so it is still under debate as to whether they represent true motion detectors (Ducommun et al., 2002; Smith et al.,

* Corresponding author. University of Rochester, 500 Joseph C. Wilson Blvd. Box 270168, Rochester, NY, 14611, USA.
E-mail address: bednara@tcd.ie (A. Bednar).

https://doi.org/10.1016/j.neuroimage.2018.07.054
Received 2 June 2018; Received in revised form 10 July 2018; Accepted 23 July 2018
Available online 24 July 2018

2007; Getzmann and Lewald, 2012; Poirier et al., 2017).

In the abovementioned neural structures, it is thought that auditory space is encoded by broadly tuned neurons within each hemisphere of auditory cortex, which are activated primarily by sounds coming from the contralateral hemifield. This was shown in mammals (Middlebrooks and Pettigrew, 1981; Stecker et al., 2003; Werner-Reiss and Groh, 2008; Ortiz-Rios et al., 2017), and in humans (Palomaki et al., 2000, 2005; Krumbholz et al., 2005b; McLaughlin et al., 2015; Derey et al., 2016). At the more fundamental level, there is uncertainty as to whether this contralateral neural tuning applies both to ILD and ITD cues (Johnson and Hautus, 2010; McLaughlin et al., 2015), and whether these cues are processed independently at the level of cortex (Ungan et al., 2001; Altmann et al., 2007; Johnson and Hautus, 2010; Edmonds and Krumbholz, 2014; Higgins et al., 2017). It is also unclear how this distributed opponent encoding might reflect the dynamics of sound motion. EEG studies have shown that the brain responds to a sudden stimulus motion onset (MOR) (Krumbholz et al., 2007; Getzmann and Lewald, 2010; Getzmann, 2011); or to an onset of mismatching moving stimuli (MMN) (Altman et al., 2010; Shestopalova et al., 2012). However, it remains unknown how cortex tracks spatial changes in more complex auditory situations, where sound source movements are continuous. In particular, it is unclear which patterns of neural activity reflect the temporal dynamics of a moving sound source.

Here, we seek to investigate this using EEG. In particular, we aim to use a model-based decoding approach to attempt to learn which EEG features encode information about the moving stimulus by attempting to reconstruct an estimate of the azimuth of that stimulus from the recorded data. Previously, we have shown that we can successfully predict positions of static noise bursts using a linear classifier (Bednar et al., 2017). However, to the best of our knowledge this is the first study to attempt to "decode" EEG signals to determine the time-varying location of a moving auditory input. Such decoding approaches have been used previously for stationary auditory inputs, based on their intensity (Lalor et al., 2009; Mesgarani et al., 2009) and to assess how attention is deployed in complex environments (Ding and Simon, 2012; O'Sullivan et al., 2015). The present work complements these findings, particularly insofar as it could be combined with some of these aforementioned measures in attempts to realize "cognitively-steered" hearing aids or in the evaluation of virtual acoustic environments.

2. Material and methods

Participants. In total thirty-three participants (median = 21 years; min = 18 years; max = 27 years; 22 females; 26 right handed) participated in this study with informed consent. Sixteen subjects took part in the first experiment and seventeen in the second. All subjects reported no neurological diseases and normal hearing. The experiments were conducted in accordance with the Declaration of Helsinki and were approved by the Research Subjects Review Board of the University of Rochester.

Experimental procedure. Participants listened to auditory stimuli (described below) presented via headphones while performing a simple target detection task. The task required subjects to respond with a button press to infrequent tremolo targets (modulation frequency 4 Hz, 2 s long), which were embedded in the stimuli. The number of targets within each trial ranged from 1 to 6 per trial. During the experiment, the subjects sat in a dark soundproof room. To minimize movement, the participants were asked to look at a fixation cross displayed on a computer screen directly in front of them. In both experiments, subjects undertook a total of 24 trials. Each trial lasted 3 min and consisted of one continuous sound stimulus that was perceived as randomly moving within the horizontal plane. The stimuli were different for the first and the second experiment.

Stimuli. In both experiments, the auditory stimulus was perceived as moving within the frontal part of the horizontal plane. The trajectory was pseudo-random, simulating smooth but unpredictable sound movement. The stimulus and method of sound spatialization was different for the first and the second experiment.

In the first experiment, the sound stimuli consisted of continuous pink noise with frequency roll-off 10 dB/decade. The sound source motion was implemented by virtual acoustic space (VAS) using Oculus Audio SDK, which simulates head-related transfer function (HRTF) filtering. The sound was spatialized so that it was perceived to be pseudo randomly moving on a semi-circular trajectory in the horizontal plane between -90° (left) and $+90^\circ$ (right) relative to the subject (see Fig. 1A). The simulated motion of the source had an average angular velocity of 80°/second. Each trial (sound stimulus) lasted exactly 180 s.

In the second experiment, we used trains of short pink noise bursts. The noise bursts were presented at a rate of 100 Hz, each burst lasted 1 ms and the silence interval between the bursts was 9 ms (see Fig. 1B). The stimuli were spatialized using either interaural level differences (ILDs) or interaural time differences (ITD). Similar to the first experiment, the acoustic cues were manipulated so that the sound appeared to be moving smoothly in a pseudo-random manner between the left and the right ear. The ITD cues varied between ± 750 μs and ILD cues within the range of ± 20 dB. From the total of 24 trials, 12 trials were spatialized using only ILDs and 12 trials contained only ITDs. The ILD and ITD trials were alternating. As in the first experiment, all trials were 180 s long.

EEG data preprocessing. The EEG data were recorded using a 128-channel ActiveTwo acquisition system (BioSemi, The Netherlands) at a sampling rate of 512 Hz. Preprocessing was done in MATLAB using custom written scripts and the EEGLAB toolbox (Delorme and Makeig, 2004). The data were filtered between 0.02 and 30 Hz and downsampled to 64 Hz. Bad channels were interpolated from the surrounding channels using the spline function from EEGLAB. Finally, the data were re-referenced to the average of all electrodes.

Data analysis. In order to assess if and how the location of the moving stimuli was represented in the EEG, we used a multivariate linear reconstruction model g to reconstruct the sound stimulus trajectory S from the neural data R (Crosse et al., 2016). This mapping can be described as:

$$\widehat{S}(t) = \sum_{n=1}^{N} \sum_{\tau} g(\tau, n) R(t+\tau, n),$$

where $\widehat{S}(t)$ is the reconstructed estimate of the stimulus position at time $t = 1 \ldots T$, $g(\tau, n)$ is the decoder model which is a function of the time lag τ and the electrode channel $n = 1 \ldots N$, and $R(t, n)$ is the neural response at time t and electrode n.

The stimulus trajectory, $\widehat{S}(t)$, is a measure of the azimuth, where azimuth of $+90^\circ$ corresponds to the right and -90° corresponds to the left (see Fig. 1A). The decoder $g(\tau, n)$ integrates EEG over time lags τ from 0 ms to 250 ms poststimulus to reconstruct each sample of the stimulus trajectory. This range of lags was selected as it typically encompasses major cortical ERP components to an auditory stimulus. As indicated by the equation, the decoder $g(\tau, n)$ is essentially a multivariate impulse response function calculated from all 128 EEG electrodes and all time-lags simultaneously and the stimulus trajectory is estimated by convolving this impulse response with the EEG data.

Leave-one-out cross-validation was used as a performance measure of our stimulus reconstruction. We fitted the model (decoder) on all but one trial and then evaluated this decoder on the remaining left out trial. Afterwards, we measured the similarity between the reconstructed and the original sound trajectory using Pearson's r. We repeated this for each trial. These correlation coefficients were then averaged across trials for each subject.

To measure the within-subject statistical significance of our decoding, we used a non-parametric permutation approach (Combrisson and Jerbi, 2015). First, we established a null distribution of the predictions by running the decoding process 1000 times with randomly permuted stimulus trajectories. After, we used the tail of this empirical distribution to calculate the p-value for the original classification. This was done for trial-averaged data and for each subject separately.

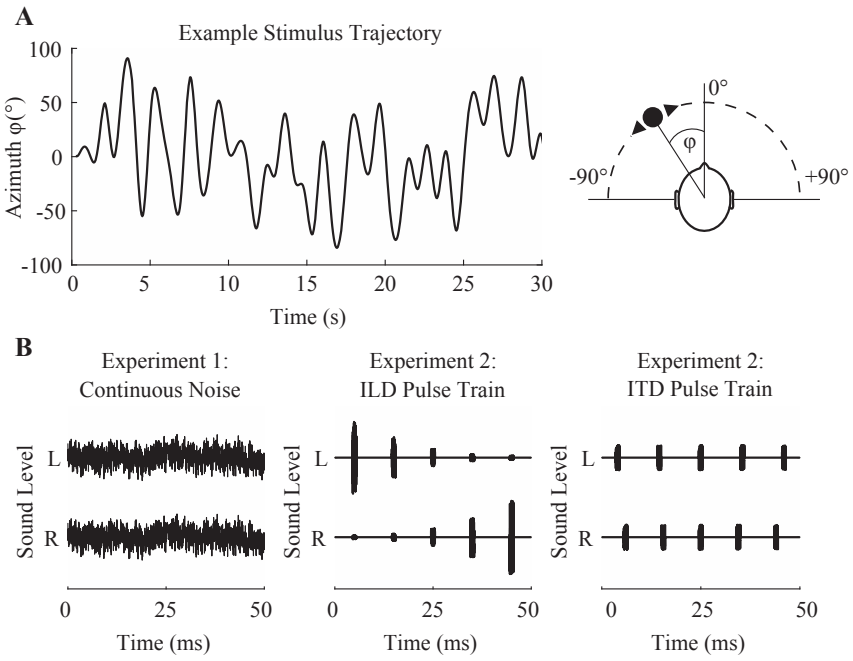

Fig. 1. Continuously moving stimuli. In Experiment 1, the pink noise stimulus was HRTF-filtered to be perceived as randomly moving on a semi-circular trajectory around the listener between the left and right. In Experiment 2, the noise stimuli were replaced by ILD- and ITD-only spatialized pulse trains. (A) Example of the sound trajectory. The sound source simulated random and smooth motion between the left (−90 deg) and the right (+90 deg) ear. (B) Sound stimulus waveform examples. Left: HRTF-filtered pink noise was used in Experiment 1. Center: Experiment 2- ILD condition used a pulse train with embedded level differences. Right: Experiment 2- ITD condition used a pulse train with introduced time differences between pulses.

All statistical comparisons on a group level were conducted using two-sided Wilcoxon signed-rank tests.

3. Results

3.1. Experiment 1: reconstructing azimuth of moving sound

We investigated whether the scalp recorded EEG reflects the dynamics of a moving sound source in the horizontal plane. As we have done in the past for other sound features e.g. sound intensity (O'Sullivan et al., 2015), we used the decoding approach to reconstruct the time-varying azimuth of the moving sound. In the first experiment, the subjects listened to spatialized continuous pink noise. We used VAS to simulate a moving sound source and so listeners had both binaural cues available, as well as spectral cues.

Predicting stimulus trajectory from broadband EEG. Initially, we tested the feasibility of reconstructing the sound trajectory using broadband EEG (0.02–30 Hz). As input to the decoder, we used all 128 EEG channels and lags within the range of 0 and 250 ms.

We used Pearson's r to compare the reconstructed sound trajectories with the original ones ('decoding'). The mean reconstruction correlation averaged across all subjects and trials was $r = 0.076$. The mean chance-level correlation ('control') was $r = 0.008$ and was computed by averaging the permutation test correlations across all trials and repetitions (Fig. 2A). On a group-level, the decoding correlation values were significantly larger than the control condition ($p = 4.4e\text{-}4$). See examples

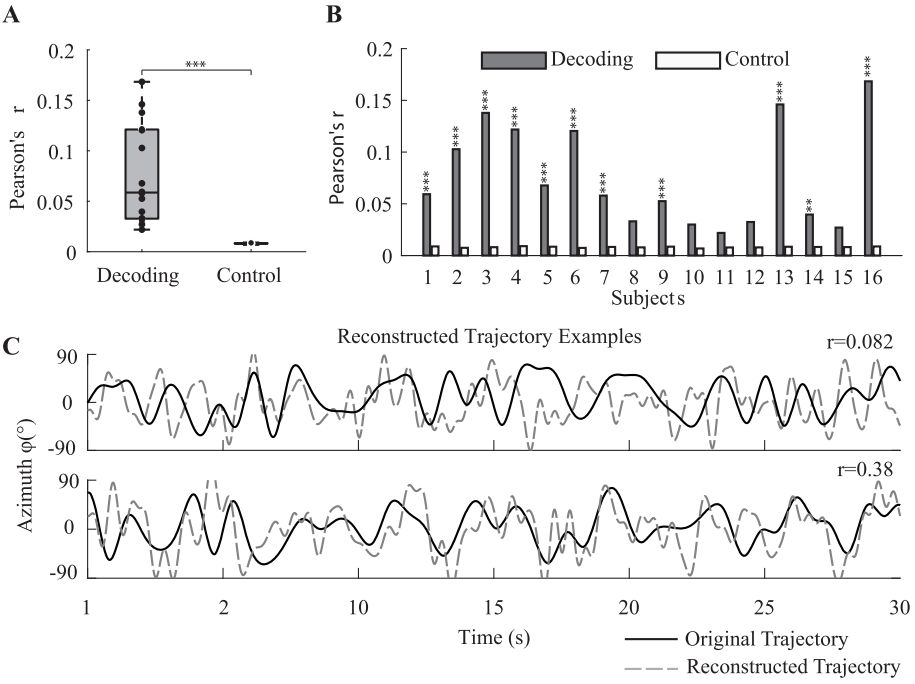

Fig. 2. Azimuth reconstruction from broadband EEG. (A) Decoding results shown as an average across subjects and trials. The decoding performance is represented as Pearson's r between the reconstructed and the original sound trajectory. As a control condition, to estimate whether our decoding is above the chance-level, we trained the decoders with randomly permuted trajectories between the trials. We found there is a significant difference between the decoding and the control condition ($p < 0.001$). (B) Decoding results shown for each individual subject. The black bars show the reconstruction correlation values averaged across all trials. The white bars show the correlation values corresponding to the permutation test ('control'). (C) Examples of original sound stimulus trajectories along with their reconstructions shown for two different trial segments of one subject. The correlation values between the original and reconstructed trajectories are shown above the plots.

of reconstructed trajectories in Fig. 2C. Although the reconstruction correlation values are relatively low, it is important to point out that they represent stimulus trajectory predictions from novel unaveraged EEG and should be interpreted in this context given that background EEG can be an order of magnitude larger than stimulus evoked activity.

We also ran the analysis on a single-subject level. For all subjects, the correlations averaged across all trials were higher than the control condition, and for eleven out of sixteen subjects, this difference was significant (see Fig. 2B).

Decoding at different EEG bands. The next stage of our analysis focused on investigating what EEG frequencies are best for stimulus angle reconstruction. We extracted different EEG bands using a sliding bandpass filter from 0.02 to 30 Hz which had a passband window with a width of 2 Hz. We then ran our decoding analysis using each of these EEG bands (Fig. 3A). The results revealed that the decoding is driven by the lowest frequency EEG delta band (0.02–2 Hz). We also ran the analysis after calculating the analytic envelope of the EEG signal using the Hilbert transformation following the bandpass filtering. In this case, the greatest decoding accuracy was achieved within the alpha band of the EEG (i.e., the EEG power in the range 8–12 Hz). In general, the performance decreased with increasing frequency, except for a small improvement around 18 Hz. In the following analyses, we therefore focused on the delta and alpha power EEG decoders, from now on abbreviated as 'Delta' and 'Alpha Pwr'.

The decoding results for Delta and Alpha Pwr decoders are shown in Fig. 3B. The average reconstruction accuracies for both the Delta and Alpha Pwr were significantly above the chance level r = 0.08, p = 4.3e-04 and r = 0.046, p = 4.3e-04 respectively.

The Delta decoder performed better than the Alpha Pwr decoder (p = 9.7e-3). The results of single subject analysis are shown in Appendix A. For the Delta decoder, eleven subjects had decoding accuracy significantly above the chance level. For the Alpha Pwr decoder, eight out of sixteen subjects showed significant decoding on single subject level. In comparison to broadband EEG decoders, both the Delta decoder and the Alpha Pwr decoder had higher reconstruction accuracies than the broadband EEG decoder (p = 4.3e-4 and p = 5.2e-3 respectively).

To test whether delta and alpha power EEG carry independent information about the source azimuth, we combined the Delta and Alpha Pwr decoders. This was done by performing the regression using 256 channels (128 channel delta EEG + 128 channels alpha power EEG). Then we compared its performance with each of the decoders separately. On average, this combined 'Delta + Alpha Pwr' decoder performed better than the individual Delta and Alpha Pwr decoders. However, this improvement was only significant when compared with Alpha Pwr decoder (p = 4.3e-4), not with the Delta decoder (p = 0.063). See results in Fig. 3B.

Spatiotemporal decoder characteristics. In order to obtain some insight into the possible cortical areas involved in encoding spatial information, we examined which scalp regions were modulated by sound source position and driving our decoding. To do this we plotted the forward transformations of the decoder weights 'activation patterns', which are more interpretable in terms of the underlying physiology (Haufe et al., 2014). As shown in Fig. 3C, the Delta decoder showed strong activations over temporal scalp bilaterally, indicating the likely involvement of auditory cortex. The activation patterns were opposite between the hemispheres, i.e., an increase in activity in one hemisphere was accompanied by a decrease in the other. The Alpha Pwr decoder indicated more posterior activation. The stimulus caused a relative decrease in parietal activity contralateral to the stimulus position, and a relative increase over ipsilateral scalp.

To further compare the Delta and Alpha Pwr markers of spatial audio encoding, we then evaluated the two performances of the two decoders at individual time lags. The temporal window of lags where the decoding was significantly above chance was relatively long. For the Delta decoder, the stimulus angle reconstruction was best when EEG was lagging the stimulus by 78 ms. For the Alpha Pwr decoder the responses were more delayed and decoding peaked at a lag of 187 ms (see Fig. 4A). The single-lag decoder activation patterns are shown in Fig. 4B.

We also investigated if we can reconstruct sound azimuth using EEG channels from a single hemisphere and tested whether this 'simulated lesion' results in a degradation of decoding performance in comparison to using information from the entire scalp. We did this by training our decoders on electrode subsets (n = 55) covering either left or right scalp, 'L-hemi' and 'R-hemi', and comparing the results with the decoder that was trained on electrodes covering the entire scalp 'Full Scalp'. To control for the greater number of electrodes in the 'Full Scalp' decoder, this decoder was trained 1000-times on a random subset of 55 electrodes from the entire scalp and the decoding results were averaged together. All midline electrodes (n = 18) were omitted.

For both 'L-hemi' and 'R-hemi' decoders using both delta and alpha EEG bands, the azimuth reconstructions were significantly better than chance level (all p < 0.005). Also, the results showed that the 'Full Scalp' decoders were significantly better than the decoders trained on a single hemisphere. This was true for 'L-hemi' and 'R-hemi' using delta filtered EEG (p = 8e-4 and p = 0.016) and for the power of alpha EEG (p = 4.5e-3 and p = 6.1e-3). See decoding results in Fig. 5.

Also, previous literature has suggested a possible asymmetric representation of auditory spatial information across the hemispheres (Kaiser

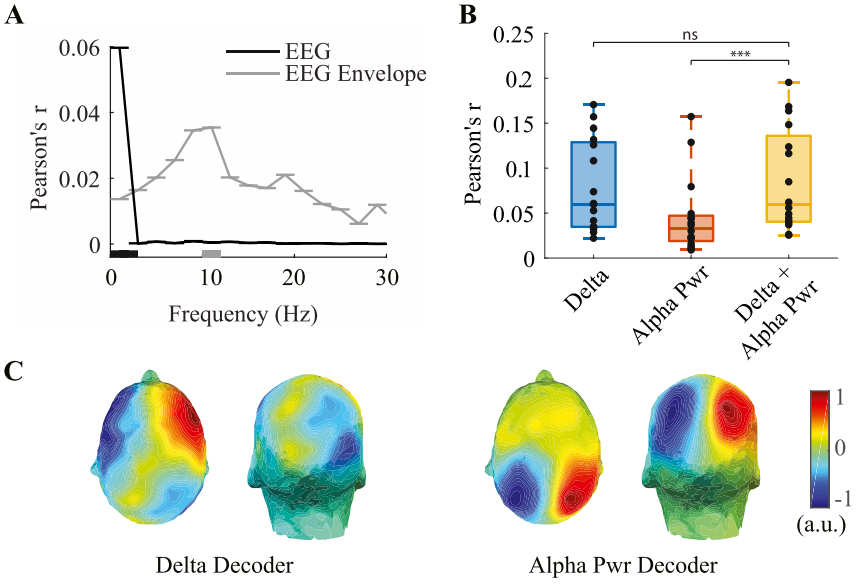

Fig. 3. Decoding azimuth for different EEG bands. (A) The decoding performance dependency on frequency bandwidth of raw EEG signal (black) and EEG envelope (grey). Significant frequency windows are marked using thick lines at the bottom of the plot (p < 0.001). (B) Azimuth reconstruction accuracies shown for Delta (blue), Alpha Pwr (red) and Delta + Alpha Pwr (yellow) decoders. All reconstruction accuracies were significantly above the chance level (p < 0.001). *** indicates reconstruction differences at the level of p < 0.01. (C) Head plots of decoder activation patterns are shown for Delta and Alpha Pwr decoders.

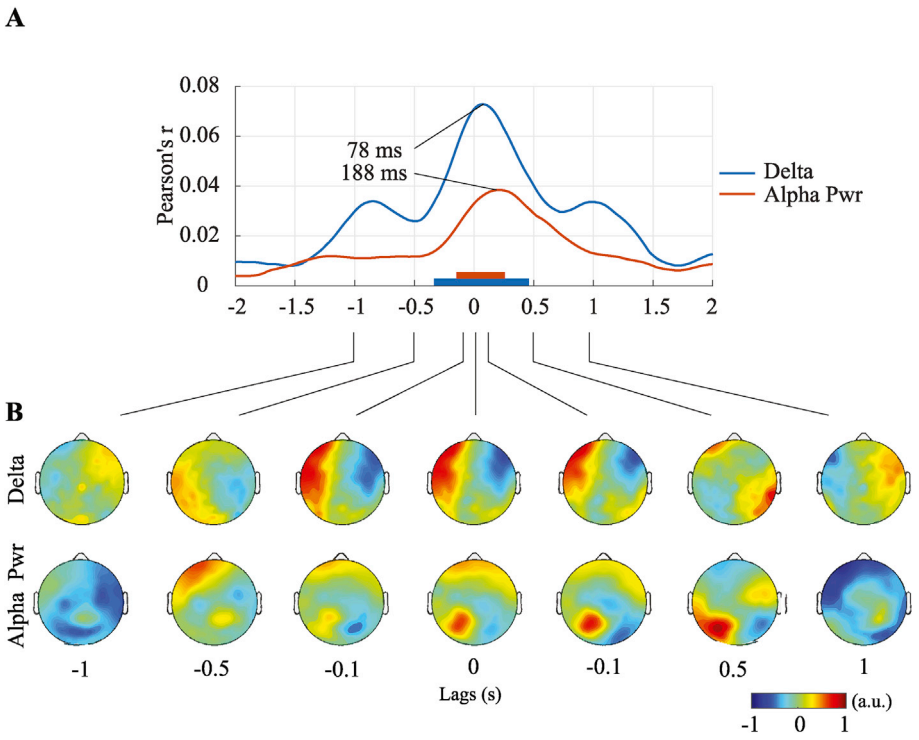

Fig. 4. Single lag decoding with decoder activation patterns (A) Single lag decoding shows reconstruction correlations as a function of time lag between stimulus and EEG shown for Delta (blue) and Alpha Pwr (red) decoders. Significant time lags are marked using thick lines at the bottom of the plot ($p < 0.001$). (B) Decoder activation patterns at selected time lags shown for Delta and Alpha Pwr decoders.

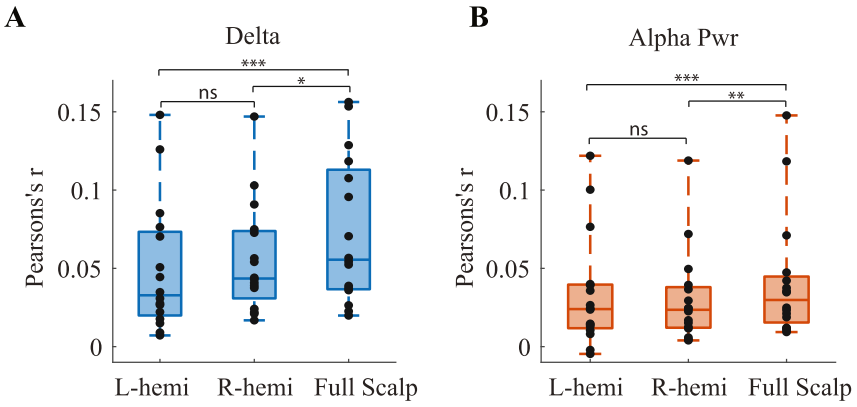

Fig. 5. Reconstructing stimulus azimuth using a single hemisphere shown for Delta (A) and Alpha Pwr decoders (B). The decoders were trained using a subset of 55 EEG electrodes corresponding to the left hemisphere 'L-hemi' or the right hemisphere 'R-hemi'. 'Full Scalp' represents a decoder that was trained using 55 randomly selected EEG electrodes from the entire scalp covering both hemispheres (shown as an average of 1000 randomly selected electrode subsets). * ** *** indicates prediction differences at the level of $p < 0.05$, $p < 0.01$ and $p < 0.005$ respectively.

et al., 2000; Zatorre and Penhune, 2001; Krumbholz et al., 2005b). While our decoder activation patterns (Fig. 3C) didn't show any obvious lateralization, we decided to test this formally by comparing the reconstruction accuracies between the left and the right hemisphere. The accuracies of these different reconstructions were not significantly different for either delta ($p = 0.26$) or the power of alpha EEG ($p = 0.72$). See Fig. 5.

Sensitivity to specific position in azimuth. Finally, we tested whether the EEG tracking of sound position relies on specific position in space or whether it only exhibits sensitivity to stimulus laterality i.e., stimulus on the left or right. We used a forward modelling approach (see Crosse et al., 2016), in which we tried to predict previously unseen EEG data from the stimulus trajectory. Specifically, we used either full trajectory of the moving stimulus 'Full' (continuously varying angle between −90° and +90°, as shown in Fig. 1A) or left/right-only trajectory 'L/R' (rectangular waveform indicating sound position within left or right hemifield).

The EEG prediction accuracies for both Full and impoverished L/R trajectories are shown in Fig. 6A. The comparison of EEG prediction accuracies averaged across all electrode channels show we can predict the EEG signal significantly better using the Full than the L/R trajectory. This was shown for both Delta ($p = 6.4e\text{-}4$) and Alpha Pwr EEG ($p = 2.3e\text{-}3$). The scalp distributions of EEG prediction accuracies for Full and L/R trajectories are shown in Fig. 6B.

Correlation with behaviour. We compared the azimuth reconstruction accuracy with the performance in the tremolo target detection task, which subjects performed during the experiment. On a group level, there was no significant correlation between the trajectory reconstruction accuracy and behaviour performance ($r = -0.21$, $p = 0.44$) and ($r = -0.28$, $p = 0.31$) for Delta and Alpha Pwr decoders respectively.

3.2. *Experiment 2: reconstructing ITD/ILD of moving stimuli*

In the second experiment, we examined which features of the acoustic signal were driving our stimulus trajectory reconstruction. We did so by running the decoding on spatially impoverished stimuli that contained either only ITD or only ILD acoustic cues. Similar to the first experiment

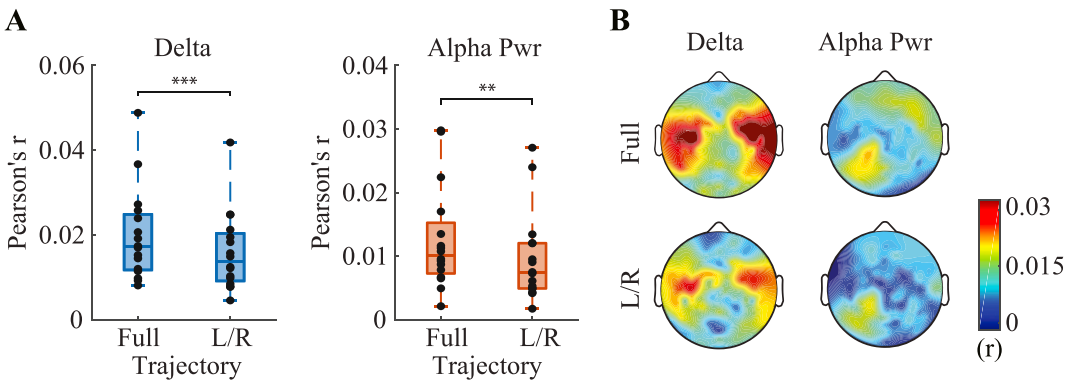

Fig. 6. Predicting EEG using full trajectory and left/right-only trajectory. To test whether the EEG tracking of sound position is sensitive to specific position in azimuth rather than just left/right lateralization, we used a forward modelling approach to predict previously unseen EEG from the trajectory data. The sound position was either represented as full trajectory 'Full' (continuously varying angle between −90° and +90°) or left/right-only trajectory 'L/R' (rectangular waveform indicating sound position within left or right hemifield). (A) Comparison of EEG prediction accuracies averaged across all electrodes shown for both trajectories and Delta and Alpha Pwr EEG. ** and *** indicates prediction differences at the level of $p < 0.01$ and $p < 0.005$ respectively. (B) Topographical plots indicating EEG prediction accuracies across scalp shown for both trajectories and Delta and Alpha Pwr EEG.

we simulated smooth movement of the sound in the horizontal plane by manipulating these acoustic cues.

We used the same approach as in the first experiment and we attempted to reconstruct stimulus position (time series of the ILD/ITD stimulus azimuth) from delta band and alpha power of EEG. We found that both ILD and ITD stimuli can be decoded with reconstruction accuracy significantly above the chance level. For Delta decoder, the mean reconstruction correlations of ILD and ITD stimuli were $r = 0.050$, $p = 9.9e\text{-}3$ and $r = 0.045$, $p = 0.017$. For Alpha Pwr decoder, the reconstruction values were $r = 0.042$, $p = 4.2e\text{-}4$and $r = 0.034$, $p = 3.1e\text{-}3$ for ILD and ITD respectively (see Fig. 7A). The results of single subject analysis are shown in Appendix B.

ILD/ITD decoder activation patterns. To characterize the spatial distribution of neural responses to each acoustic cue, we calculated the decoder activation patterns for ILD and ITD decoders (Fig. 7B). For both ITD and ILD trials, the Delta decoder weights had a lateralized distribution and the activation patterns had an opposite polarity between the hemispheres. Although being noisier, the activation patterns, and especially the patterns corresponding to ILD cue processing, resembled the activation patterns seen in the first experiment. For the Alpha Pwr decoders, the lateralized responses were again more posterior, showing opposite patterns between the left and the right hemispheres. In comparison to the ILD decoder, the ITD weights are less central and located closer to the temporal areas.

Cross-modal decoding. We observed that the activation patterns for ILD and ITD decoders were relatively similar. We attempted to quantify this similarity by performing cross-cue classification e.g. evaluating ILD-trained model performance on ITD trials and ITD-trained model performance on ILD trials.

When evaluating the ILD-trained decoder on ITD data, the reconstruction accuracies were $r = 0.045$, $p = 0.025$ and $r = 0.032$, $p = 3.6e\text{-}3$ for Delta and Alpha Pwr. When evaluating ITD-trained decoder on ILD

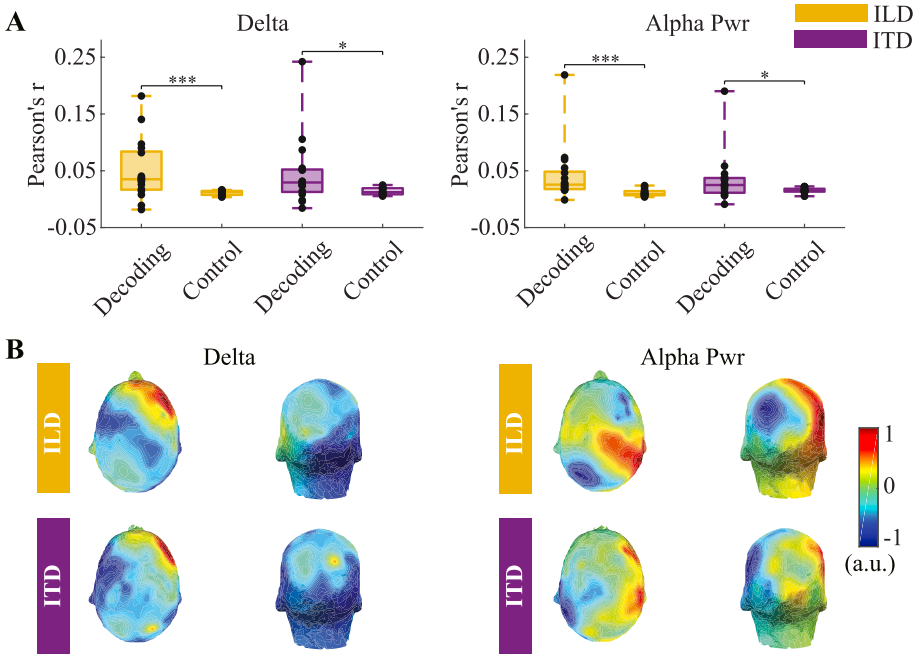

Fig. 7. Decoding ITDs and ILDs. (A) Reconstruction accuracy of ILD and ITD trials shown for delta (0.02–2 Hz) and alpha power EEG (8–12 Hz) shown along with chance level correlation values ('control'). * ** *** indicates classification with decoding accuracy significantly above the chance level for $p < 0.05$ $p < 0.01$ and $p < 0.005$ respectively. (B) Head plots of Delta and Alpha decoder activation patterns shown for ILD and ITD trials.

trials, the decoding results were r = 0.051, p = 1.2e-3 for Delta and r = 0.033, p = 0.044 for Alpha Pwr. The reconstruction accuracies were significantly above the chance level for all conditions. Pair-wise comparisons between the within-modal (e.g. ILD decoder evaluated on ILD data) and cross-modal reconstruction (e.g. ILD decoder evaluated on ITD trials) are shown in Fig. 8. On a group level, there were no significant differences in reconstruction accuracies between the within-modal and cross-modal decoding.

Correlation with behaviour. We compared the azimuth reconstruction accuracy for both ILD and ITD trials with the performance in the tremolo target detection task. For ILD, there was no significant correlation between the trajectory reconstruction accuracy and behaviour performance (r = 0.34, p = 0.18) and (r = 0.21, p = 0.41) for Delta and Alpha Pwr decoders respectively. For ITD, there was a significant correlation between azimuth reconstruction and behaviour for Delta (r = 0.55, p = 0.02) but not for Alpha Pwr (r = 0.30, p = 0.23) decoders.

4. Discussion

EEG decoding reveals cortical tracking of auditory motion. Recent studies have shown that cortex is sensitive to sound position, yet it remains unclear to what extent neural activity reflects the on-going dynamics of a moving sound. Here, we demonstrated that it is possible to decode the trajectory of a moving sound source from low frequency delta EEG (0.02–2 Hz) and by the power of alpha EEG (8–12 Hz). This suggests that cortical activity tracks the time-varying azimuth of a moving sound. We also showed that this tracking does not rely only on the stimulus lateralization i.e., whether sound is on the left or right side of the acoustic space but is rather sensitive to specific sound source location in azimuth.

Auditory motion is likely encoded in auditory cortex. We found that the low-frequency delta EEG (<2 Hz) track the fluctuating stimulus azimuth. This is somewhat in line with previous studies, which demonstrated that low frequency EEG oscillations (<10 Hz) entrain to temporal modulations of sensory input. Specifically, in the auditory domain, this was shown for the speech envelope (Lalor and Foxe, 2010; Luo et al., 2010).

One reason why our decoding performs well in this relatively low and narrow frequency band might be explained by the frequency characteristics of the stimulus. The stimulus consisted of pink noise that was spatialized to be perceived as moving smoothly around the listener. To ensure that the stimulus movement would be comfortable to follow and perceived as continuous, the maximum angular velocity was kept under 280 deg/s. Translating this into the frequency domain, the spectral analysis of the stimulus trajectory time-series showed that the majority of signal power is distributed below 1 Hz (see appendix C). And if, as the results indicate, the cortical activity entrains to the temporally varying sound azimuth in a (quasi-) linear manner, then we would expect that the frequency band of EEG that encodes the stimulus would correspond to the frequency bandwidth of the stimulus trajectory.

The delta EEG decoder weights indicated a contralateral bias. The topographical distribution was relatively symmetrical between the hemispheres but the polarity of weights was opposite between the left and the right hemisphere. Similar to this, predominately contralateral tuning was also shown for lateralized sounds in animals (Stecker et al., 2003, 2005; Werner-Reiss and Groh, 2008; Ortiz-Rios et al., 2017; Poirier et al., 2017), as well as in humans (Palomaki et al., 2000, 2005; Fujiki et al., 2002; Pavani et al., 2002; Getzmann and Lewald, 2010; Lewald and Getzmann, 2011; Derey et al., 2016).

Using single time-lags to decode the trajectory showed that the sound position is best reconstructed when delta filtered EEG is lagging the stimulus trajectory by approximately 100 ms. The cortical sensitivity to sound azimuth at this latency was also reported by previous studies (Palomaki et al., 2000, 2005; Fujiki et al., 2002; Lewald and Getzmann, 2011; Bednar et al., 2017). This would correspond to the latency of the N100 waveform in classical ERP analysis, which is generated from Heschel's gyrus and planum temporale (Godey et al., 2001).

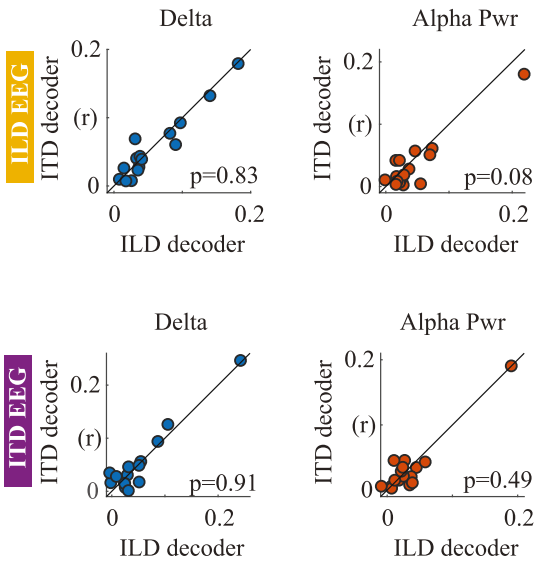

Fig. 8. Comparison of within modal and cross-modal classification for Delta and Alpha Pwr decoders. The cross-modal classification was done by evaluating ILD-trained model on ITD trials and ITD-trained model on ILD trials. The p-values correspond to pairwise comparison between the within-modal and cross-modal decoders.

The second EEG measure that was found to track the sound azimuth was the parieto-occipital alpha power, which is often related to deployment of selective attention (Kerlin et al., 2010; Ahveninen et al., 2013; Wöstmann et al., 2016). Unlike these studies, our experiment did not explicitly employ selective attention as the subjects listened only to a single moving sound stimulus and no competing audio stream was present. However, our results are compatible with recent studies that reported lateralization of occipital alpha oscillations in response to salient but spatially unpredictable sounds (Störmer et al., 2016; Feng et al., 2017).

Different spatio-temporal characteristics of delta and the power of alpha EEG encoding suggests they potentially reflect different aspects of auditory motion processing. We tested this, by combining the Delta and Alpha Pwr decoders and comparing that with the results of the Delta and Alpha Pwr decoders individually. More specifically, if we could show that combining the two decoders led to improved decoding performance, we could argue that the two neural signatures carried complementary information. On average, the combined decoder performed better than the Delta and Alpha decoders separately. However, this result was only significant in comparison to the Alpha Pwr decoder. Therefore, we were unable to confirm if delta and alpha EEG each carry different information about the sound source position.

We then investigated the involvement of each hemisphere in sound location encoding. As shown in a human lesion study (Zatorre and Penhune, 2001), a unilateral lesion caused impairment in sound localization ability. In line with that, we found that the azimuth reconstruction accuracies decreased when we trained the decoders on a subset of EEG channels covering either left or right scalp only. However, the performances of decoders that were trained on a single hemisphere were still significantly higher than the chance level. This suggest that the sound source position is at least partially encoded in a single hemisphere, however, the inter-hemispheric interactions are necessary to fully describe the sound source position. This is also compatible with the opponent-channel model (Stecker et al., 2005), which predicts that sound location is calculated based on a difference of contra- and ipsilaterally tuned neural channels and assumes that both contralateral and ipsilateral channels are present within each hemisphere.

Interestingly, the comparison of decoding performance between hemispheres did not reveal any significant lateralization, which is in

contrast with a number of studies that suggested a larger involvement of right hemisphere in auditory localization (Burke et al., 1994; Kaiser et al., 2000; Zatorre and Penhune, 2001; Palomaki et al., 2005). However, our lack of any inter hemispheric differences may be due to the limited spatial sensitivity of our approach.

Are we decoding sound position or just fluctuations in the envelope of the sound in each ear?. It is possible that the decoding in our first experiment was not driven by sound position per se but by the fluctuating envelope of the signal that was caused by ILDs. In the second experiment, we sought to test this by investigating if we can decode sound position independently based on the type of the acoustic localization cue. Analogous to the first experiment, we simulated a smoothly moving sound source in the horizontal plane. However, the stimuli were spatially impoverished and contained either ILD or ITD cues.

The results showed that we can indeed reconstruct the trajectory of a sound source containing only ILD or ITD cues from the EEG with accuracy significantly above the chance level. This indicates that we decoded genuine sound motion trajectory and not only sound envelope changes that correspond to varying ILDs. Interestingly, the performances of ILD and ITD decoders were relatively similar. The correlation values were generally lower than in the first experiment, which could be attributed to several factors. First, the spatially impoverished character of the stimuli possibly caused a reduction in neural responses compared with the stimuli that contained all forms of sound localization cues. This was shown in previous MEG studies (Palomaki et al., 2000, 2005). Another possible reason for the reduction in accuracy may be that there were two conditions (ILD and ITD) in the latter experiment and so the decoders were trained on only half the number trials. Indeed, when sub-sampling half the number of trials of Experiment 1 and comparing the decoding performance results with Experiment 2, we found only significant differences for delta EEG and ITD trials ($p = 0.048$) but not for ILD trials and there were no differences for alpha filtered EEG (all $p > 0.05$).

ITD and ILD cues may not be represented independently at the level of cortex. To determine whether or not ILD and ITD are represented differently in the EEG activity, we performed cross-cue decoding. A similar approach to this was used in an fMRI study by (Higgins et al., 2017). The results presented here show that we can successfully exchange ILD- and ITD-trained models without any significant change in reconstruction accuracy, despite only approximate ILD-ITD mapping. This suggests that activity at the cortical level represents sound location in an acoustic cue-independent manner. Or, considering the possibility that ILD and ITD are processed by separate cortical networks, the ILD and ITD may activate similar brain regions, whose separation is not detectable by EEG. Whether and where ITD and ILD cues integrate in human cortex is still under debate. And one could interpret the relatively inconsistent electrophysiological evidence seen to date as meaning that although cortical responses to perceptually matched ITD and ILD stimuli show some differences, a cortical region sensitive to sound location independent of the type of acoustic cue possibly exists. See (Ungan et al., 2001; Johnson and Hautus, 2010; Edmonds and Krumbholz, 2014; Salminen et al., 2015; Altmann et al., 2017; Higgins et al., 2017).

Limitations. In this study, we showed that one could decode EEG signals to investigate the neural underpinnings of auditory motion processing. Nevertheless, there are several limitations in the current study.

First, we discuss some limitations arising from the stimulus used. The stimuli were presented over headphones, and the motion was simulated by manipulating the acoustic cues that were embedded in the signal. In the first experiment, VAS was used which is based on non-individualized HRTFs. In the second experiment, subjects were presented with artificially impoverished sound that contained only ILD or ITD binaural cues. Therefore, in both experiments, the fidelity of sound spatialization was imperfect and the neural responses we obtained might not fully represent processing of 'true' auditory motion as it would be when using free-field sound presentation.

Second, the results of the cross-modal decoding in the second experiment should be interpreted with caution since the ITD and ILD were not perceptually matched. We used only approximate linear mapping between the acoustic cues where the maximum ITDs of ± 750 μs corresponded to ILD ± 20 dB. The absence of evidence for a change in performance when exchanging decoders cannot necessarily be construed as evidence for absence of any difference in how ITD and ILD are represented in cortex. Especially given the inherent spatial limitations and signal-to-noise ratio of EEG.

Finally, the sound trajectory time-series used here could also be improved to better elucidate the underlying neural dynamics of tracking sound motion. In particular, the stimuli trajectories happened to be relatively strongly autocorrelated (see appendix D). This caused 'smearing' of the temporal dimension of our decoding and prevented us from providing more detailed information on the neural dynamics of sound motion processing.

5. Conclusion

We demonstrated that cortical activity fluctuates in a phase-locked manner with the dynamics of a moving sound source. Specifically, the delta phase and alpha power of EEG were both found to contribute to the decoding of sound azimuth. Moreover, using spatially impoverished stimuli, we found that this cortical tracking is equally present for both binaural acoustic cues ILD and ITD.

Declarations of interest

None.

Acknowledgments

Irish Research Council for Science, Engineering and Technology (GOIPG/2015/1656).

Appendix E. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2018.07.054.

References

Ahveninen, J., Huang, S., Belliveau, J.W., Chang, W.-T., Hämäläinen, M., 2013. Dynamic oscillatory processes governing cued orienting and allocation of auditory attention. J. Cognit. Neurosci. 25, 1926–1943.

Altman, J., Vaitulevich, S.P., Shestopalova, L., Petropavlovskaia, E., 2010. How does mismatch negativity reflect auditory motion? Hear. Res. 268, 194–201.

Altmann, C.F., Bledowski, C., Wibral, M., Kaiser, J., 2007. Processing of location and pattern changes of natural sounds in the human auditory cortex. Neuroimage 35, 1192–1200.

Altmann, C.F., Ueda, R., Bucher, B., Furukawa, S., Ono, K., Kashino, M., Mima, T., Fukuyama, H., 2017. Trading of dynamic interaural time and level difference cues and its effect on the auditory motion-onset response measured with electroencephalography. Neuroimage 159, 185–194.

Baumgart, F., Gaschler-Markefski, B., Woldorff, M.G., Heinze, H.-J., Scheich, H., 1999. A movement-sensitive area in auditory cortex. Nature 400, 724–726.

Bednar, A., Boland, F.M., Lalor, E.C., 2017. Different spatio-temporal electroencephalography features drive the successful decoding of binaural and monaural cues for sound localization. Eur. J. Neurosci. 45, 679–689.

Burke, K.A., Letsos, A., Butler, R.A., 1994. Asymmetric performances in binaural localization of sound in space. Neuropsychologia 32, 1409–1417.

Combrisson, E., Jerbi, K., 2015. Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. J. Neurosci. Meth. 250, 126–136.

Crosse, M.J., Liberto, D.,M.,G., Bednar, A., Lalor, E.C., 2016. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Front. Hum. Neurosci. 10.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Meth. 134, 9–21.

Derey, K., Valente, G., de Gelder, B., Formisano, E., 2016. Opponent coding of sound location (azimuth) in planum temporale is robust to sound-level variations. Cerebr. Cortex 26, 450–464.

Ding, N., Simon, J.Z., 2012. Emergence of neural encoding of auditory objects while listening to competing speakers. Proc. Natl. Acad. Sci. Unit. States Am. 109, 11854–11859.

Ducommun, C.Y., Murray, M.M., Thut, G., Bellmann, A., Viaud-Delmon, I., Clarke, S., Michel, C.M., 2002. Segregated processing of auditory motion and auditory location: an ERP mapping study. Neuroimage 16, 76–88.
Edmonds, B.A., Krumbholz, K., 2014. Are interaural time and level differences represented by independent or integrated codes in the human auditory cortex? J Assoc Res Otolaryngol 15, 103–114.
Feng, W., Störmer, V.S., Martinez, A., McDonald, J.J., Hillyard, S.A., 2017. Involuntary orienting of attention to a sound desynchronizes the occipital alpha rhythm and improves visual perception. Neuroimage 150, 318–328.
Fujiki, N., Riederer, K.A.J., Jousmaki, V., Makela, J.P., Hari, R., 2002. Human cortical representation of virtual auditory space: differences between sound azimuth and elevation. Eur. J. Neurosci. 16, 2207–2213.
Getzmann, S., 2011. Auditory motion perception: onset position and motion direction are encoded in discrete processing stages. Eur. J. Neurosci. 33, 1339–1350.
Getzmann, S., Lewald, J., 2010. Shared cortical systems for processing of horizontal and vertical sound motion. J. Neurophysiol. 103, 1896–1904.
Getzmann, S., Lewald, J., 2012. Cortical processing of change in sound location: smooth motion versus discontinuous displacement. Brain Res. 1466, 119–127.
Godey, B., Schwartz, D., de Graaf, J.B., Chauvel, P., Liegeois-Chauvel, C., 2001. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. Clin. Neurophysiol. 112, 1850–1859.
Grothe, B., Pecka, M., McAlpine, D., 2010. Mechanisms of sound localization in mammals. Physiol. Rev. 90, 983–1012.
Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., Bießmann, F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage 87, 96–110.
Higgins, N.C., McLaughlin, S.A., Rinne, T., Stecker, G.C., 2017. Evidence for cue-independent spatial representation in the human auditory cortex during active listening. Proc. Natl. Acad. Sci. Unit. States Am. 201707522.
Johnson, B.W., Hautus, M.J., 2010. Processing of binaural spatial information in human auditory cortex: neuromagnetic responses to interaural timing and level differences. Neuropsychologia 48, 2610–2619.
Kaiser, J., Lutzenberger, W., Preissl, H., Ackermann, H., Birbaumer, N., 2000. Right-hemisphere dominance for the processing of sound-source lateralization. J. Neurosci. 20, 6631–6639.
Kerlin, J.R., Shahin, A.J., Miller, L.M., 2010. Attentional gain control of ongoing cortical speech representations in a “cocktail party”. J. Neurosci. 30, 620–628.
Krumbholz, K., Hewson-Stoate, N., Schönwiesner, M., 2007. Cortical response to auditory motion suggests an asymmetry in the reliance on inter-hemispheric connections between the left and right auditory cortices. J. Neurophysiol. 97, 1649–1655.
Krumbholz, K., Schönwiesner, M., Rübsamen, R., Zilles, K., Fink, G.R., von Cramon, D.Y., 2005a. Hierarchical processing of sound location and motion in the human brainstem and planum temporale. Eur. J. Neurosci. 21, 230–238.
Krumbholz, K., Schonwiesner, M., von Cramon, D.Y., Rubsamen, R., Shah, N.J., Zilles, K., Fink, G.R., 2005b. Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe. Cerebr. Cortex 15, 317–324.
Lalor, E.C., Foxe, J.J., 2010. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. Eur. J. Neurosci. 31, 189–193.
Lalor, E.C., Power, A.J., Reilly, R.B., Foxe, J.J., 2009. Resolving precise temporal processing properties of the auditory system using continuous stimuli. J. Neurophysiol. 102, 349–359.
Lewald, J., Getzmann, S., 2011. When and where of auditory spatial processing in cortex: a novel approach using electrotomography. PLoS One 6, e25146.
Lewis, J.W., Beauchamp, M.S., DeYoe, E.A., 2000. A comparison of visual and auditory motion processing in human cerebral cortex. Cerebr. Cortex 10, 873–888.
Luo, H., Liu, Z., Poeppel, D., 2010. Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. PLoS Biol. 8, e1000445.
McAlpine, D., Jiang, D., Shackleton, T.M., Palmer, A.R., 2000. Responses of neurons in the inferior colliculus to dynamic interaural phase cues: evidence for a mechanism of binaural adaptation. J. Neurophysiol. 83, 1356–1365.
McLaughlin, S.A., Higgins, N.C., Stecker, G.C., 2015. Tuning to binaural cues in human auditory cortex. J Assoc Res Otolaryngol 1–17.
Mesgarani, N., David, S.V., Fritz, J.B., Shamma, S.A., 2009. Influence of context and behavior on stimulus reconstruction from neural activity in primary auditory cortex. J. Neurophysiol. 102, 3329–3339.
Middlebrooks, J.C., Pettigrew, J.D., 1981. Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. J. Neurosci. 1, 107–120.
O'Sullivan, J.A., Power, A.J., Mesgarani, N., Rajaram, S., Foxe, J.J., Shinn-Cunningham, B.G., Slaney, M., Shamma, S.A., Lalor, E.C., 2015. Attentional selection in a cocktail party environment can Be decoded from single-trial EEG. Cerebr. Cortex 25, 1697–1706.
Ortiz-Rios, M., Azevedo, F.A., Kusmierek, P., Balla, D.Z., Munk, M.H., Keliris, G.A., Logothetis, N.K., Rauschecker, J.P., 2017. Widespread and opponent fMRI signals represent sound location in macaque auditory cortex. Neuron 93, 971–983 e974.
Palomaki, K., Alku, P., Makinen, V., May, P., Tiitinen, H., 2000. Sound localization in the human brain: neuromagnetic observations. Neuroreport 11, 1535–1538.
Palomaki, K.J., Tiitinen, H., Makinen, V., May, P.J., Alku, P., 2005. Spatial processing in human auditory cortex: the effects of 3D, ITD, and ILD stimulation techniques. Brain Res Cogn Brain Res 24, 364–379.
Pavani, F., Macaluso, E., Warren, J.D., Driver, J., Griffiths, T.D., 2002. A common cortical substrate activated by horizontal and vertical sound movement in the human brain. Curr. Biol. 12, 1584–1590.
Poirier, C., Baumann, S., Dheerendra, P., Joly, O., Hunter, D., Balezeau, F., Sun, L., Rees, A., Petkov, C.I., Thiele, A., Griffiths, T.D., 2017. Auditory motion-specific mechanisms in the primate brain. PLoS Biol. 15, e2001379.
Salminen, N.H., Takanen, M., Santala, O., Lamminsalo, J., Altoe, A., Pulkki, V., 2015. Integrated processing of spatial cues in human auditory cortex. Hear. Res. 327, 143–152.
Shestopalova, L.B., Petropavlovskaia, E.A., Vaitulevich, S.P., Vasilenko, Y.A., Nikitin, N.I., Altman, J.A., 2012. Discrimination of auditory motion patterns: the mismatch negativity study. Neuropsychologia 50, 2720–2729.
Smith, K.R., Saberi, K., Hickok, G., 2007. An event-related fMRI study of auditory motion perception: No evidence for a specialized cortical system. Brain Res. 1150, 94–99.
Spitzer, M., Semple, M., 1993. Responses of inferior colliculus neurons to time-varying interaural phase disparity: effects of shifting the locus of virtual motion. J. Neurophysiol. 69, 1245–1263.
Stecker, G.C., Harrington, I.A., Middlebrooks, J.C., 2005. Location coding by opponent neural populations in the auditory cortex. PLoS Biol. 3, e78.
Stecker, G.C., Mickey, B.J., Macpherson, E.A., Middlebrooks, J.C., 2003. Spatial sensitivity in field PAF of cat auditory cortex. J. Neurophysiol. 89, 2889–2903.
Störmer, V.S., Feng, W., Martinez, A., McDonald, J.J., Hillyard, S.A., 2016. Salient, irrelevant sounds reflexively induce alpha rhythm desynchronization in parallel with slow potential shifts in visual cortex. J. Cognit. Neurosci.
Ungan, P., Yagcioglu, S., Goksoy, C., 2001. Differences between the N1 waves of the responses to interaural time and intensity disparities: scalp topography and dipole sources. Clin. Neurophysiol. 112, 485–498.
Warren, J.D., Zielinski, B.A., Green, G.G.R., Rauschecker, J.P., Griffiths, T.D., 2002. Perception of sound-source motion by the human brain. Neuron 34, 139–148.
Werner-Reiss, U., Groh, J.M., 2008. A rate code for sound azimuth in monkey auditory cortex: implications for human neuroimaging studies. J. Neurosci. 28, 3747–3758.
Wöstmann, M., Herrmann, B., Maess, B., Obleser, J., 2016. Spatiotemporal dynamics of auditory attention synchronize with speech. Proc. Natl. Acad. Sci. U.S.A. 113, 3873–3878.
Zatorre, R.J., Penhune, V.B., 2001. Spatial localization after excision of human auditory cortex. J. Neurosci. 21, 6321–6328.