

RESEARCH ARTICLE

Across-subjects classification of stimulus modality from human MEG high frequency activity

Britta U. Westner^{1,2*}, Sarang S. Dalal², Simon Hanslmayr³, Tobias Staudigl⁴

1 Department of Psychology, University of Konstanz, Konstanz, Germany, **2** Center of Functionally Integrative Neuroscience, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark, **3** School of Psychology, University of Birmingham, Birmingham, United Kingdom, **4** Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

* britta.wstnr@gmail.com



OPEN ACCESS

Citation: Westner BU, Dalal SS, Hanslmayr S, Staudigl T (2018) Across-subjects classification of stimulus modality from human MEG high frequency activity. *PLoS Comput Biol* 14(3): e1005938. <https://doi.org/10.1371/journal.pcbi.1005938>

Editor: Daniel Bush, University College London, UNITED KINGDOM

Received: June 27, 2017

Accepted: December 20, 2017

Published: March 12, 2018

Copyright: © 2018 Westner et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data is hosted on Open Science Framework, DOI [10.17605/OSF.IO/M25N4](https://doi.org/10.17605/OSF.IO/M25N4): <https://osf.io/M25N4/>.

Funding: This work was supported by ERA-Net NEURON via the German Federal Ministry of Education and Research, <http://www.neuron-eranet.eu/>, grant 01EW1307, to SSD; European Research Council, <https://erc.europa.eu/>, Starting Grant 640488, to SSD; Deutsche Forschungsgemeinschaft, <http://www.dfg.de/>, Emmy Noether Programme Grant HA 5622/1-1, to

Abstract

Single-trial analyses have the potential to uncover meaningful brain dynamics that are obscured when averaging across trials. However, low signal-to-noise ratio (SNR) can impede the use of single-trial analyses and decoding methods. In this study, we investigate the applicability of a single-trial approach to decode stimulus modality from magnetoencephalographic (MEG) high frequency activity. In order to classify the auditory versus visual presentation of words, we combine beamformer source reconstruction with the random forest classification method. To enable group level inference, the classification is embedded in an across-subjects framework. We show that single-trial gamma SNR allows for good classification performance (accuracy across subjects: 66.44%). This implies that the characteristics of high frequency activity have a high consistency across trials and subjects. The random forest classifier assigned informational value to activity in both auditory and visual cortex with high spatial specificity. Across time, gamma power was most informative during stimulus presentation. Among all frequency bands, the 75 Hz 95 Hz band was the most informative frequency band in visual as well as in auditory areas. Especially in visual areas, a broad range of gamma frequencies (55 Hz 125 Hz) contributed to the successful classification. Thus, we demonstrate the feasibility of single-trial approaches for decoding the stimulus modality across subjects from high frequency activity and describe the discriminative gamma activity in time, frequency, and space.

Author summary

Averaging brain activity across trials is a powerful way to increase signal-to-noise ratio in MEG data. This approach, however, potentially obscures meaningful brain dynamics that unfold on the single-trial level. Single-trial analyses have been successfully applied to time domain or low frequency oscillatory activity; its application to MEG high frequency activity is hindered by the low amplitude of these signals. In the present study, we show that stimulus modality (visual versus auditory presentation of words) can successfully be

SH; European Research Council, <https://erc.europa.eu/>, Consolidator Grant 647954, to SH; Wolfson Foundation and Royal Society, <https://royalsociety.org/grants-schemes-awards/grants/wolfson-research-merit/>, to SH; and European Union's Horizon 2020, <https://ec.europa.eu/programmes/horizon2020/>, 661373, to TS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

decoded from single-trial MEG high frequency activity by combining source reconstruction with a random forest classification algorithm. This approach reveals patterns of activity above 75 Hz in both visual and auditory cortex, highlighting the importance of high frequency activity for the processing of domain-specific stimuli. Thereby, our results extend prior findings by revealing high-frequency activity in auditory cortex related to auditory word stimuli in MEG data. The adopted across-subjects framework furthermore suggests a high inter-individual consistency in the high frequency activity patterns.

Introduction

Since the first reports of cortical gamma band activity [1, 2], these high frequency responses have been linked to a plethora of brain processes and mental tasks, for example visual perception and processing [3–6], auditory perception [7, 8] or memory [9–12]. Although numerous theories about the origin and function of these high frequency oscillations and their relation with lower frequencies like theta and alpha have been proposed (e.g., [13–15]), there is an ongoing debate about whether gamma band responses reflect narrowband oscillations or broadband power increases, possibly echoing an increase in spiking activity [6, 16, 17]. One obstacle in this quest is the $1/f$ characteristic of the brain's frequency power spectrum and a low signal-to-noise ratio (SNR) gamma band activity in magnetoencephalography (MEG) or electroencephalography (EEG) recordings. To increase SNR, trial averaging is a frequently used tool to cancel out random variance. However, this approach can potentially obscure or cancel meaningful brain activity [18]. Indeed, local field potentials and electrocorticographic data from monkeys revealed systematic trial-to-trial variations in gamma power and frequency in a visual [19] and a memory task [20]. Importantly, the averages across trials in these studies displayed the classic sustained gamma effect, indicating that single-trial responses are crucial to understand the brain's dynamics [18]. One powerful approach to assess single-trial information are multivariate decoding techniques. Whether such methods are applicable to low SNR gamma band MEG data, however, remains unclear. In the present paper, we investigate the predictive value of single-trial gamma power regarding the modality of stimulus presentation (auditory or visual presentation of words) in human MEG data. While comparable contrasts have been used to test classifier performance or as example datasets (e.g., [21, 22]), our aim was to unravel single-trial high frequency patterns in human MEG data. To decode information about stimulus-modality from the time-frequency data, we used a combination of beamforming [23] and random forest classification [24]. This approach was embedded into an across-subjects cross-validation framework, where the classifier was tested on single trials of unseen subjects to assess the generality of the spatial time-frequency pattern. Our results confirm that gamma SNR in single trials is high enough to achieve stable classification accuracy significantly above chance. Interestingly, the classification model yields high informational value to a broad bandwidth in the gamma range. Furthermore, we show that the characteristics of the gamma activity are similar enough across trials and even subjects to yield reliable classification performance.

Materials and methods

Ethics statement

The study was approved by the Institutional Review Board of the University of Konstanz and in accordance with the Declaration of Helsinki.

Participants

A total of 24 participants (17 female; mean age = 22 years, range = 19–26 years; 21 right-handed) took part in this MEG experiment. Three participants were excluded due to technical problems, one due to excessive environmental noise. The data from the remaining 20 participants are presented here. All of the participants gave written informed consent prior to the experiment and received course credits or nominal financial compensation for participation. All participants were German native speakers and reported normal or corrected-to-normal vision, and no history of neurological disease.

Parts of this data have been published in [12], with respect to independent research questions and analyses.

Design, procedure, and material

The experiment consisted of a study phase and a subsequent recognition test. Only data from the study phase are reported here. In the study phase, participants were presented with words either visually (projected centrally on a screen) or auditorily (via nonferromagnetic tubes to both ears). The duration of the visual word presentation was determined by the duration of the respective audio file, i.e., the time to pronounce the word (mean duration = 697 ms, *s.d.* = 119 ms). Each word was followed by a fixation cross. The duration of the word and fixation cross together added up to 2000 ms. Participants were instructed to count the syllables of the word and indicate via button press whether the word had two syllables. A question mark (max. duration of 1500 ms) prompted the subject's response. The button press ended the presentation of the question mark. A fixation cross with variable duration (1000 ms to 1500 ms) was presented before each item. After the encoding phase, participants performed a distractor task and a surprise recognition test phase.

The stimuli consisted of 420 unrelated German nouns, grouped into three lists with 140 words. Half of each list's words had two syllables, the other half had one, three or four syllables. Two lists were presented during the study phase and one list during the test phase. The assignment of the lists to study or test phase was counterbalanced across participants. Items were presented in random order, with the constraint that not more than 5 words of the same modality and not more than 5 words from the same condition were presented sequentially.

MEG data acquisition and preprocessing

MEG data was recorded with a 148-channel magnetometer (MAGNES 2500 WH, 4D Neuro-imaging, San Diego, USA) in a supine position inside a magnetically shielded room. Data was continuously recorded at a sampling rate of 678.17 Hz and bandwidth of 0.1 Hz to 200 Hz, and later downsampled to 300 Hz to reduce computational load. All data processing prior to classification was done using FieldTrip [25], an open-source MATLAB toolbox for MEEG data analysis. Data was epoched into single trials, with epochs ranging from 1500 ms before item presentation to 4000 ms after item presentation. Trials were visually inspected for artifacts, contaminated trials were rejected. Thereafter, trials were corrected for blinks, eye movements, and cardiac artifacts using independent component analysis (ICA).

Source reconstruction

For coregistration with the individual structural magnetic resonance image (available for 17 out of 20 participants; for the remaining three participants we used an affine transformation of an MNI-template brain; Montreal Neurological Institute, Montreal, Canada), the shape of the participant's head as well as three markers (nasion, left and right ear canal) and the location of

the head position indicator (HPI) coils were digitized prior to the experiment using a Fastrak Polhemus 3D scanner (Polhemus, Colchester, VT, USA).

Single-trial source space activity was reconstructed using a linearly constrained minimum variance (LCMV) beamformer [23] with weight normalization (neural activity index; [23, 26]). First, the spatial filter was computed adopting a realistic single shell head model [27] based on the individual structural magnetic resonance image (MRI) and a source model with grid points covering the whole brain volume (resolution: 15 mm). The data covariance matrix was computed for -500 ms to 1000 ms relative to stimulus presentation. To account for the rank-deficiency of the data that was introduced by the application of the ICA, the covariance matrix was regularized by loading its diagonal with 5% of the sensor power. Subsequently, the spatial filter was applied to the single trials to obtain virtual electrodes at all grid point locations.

For the classification of the oscillatory activity, single-trial time frequency representations were calculated at every virtual electrode applying a Fast Fourier Transform. Gamma band activity was estimated using frequency smoothing (Slepian sequence multi taper approach), yielding 20 Hz-wide frequency bands centered at 35 Hz, 65 Hz, 85 Hz, 115 Hz and 135 Hz. The power was calculated separately for 250 ms long time windows from -500 ms to 1000 ms and the post-stimulus activity was then expressed as relative change to baseline power, as using relative change helps to overcome issues arising from the $1/f$ shape of MEG data.

Random forest classification

The random forest algorithm [24], an ensemble method, aggregates the results of several classifiers. These so-called base learners are classification and regression trees [28], which partition the data by adopting binary splits. The aim of this partitioning process is to reduce the impurity regarding the class labels in the daughter nodes that result from this split: preferably, all observations from one class should arrive in the same node. In every split, the tree algorithm searches first for the predictor that maximizes the purity of the daughter nodes and then for the best split point within that predictor. Random forests now grow numerous trees; each of these trees, however, is built on a bootstrap sample of the original data and in every split only a random subsample of all predictors is searched. The variance introduced by this randomness leads to a robust prediction by the aggregated model. This approach furthermore enables random forest to cope particularly well with highly correlated predictor variables [29], which is of special interest when working with MEEG data. Additionally, data with more predictors than observations (small n large p problems) are also handled effectively since the predictor variables are searched successively [30], which makes this approach particularly interesting when dealing with high-dimensional source-space MEEG data. For every predictor, the algorithm returns an estimate of how important this variable was for the model's prediction. The version used here is based on the impurity reduction introduced by a predictor variable across all trees, which is measured by the Gini index [28, 29, 31].

Random forest classification was performed using the scikit-learn module for Python [32]. The aim of the decoding was to classify trials regarding their stimulus modality: visual or auditory. The predictors were [voxel, time point, frequency band]-triplets, providing 16 624 predictors, overall. For every subject, the more prevalent class (auditory or visual stimulation) was downsampled such that every dataset contained equal trial numbers for both cases. The total trial number across all subjects was 4270 trials.

The classification was embedded in a cross-validation framework across subjects: the classifier was trained on the data from all but one subject and then tested on the data of this left-out subject. This procedure was repeated for all 20 subjects, such that every dataset was used as test set once. This approach ensures that the classifier is never tested on data it was trained on and

thus controls for possible overfitting of the classifier. Moreover, it allows the assessment of across-subjects predictability of the data regarding the response variable.

Each of the 20 cross-validation models aggregated the results of 15 000 classification trees, where every tree was built on a bootstrap sample of all observations in the trainings set. To ensure that the model incorporated a sufficient number of trees, classification performance was assessed with 25 000 trees for two folds [33], yielding comparable results as the sparser model. At each binary split, the algorithm considered $\sqrt{N_{\text{features}}}$ predictors to find the best split. The accuracies on the test datasets as well as the variable importances were merged across the cross-validation folds. The performance of the classifier was then tested against 50% chance level using a binomial test [34], since a permutation based test was computationally not feasible.

Comparison to support vector machines

The performance of the random forest algorithm was compared to the accuracies obtained by fitting two support vector machine (SVM) models [35, 36] on the data. The general procedure was as described above, instead of the random forest model, however, either a linear SVM or a non-linear SVM with a radial basis function (RBF) kernel was fitted using the scikit-learn module in Python. The penalty parameter was set to 1.0 in both models and a kernel coefficient of $1/N_{\text{features}}$ for the RBF kernel.

The performance of the random forest model was subsequently tested against the two SVM models by applying a Fisher's exact test to the obtained classification accuracies.

Results

To assess the predictive value of single-trial gamma power towards stimulus modality, we used MEG data from 20 subjects and adopted an across-subjects classification scheme. Data was first source reconstructed with a linearly constrained minimum variance (LCMV) beamformer, subsequently, we used the random forest algorithm to classify the modality of stimulus presentation (auditory or visual).

The random forest model classified auditory versus visual trials with 66.44% accuracy, which is significantly better than chance (binomial test, $n_{\text{trials}} = 4270$, $p < 0.001$). As the confusion matrix in Fig 1A shows, the accuracy was slightly better for auditory trials (69.60%) than for visual trials (63.19%). In the adopted 20-fold cross-validation scheme, every fold corresponded to the data of one subject, hence, the classifier was always tested on data of one subject which was not included in building the model. The classifier accuracy on the 20 cross-validation folds is depicted in Fig 1B. The performance on the different folds is diverse, ranging from 50.98% to 84.86%, however, the accuracy for all but three folds is above 60% (note that the folds, since they are part of the whole classifier model, are not tested for significance). The good classifier performance indicates that the gamma power patterns are remarkably stable across trials and even subjects. For a comparison of this across-subjects approach to within-subject analyses, see supplementary Figure in the S1 Appendix.

The random forest classifier provides the variable importance as an importance estimate for every predictor in the model. This measure indicates the informational value of a given predictor towards the discrimination of the two classes, auditory and visual modality. Fig 2 reports the highest 2% of variable importance values, i.e., those [voxel, time point, frequency band]-triplets that were most informative for partitioning the data. This cutoff was chosen because cutoffs of higher values ($> 2\%$) would have included features with variable importances equal to zero, i.e., variables that did not contribute information to the classifier. Not only visual, but also auditory regions contributed to the model, even though visual areas yielded more

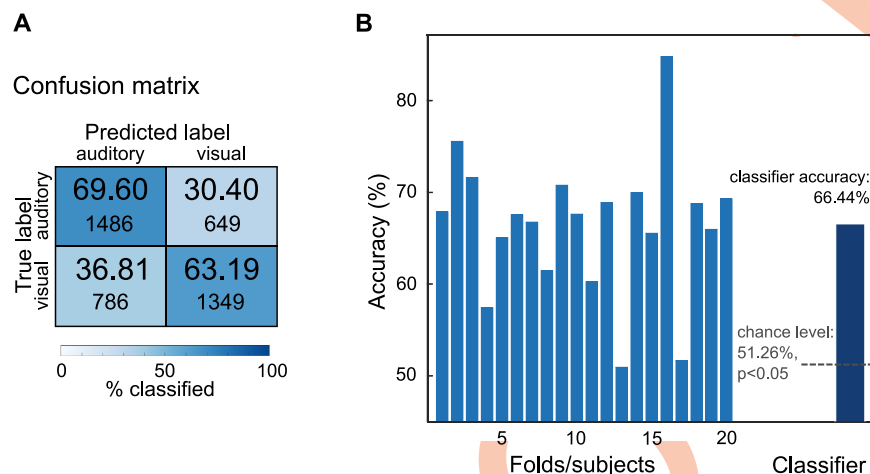


Fig 1. Classifier results. **A** Confusion matrix showing the proportion of correctly classified trials (diagonal) and misclassified trials. **B** Classifier accuracy. The mid blue bars represent the test accuracy in the 20 folds, where every fold is a subject the classifier was not trained on. The dark blue bar shows the overall classifier accuracy, which was tested against chance level.

<https://doi.org/10.1371/journal.pcbi.1005938.g001>

information than the auditory cortex. Interestingly, the lower frequency bands of 25 Hz to 45 Hz and 55 Hz to 75 Hz did not rank as important as the 75 Hz to 95 Hz band. Even frequencies above 100 Hz contributed to the model in both visual and right auditory cortex. Gamma power beyond 125 Hz, however, did not add substantially to the classification model.

All time windows but the last one (750 ms to 1000 ms) supplied information to the classifier, in higher frequencies, the earlier time windows seemed to play a more pronounced role compared to the lower gamma frequencies. Fig 3 shows the time-frequency representations of

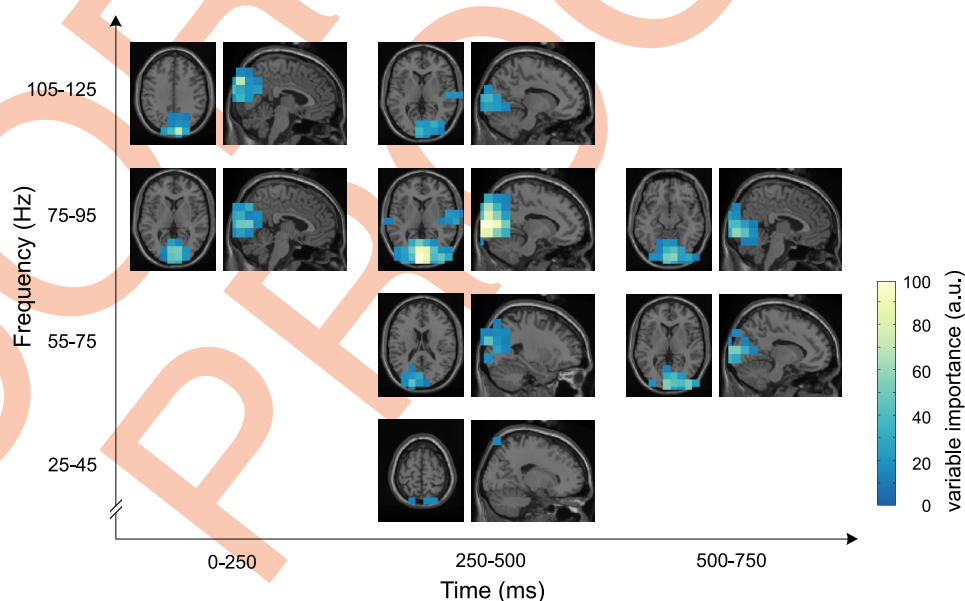


Fig 2. Variable importances. The 2% most important predictors are shown across time, frequency and space. A higher variable importance score implies that this predictor had a higher informative value in the random forest model to partition the data into trials with auditory and visual perception. The orthogonal views are centered on the voxel showing the highest variable importance.

<https://doi.org/10.1371/journal.pcbi.1005938.g002>

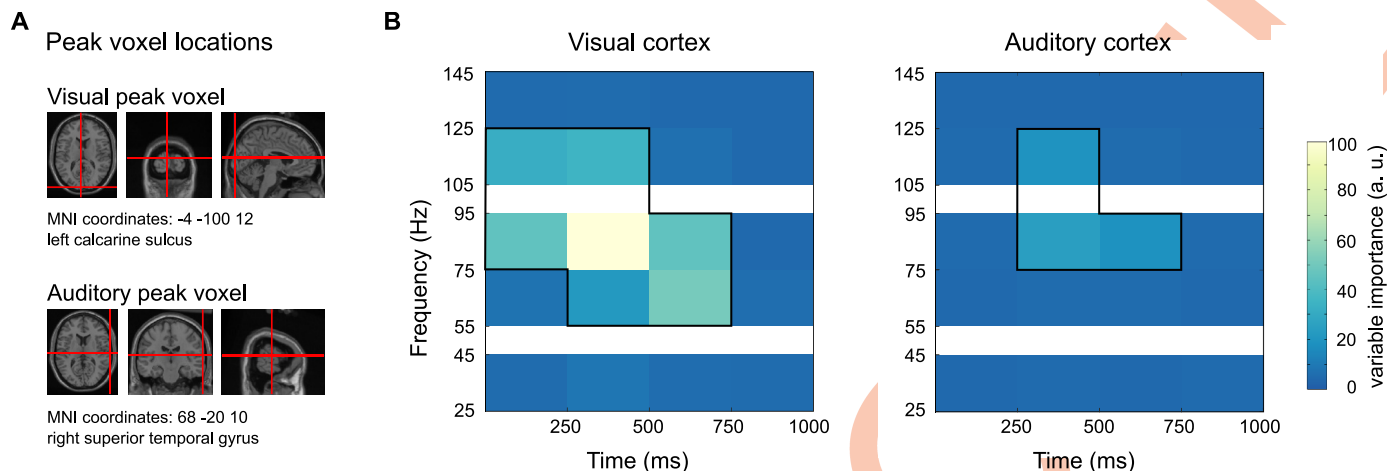


Fig 3. Variable importances in visual and auditory peak voxels. A Peak voxel locations for auditory and visual cortex (compare peak voxels from Fig 2) B Time-frequency representation of variable importances in those peak voxels. Black boxes indicate those variables which were among the 2% most informative predictors.

<https://doi.org/10.1371/journal.pcbi.1005938.g003>

variable importance for the visual and auditory peak voxels: the visual peak voxel (MNI coordinates: [-4 -100 12]) falls into left calcarine sulcus, the auditory peak voxel (MNI coordinates: [68 -20 10]) into right superior temporal gyrus (labels determined with the Automated Anatomical Labeling (AAL) atlas [37]). The time-frequency representations for those two peak voxels (Fig 3) confirm the pattern evident across all voxels (Fig 2). Thus, the 75 Hz to 95 Hz band yielded a characteristic and stable activity pattern in both the auditory and visual cortex. The visual response was specifically characterized by a broadband gamma increase in the range of 55 Hz to 125 Hz. The auditory response yielded informational value in an overlapping but narrower frequency range (75 Hz to 125 Hz).

To investigate the underlying gamma power changes, the variable importance rankings were compared to the power differences between auditory and visual trials. To this end, auditory and visual power changes relative to baseline were averaged across trials and subjects, and the difference between the visual and the auditory condition was computed. These differences are depicted in Fig 4A: the spatial pattern of power is shown for the 250 ms to 500 ms time window and two frequency bands (75 Hz to 95 Hz, top, and 105 Hz to 125 Hz, bottom in Fig 4A). Red colors refer to higher gamma power in the visual condition and blue colors to higher power in the auditory condition. The black lines encircle those voxels which were among the 2% most informative predictors for the classifier. Fig 4B shows the underlying gamma power relations for the same peak voxels as presented in Fig 3. Interestingly, the classifier analysis based on single trials also rated predictors as highly informative where a difference in the averages is small, as is most evident for the time-frequency representation of the auditory condition (75 Hz to 95 Hz, 500 ms to 750 ms).

To get a general assessment of the relative performance of random forests on MEG data, we compared this approach to the outcome of support vector machines (SVM), a widely used method in multivariate data analysis. We applied two SVM types, a linear SVM and a non-linear SVM with a radial basis function (RBF) kernel. The results of this comparison are shown in Fig 5. The linear SVM yielded an accuracy of 63.07% and performed significantly worse than the random forest model (two-sided Fisher's exact test, *odds ratio* = 86.79%, *p* = 0.002, 0.004 corrected). The RBF SVM performed slightly better than the random forest model with an accuracy of 68.34% (two-sided Fisher's exact test, *odds ratio* = 91.1%, *p* = 0.047, 0.094 corrected). Note that the comparison between the non-linear SVM and the random forest model

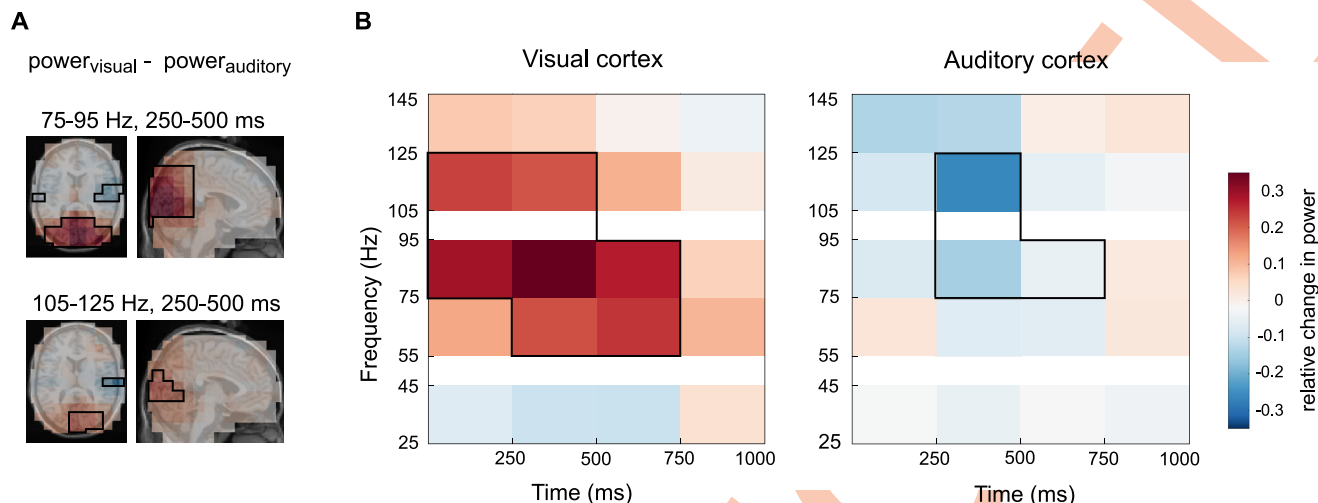


Fig 4. Underlying gamma power. This figure shows the difference in averaged gamma power between visual and auditory word presentation trials. **A** Spatial representation of gamma power for two frequency bands (75-95 Hz, top, and 105-125 Hz, bottom). Red hues represent a higher gamma power in the average of visual trials, the blue colors depict higher gamma power in the average of the auditory condition. Black boxes indicate the 2% most informative predictors as shown in Fig 2. **B** Gamma power in visual and auditory peak voxels. Shown is the difference between the visual and auditory condition, black boxes again indicate the most informative predictors for the classifier model.

<https://doi.org/10.1371/journal.pcbi.1005938.g004>

does not survive a correction for multiple comparisons (Bonferroni-correction for two Fisher's exact tests).

Discussion

In the present work, we investigated the predictive value of single-trial gamma power to classify the stimuli's modality. This was done in an across-subjects cross-validation framework which allowed us to estimate not only the gamma pattern stability across trials but also across subjects.

The decoding of MEEG high frequency activity on a single-trial basis can be challenging due to the low SNR: while intracranially recorded high frequency activity up to 180 Hz has been used to decode movements [38, 39], comparable approaches with MEEG data were not successful [40, 41]. Some studies could show a contribution of high gamma power (along with lower oscillatory activity) to the overall classifier performance [42, 43]. In this study, we successfully decoded stimulus modality exclusively from high frequency activity: the classifier model was able to correctly classify 66.44% of the trials based on their source reconstructed gamma activity pattern, reliably distinguishing visual from auditory word presentation. Thus, the SNR of single-trial gamma power in source-level MEG data was high enough to successfully apply single-trial multivariate analyses. Interestingly, more auditory (69.60%) than visual trials (63.19%) were classified correctly, although visual areas yielded more information to the classifier. One possible explanation for this could be that the classifier-inherent cutoff values for gamma power in the visual voxels were rather conservative and therefore missed small gamma increases in visual cortex in visual trials, but still reliably detected the absence of visual activity in auditory trials.

The classification model was built across subjects, adopting a 20-fold across-subjects cross-validation, where the classifier was trained on 19 subjects and then tested on the data of the left-out 20th subject. Hence, the trials of any given subject were classified by a model which was built on the data from different subjects. Using this approach, we assessed the common

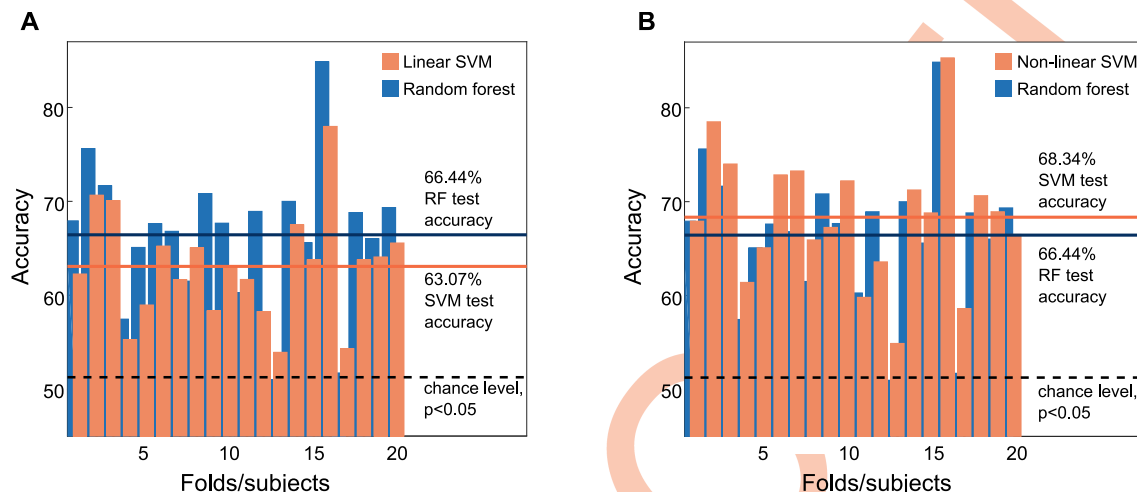


Fig 5. Comparison to SVM models. Performance comparison of the random forest algorithm to SVM models. **A** Results from the linear SVM compared to the results from the random forest algorithm across cross-validation folds. Random forest performed significantly better than the linear SVM ($p = 0.002$, 0.004 with Bonferroni correction). **B** Results from a non-linear SVM with RBF kernel compared to the random forest results. The non-linear SVM yielded a higher classification accuracy than the random forest model ($p = 0.047$), applying a Bonferroni-correction for multiple comparisons renders this effect insignificant ($p_{corrected} = 0.094$).

<https://doi.org/10.1371/journal.pcbi.1005938.g005>

patterns across trials and subjects. The accuracy pattern across the different folds was higher than 60% for all but three subjects. Low accuracies indicate either higher noise levels in these participants or activity patterns which deviate from the across-subjects consensus as uncovered by the random forest model. The overall classification accuracy of 66.44% is comparable to previous reports of across-subjects MEG data classification (e.g., [44]).

The variable importance indicates which predictors were used by the model to yield the classification performance, by providing the common pattern across trials and subjects that differentiated between the two conditions. Clearly, gamma band activity from both visual and auditory areas was exploited by the model, although the visual cortex was more important than the auditory cortices, expressed by higher ranking variable importances. Overall, a broad range of frequencies and a time span of 750 ms included gamma band activation relevant to the random forest model.

In this study, we show the feasibility of applying the random forest algorithm [24] to single-trial source-localized time-frequency data. With its non-parametric, non-linear approach and its capability to handle high dimensional datasets with highly correlated predictors, this method is well suited for MEG data (also see [45–48]) and can detect subtle differences concealed in the averaged data. In contrast to approaches comparing averaged time courses, the multivariate analysis presented here furthermore solves the multiple comparison problem that is inherent to the application of univariate tests for time points or similarly high-dimensional data.

Another advantage of this method is the possibility to directly compare predictors (e.g., frequency bands) to each other regarding their importance in the model: for example, we are able to state that the 75 Hz to 95 Hz frequency band is the most important frequency band, and that the visual cortex has higher informational value for the classification than the auditory cortex.

Comparing the random forest algorithm to SVM methods shows that it clearly outperforms a linear SVM model in this dataset. The SVM with the non-linear kernel showed a comparable performance to the random forest, however, using non-linear SVMs comes at the expense of interpretability and usability. While it is possible to assess the relative importance of the different predictors with random forests, this is hindered with non-linear SVMs, since the classification is

not computed in the original feature space. Furthermore, while SVMs often require an extensive search for optimal parameter settings, the application of the random forest algorithm is less complex and needs less fine-tuning.

In our data, the left primary visual cortex was most informative for the classification among all brain regions. Additionally, also higher visual areas ranked as highly informative, which is concordant with the localization of visual gamma band responses in intracranial electroencephalography (iEEG) and MEG studies (e.g., [5, 49–51]). The classification further identified auditory regions as informative. Although iEEG reliably shows high gamma responses to auditory stimuli [52–55], auditory high frequency activity above 75 Hz has only rarely been shown in MEG studies: examples include high gamma responses to sound and pitch perception [56, 57]. Within the auditory regions, the most important voxel in our data was located in the right superior temporal gyrus, which is in line with iEEG studies investigating phoneme and word processing [7, 55] and the above-mentioned MEG studies. Further important regions included Heschl's gyrus and the planum temporale. Interestingly, the right auditory cortex showed higher importance with more voxels involved compared to the left auditory cortex, although the stimuli were words and should typically evoke language-related activity localized to the left hemisphere [54, 55]. This may be explained by the fact that both conditions used words as stimuli and thus, left-hemispheric language related activity is not able to distinguish between auditory and visual trials.

The time windows most important to the classification covered 0 ms to 750 ms after stimulus onset, while the last time window (750 ms to 1000 ms) did not show any high ranking variable importance values, implying that gamma activity was most informative to the classifier during presentation of a word (mean = 700 ms).

Informative predictors in auditory areas, however, were only found between 250 ms and 750 ms, although previous studies reported early auditory (high) gamma responses following phoneme or word stimuli (e.g., [7, 55, 58, 59]).

In both, visual and auditory brain areas, the most important frequency band was the 75 Hz to 95 Hz band. Especially in the 250 ms to 500 ms window, this frequency band exhibited exceeding informative value for the classification. Yet, the visual areas overall provided informative predictors across a broad frequency range (55 Hz to 125 Hz). This points to underlying broadband gamma activity in single trials rather than a narrowband response, which is typically elicited by high contrast stimuli such as gratings, (e.g., [5, 51]). The high frequency activity beneficial for classification is similar to visually induced broadband gamma activity reported in iEEG and MEG studies [6, 49, 60–62]. Vidal et al. [61], for example, describe a lower frequency band of 45 Hz to 65 Hz and high gamma activity of 70 Hz to 120 Hz in their MEG study on visual grouping. Related to reading, broadband high frequency activity above 50 Hz has been reported in iEEG studies [63–66]. Furthermore, compared to the narrowband responses elicited by high contrast stimuli such as gratings, which are typically centered at lower frequencies, (e.g., 50 Hz [5] or 60 Hz [51]), our results yielded the 75 Hz to 95 Hz frequency band as most informative.

In the auditory areas, the most important variables were concentrated in the 75 Hz to 95 Hz and 105 Hz to 125 Hz frequency bands. This is in line with iEEG studies on syllable and word processing, which report gamma responses from 80 Hz up to 200 Hz [54, 67]. Thus, our results might reflect the lower end of the high gamma response described in these studies, potentially cropped by low SNR above 125 Hz.

To summarize, we have shown that single-trial gamma activity can be successfully used to classify stimulus modality. Importantly, the successful across-subjects classification suggested that single-trial gamma-band activity contains high inter-individual consistency. Future studies investigating discriminative rather than consistent activity across trials should explore the

possibilities provided by the present approach. The classifier identified both visual and auditory areas as informative with high spatial specificity. Our results furthermore suggest that single-trial high frequency activity after visual word presentation is characterized by a broadband rather than a narrowband response.

Supporting information

S1 Appendix. Within-subject classification. Classifier accuracies obtained when training and testing the random forest model within instead of across subjects. The procedure adopted a split-half cross-validation scheme including both the fitting of the beamformer and the classification modelling.
(PDF)

Acknowledgments

We thank Ann-Kristin Rombach, Leona Hellwig, Marina Koepfer, and Janine Weichert for help with data acquisition and Sabine Leske and Tzvetan Popov for valuable discussion.

Author Contributions

Conceptualization: Britta U. Westner, Tobias Staudigl.

Formal analysis: Britta U. Westner, Tobias Staudigl.

Funding acquisition: Sarang S. Dalal, Simon Hanslmayr, Tobias Staudigl.

Investigation: Tobias Staudigl.

Methodology: Britta U. Westner.

Resources: Sarang S. Dalal, Simon Hanslmayr, Tobias Staudigl.

Supervision: Sarang S. Dalal, Tobias Staudigl.

Visualization: Britta U. Westner.

Writing – original draft: Britta U. Westner, Sarang S. Dalal, Tobias Staudigl.

Writing – review & editing: Britta U. Westner, Sarang S. Dalal, Simon Hanslmayr, Tobias Staudigl.

References

1. Gray CM, Singer W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences*. 1989; 86:1698–1702. <https://doi.org/10.1073/pnas.86.5.1698>
2. Gray CM, König P, Engel AK, Singer W. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*. 1989; 338:334–337. <https://doi.org/10.1038/338334a0> PMID: 2922061
3. Hoogenboom N, Schoffelen JM, Oostenveld R, Parkes LM, Fries P. Localizing human visual gamma-band activity in frequency, time and space. *NeuroImage*. 2006; 29(3):764–773. <https://doi.org/10.1016/j.neuroimage.2005.08.043> PMID: 16216533
4. Dalal SS, Vidal JR, Hamamé CM, Ossandón T, Bertrand O, Lachaux JP, et al. Spanning the rich spectrum of the human brain: slow waves to gamma and beyond. *Brain Structure and Function*. 2011; 216:77–84. <https://doi.org/10.1007/s00429-011-0307-z> PMID: 21437655
5. Muthukumaraswamy SD, Singh KD, Swettenham JB, Jones DK. Visual gamma oscillations and evoked responses: variability, repeatability and structural MRI correlates. *NeuroImage*. 2010; 49(4):3349–3357. <https://doi.org/10.1016/j.neuroimage.2009.11.045> PMID: 19944770

6. Hermes D, Miller KJ, Wandell BA, Winawer J. Stimulus dependence of gamma oscillations in human visual cortex. *Cerebral Cortex*. 2014; 25(9):2951–2959. <https://doi.org/10.1093/cercor/bhu091> PMID: 24855114
7. Crone NE, Boatman D, Gordon B, Hao L. Induced electrocorticographic gamma activity during auditory perception. *Clinical Neurophysiology*. 2001; 112:565–582. [https://doi.org/10.1016/S1388-2457\(00\)00545-9](https://doi.org/10.1016/S1388-2457(00)00545-9) PMID: 11275528
8. Brosch M, Budge E, Scheich H. Stimulus-related gamma oscillations in primate auditory cortex. *Journal of Neurophysiology*. 2002; 87(6):2715–2725. <https://doi.org/10.1152/jn.2002.87.6.2715> PMID: 12037173
9. Fell J, Klaver P, Lehnertz K, Grunwald T, Schaller C, Elger CE, et al. Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. *Nature Neuroscience*. 2001; 4:1259–1264. <https://doi.org/10.1038/nn759> PMID: 11694886
10. Osipova D, Takashima A, Oostenveld R, Fernandez G, Maris E, Jensen O. Theta and gamma oscillations predict encoding and retrieval of declarative memory. *Journal of Neuroscience*. 2006; 26(28):7523–7531. <https://doi.org/10.1523/JNEUROSCI.1948-06.2006> PMID: 16837600
11. Jensen O, Kaiser J, Lachaux JP. Human gamma-frequency oscillations associated with attention and memory. *Trends in Neurosciences*. 2007; 30:317–324. <https://doi.org/10.1016/j.tins.2007.05.001> PMID: 17499860
12. Staudigl T, Hanslmayr S. Theta oscillations at encoding mediate the context-dependent nature of human episodic memory. *Current Biology*. 2013; 23(12):1101–1106. <https://doi.org/10.1016/j.cub.2013.04.074> PMID: 23746635
13. Akam T, Kullmann DM. Oscillatory multiplexing of population codes for selective communication in the mammalian brain. *Nature Reviews Neuroscience*. 2014; 15:111–122. <https://doi.org/10.1038/nrn3668> PMID: 24434912
14. Jensen O, Mazaheri A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*. 2010; 4(186).
15. Fries P. Rhythms for cognition: communication through coherence. *Neuron*. 2015; 88(1):220–235. <https://doi.org/10.1016/j.neuron.2015.09.034> PMID: 26447583
16. Miller KJ, Sorensen LB, Ojemann JG, Den Nijs M. Power-law scaling in the brain surface electric potential. *PLoS Computational Biology*. 2009; 5(12). <https://doi.org/10.1371/journal.pcbi.1000609>
17. Ray S, Maunsell JH. Differences in gamma frequencies across visual cortex restrict their possible use in computation. *Neuron*. 2010; 67(5):885–896. <https://doi.org/10.1016/j.neuron.2010.08.004> PMID: 20826318
18. Stokes M, Spaak E. The Importance of single-trial analyses in cognitive neuroscience. *Trends in Cognitive Sciences*. 2016; 20(7):483–486. <https://doi.org/10.1016/j.tics.2016.05.008> PMID: 27237797
19. Lowet E, Roberts MJ, Bosman CA, Fries P, De Weerd P. Areas V1 and V2 show microsaccade-related 3–4-Hz covariation in gamma power and frequency. *European Journal of Neuroscience*. 2016; 43:1286–1296. <https://doi.org/10.1111/ejn.13126> PMID: 26547390
20. Lundqvist M, Rose J, Herman P, Brincat SL, Buschman TJ, Miller EK. Gamma and beta bursts underlie working memory. *Neuron*. 2016; 90:152–164. <https://doi.org/10.1016/j.neuron.2016.02.028> PMID: 26996084
21. Guimaraes MP, Wong DK, Uy ET, Grosenick L, Suppes P. Single-trial classification of MEG recordings. *IEEE Transactions on Biomedical Engineering*. 2007; 54(3):436–443. <https://doi.org/10.1109/TBME.2006.888824> PMID: 17355055
22. Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*. 2013; 7:267. <https://doi.org/10.3389/fnins.2013.00267> PMID: 24431986
23. Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A. Localization of Brain Electrical Activity via linearly constrained minimum variance spatial filtering. *IEEE Transaction on Biomedical Engineering*. 1997; 44(9):867–880. <https://doi.org/10.1109/10.623056>
24. Breiman L. Random forests. *Machine Learning*. 2001; 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
25. Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*. 2011. <https://doi.org/10.1155/2011/156869>
26. Sekihara K, Nagarajan SS. Adaptive spatial filters for electromagnetic brain imaging. Springer Science & Business Media; 2008.

27. Nolte G. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in Medicine and Biology*. 2003; 48(22):3637–3652. <https://doi.org/10.1088/0031-9155/48/22/002> PMID: 14680264
28. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press; 1984.
29. Cutler A, Cutler DR, Stevens JR. Tree-based methods. In: High-Dimensional Data Analysis in Cancer Research. Springer New York; 2009. p. 1–19.
30. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*. 2009; 14(4):323. <https://doi.org/10.1037/a0016973> PMID: 19968396
31. Gini C. Variabilità e Mutuabilità: Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche. Tipografia di Cuppini; 1912.
32. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011; 12:2825–2830.
33. Liaw A, Wiener M. Classification and regression by randomForest. *R news*. 2002; 2(3):18–22.
34. Combrisson E, Jerbi K. Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*. 2015; 30:126–136. <https://doi.org/10.1016/j.jneumeth.2015.01.010>
35. Boser BE, Guyon IM, Vapnik VN. A Training Algorithm for Optimal Margin Classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. ACM; 1992. p. 144–152.
36. Cortes C, Vapnik V. Support-vector networks. *Machine Learning*. 1995; 20(3):273–297. <https://doi.org/10.1007/BF00994018>
37. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated Anatomical Labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*. 2002; 15:273–289. <https://doi.org/10.1006/nimg.2001.0978> PMID: 11771995
38. Leuthardt EC, Schalk G, Wolpaw JR, Ojemann JG, Moran DW. A brain—computer interface using electrocorticographic signals in humans. *Journal of Neural Engineering*. 2004; 1(2):63. <https://doi.org/10.1088/1741-2560/1/2/001> PMID: 15876624
39. Rickert J, de Oliveira SC, Vaadia E, Aertsen A, Rotter S, Mehring C. Encoding of movement direction in different frequency ranges of motor cortical local field potentials. *Journal of Neuroscience*. 2005; 25(39):8815–8824. <https://doi.org/10.1523/JNEUROSCI.0816-05.2005> PMID: 16192371
40. Waldert S, Preissl H, Demandt E, Braun C, Birbaumer N, Aertsen A, et al. Hand movement direction decoded from MEG and EEG. *Journal of Neuroscience*. 2008; 28(4):1000–1008. <https://doi.org/10.1523/JNEUROSCI.5171-07.2008> PMID: 18216207
41. Quandt F, Reichert C, Hinrichs H, Heinze HJ, Knight RT, Rieger JW. Single trial discrimination of individual finger movements on one hand: a combined MEG and EEG study. *NeuroImage*. 2012; 59(4):3316–3324. <https://doi.org/10.1016/j.neuroimage.2011.11.053> PMID: 22155040
42. Fuentemilla L, Penny WD, Cashdollar N, Bunzeck N, Düzel E. Theta-coupled periodic replay in working memory. *Current Biology*. 2010; 20(7):606–612. <https://doi.org/10.1016/j.cub.2010.01.057> PMID: 20303266
43. Schulz E, Zherdin A, Tiemann L, Plant C, Ploner M. Decoding an individual's sensitivity to pain from the multivariate analysis of EEG data. *Cerebral Cortex*. 2011; 22(5):1118–1123. <https://doi.org/10.1093/cercor/bhr186> PMID: 21765182
44. Olivetti E, Kia SM, Avesani P. MEG decoding across subjects. In: Pattern Recognition in Neuroimaging, 2014 International Workshop on. IEEE; 2014. p. 1–4.
45. Fraiwan L, Lweesy K, Khasawneh N, Wenz H, Dickhaus H. Automated sleep stage identification system based on time—frequency analysis of a single EEG channel and random forest classifier. *Computer Methods and Programs in Biomedicine*. 2012; 108(1):10–19. <https://doi.org/10.1016/j.cmpb.2011.11.005> PMID: 22178068
46. Lehmann C, Koenig T, Jelic V, Prichep L, John RE, Wahlund LO, et al. Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *Journal of Neuroscience Methods*. 2007; 161(2):342–350. <https://doi.org/10.1016/j.jneumeth.2006.10.023> PMID: 17156848
47. Bentleimsan M, Zemouri ET, Bouchaffra D, Yahya-Zoubir B, Ferroudji K. Random forest and filter bank common spatial patterns for EEG-based motor imagery classification. In: 5th International Conference on Intelligent Systems, Modelling and Simulation (ISMS). IEEE; 2014. p. 235–238.
48. Donos C, Dümpelmann M, Schulze-Bonhage A. Early seizure detection algorithm based on intracranial EEG and random forest classification. *International journal of neural systems*. 2015; 25(05):1550023. <https://doi.org/10.1142/S0129065715500239> PMID: 26022388

49. Tallon-Baudry C, Bertrand O, Hénaff MA, Isnard J, Fischer C. Attention modulates gamma-band oscillations differently in the human lateral occipital cortex and fusiform gyrus. *Cerebral Cortex*. 2005; 15(5):654–662. <https://doi.org/10.1093/cercor/bhh167> PMID: 15371290
50. Swettenham JB, Muthukumaraswamy SD, Singh KD. Spectral properties of induced and evoked gamma oscillations in human early visual cortex to moving and stationary stimuli. *Journal of Neurophysiology*. 2009; 102(2):1241–1253. <https://doi.org/10.1152/jn.91044.2008> PMID: 19515947
51. Tan HR, Gross J, Uhlhaas P. MEG sensor and source measures of visually induced gamma-band oscillations are highly reliable. *NeuroImage*. 2016; 137:34–44. <https://doi.org/10.1016/j.neuroimage.2016.05.006> PMID: 27153980
52. Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. *Journal of Neurophysiology*. 2005; 94(6):4269–4280. <https://doi.org/10.1152/jn.00324.2005> PMID: 16093343
53. Bidet-Caulet A, Fischer C, Bauchet F, Aguera PE, Bertrand O. Neural substrate of concurrent sound perception: direct electrophysiological recordings from human auditory cortex. *Frontiers in Human Neuroscience*. 2007; 1. <https://doi.org/10.3389/neuro.09.005.2007> PMID: 18958219
54. Canolty RT, Soltani M, Dalal SS, Edwards E, Dronkers NF, Nagarajan SS, et al. Spatiotemporal dynamics of word processing in the human brain. *Frontiers in Neuroscience*. 2007; 1:14. <https://doi.org/10.3389/neuro.01.1.1.014.2007>
55. Edwards E, Soltani M, Kim W, Dalal SS, Nagarajan SS, Berger MS, et al. Comparison of time—frequency responses and the event-related potential to auditory speech stimuli in human cortex. *Journal of Neurophysiology*. 2009; 102(1):377–386. <https://doi.org/10.1152/jn.90954.2008> PMID: 19439673
56. Schepers IM, Hipp JF, Schneider TR, Röder B, Engel AK. Functionally specific oscillatory activity correlates between visual and auditory cortex in the blind. *Brain*. 2012; 135(3):922–934. <https://doi.org/10.1093/brain/aws014> PMID: 22366801
57. Sedley W, Teki S, Kumar S, Overath T, Barnes GR, Griffiths TD. Gamma band pitch responses in human auditory cortex measured with magnetoencephalography. *NeuroImage*. 2012; 59(2):1904–1911. <https://doi.org/10.1016/j.neuroimage.2011.08.098> PMID: 21925281
58. Edwards E, Nagarajan SS, Dalal SS, Canolty RT, Kirsch HE, Barbaro NM, et al. Spatiotemporal imaging of cortical activation during verb generation and picture naming. *NeuroImage*. 2010; 50(1):291–301. <https://doi.org/10.1016/j.neuroimage.2009.12.035> PMID: 20026224
59. Steinschneider M, Nourski KV, Kawasaki H, Oya H, Brugge JF, Howard MA. Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cerebral Cortex*. 2011; 21(10):2332–2347. <https://doi.org/10.1093/cercor/bhr014> PMID: 21368087
60. Lachaux JP, George N, Tallon-Baudry C, Martinerie J, Hugueville L, Minotti L, et al. The many faces of the gamma band response to complex visual stimuli. *NeuroImage*. 2005; 25(2):491–501. <https://doi.org/10.1016/j.neuroimage.2004.11.052> PMID: 15784428
61. Vidal JR, Chaumon M, O'Regan JK, Tallon-Baudry C. Visual grouping and the focusing of attention induce gamma-band oscillations at different frequencies in human magnetoencephalogram signals. *Journal of Cognitive Neuroscience*. 2006; 18(11):1850–1862. <https://doi.org/10.1162/jocn.2006.18.11.1850> PMID: 17069476
62. Siegel M, Donner TH, Oostenveld R, Fries P, Engel AK. High-frequency activity in human visual cortex is modulated by visual motion strength. *Cerebral Cortex*. 2007; 17(3):732–741. <https://doi.org/10.1093/cercor/bhk025> PMID: 16648451
63. Jung J, Mainy N, Kahane P, Minotti L, Hoffmann D, Bertrand O, et al. The neural bases of attentive reading. *Human Brain Mapping*. 2008; 29(10):1193–1206. <https://doi.org/10.1002/hbm.20454> PMID: 17894399
64. Dalal SS, Baillet S, Adam C, Ducorps A, Schwartz D, Jerbi K, et al. Simultaneous MEG and intracranial EEG recordings during attentive reading. *NeuroImage*. 2009; 45(4):1289–1304. <https://doi.org/10.1016/j.neuroimage.2009.01.017> PMID: 19349241
65. Hamamé CM, Vidal JR, Ossandón T, Jerbi K, Dalal SS, Minotti L, et al. Reading the mind's eye: Online detection of visuo-spatial working memory and visual imagery in the inferior temporal lobe. *NeuroImage*. 2012; 59(1):872–879. <https://doi.org/10.1016/j.neuroimage.2011.07.087> PMID: 21839843
66. Hamamé CM, Szwed M, Sharman M, Vidal JR, Perrone-Bertolotti M, Kahane P, et al. Dejerine's reading area revisited with intracranial EEG: Selective responses to letter strings. *Neurology*. 2013; 80(6):602–603. <https://doi.org/10.1212/WNL.0b013e31828154d9> PMID: 23382370
67. Chang EF, Edwards E, Nagarajan SS, Fogelson N, Dalal SS, Canolty RT, et al. Cortical spatio-temporal dynamics underlying phonological target detection in humans. *Journal of Cognitive Neuroscience*. 2011; 23(6):1437–1446. <https://doi.org/10.1162/jocn.2010.21466> PMID: 20465359