NeuroImage 139 (2016) 127–135

Contents lists available at ScienceDirect

NeuroImage

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

Prediction of successful memory encoding based on single-trial rhinal and hippocampal phase information

Marlene Höhne [a], Amirhossein Jahanbekam [a], Christian Bauckhage [b,c], Nikolai Axmacher [d,e], Juergen Fell [a,*]

[a] Department of Epileptology, University of Bonn, D-53105 Bonn, Germany
[b] Bonn-Aachen International Center for Information Technology, University of Bonn, D-53113 Bonn, Germany
[c] Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS, D-53757 Sankt Augustin, Germany
[d] Department of Neuropsychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, D-44801 Bochum, Germany
[e] German Center for Neurodegenerative Diseases (DZNE), D-53175 Bonn, Germany

ARTICLE INFO

Article history:
Received 9 February 2016
Revised 9 May 2016
Accepted 12 June 2016
Available online 14 June 2016

Keywords:
Memory formation
Prediction
Intracranial EEG
Phase
Hippocampus
Rhinal cortex

ABSTRACT

Mediotemporal EEG characteristics are closely related to long-term memory formation. It has been reported that rhinal and hippocampal EEG measures reflecting the stability of phases across trials are better suited to distinguish subsequently remembered from forgotten trials than event-related potentials or amplitude-based measures. Theoretical models suggest that the phase of EEG oscillations reflects neural excitability and influences cellular plasticity. However, while previous studies have shown that the stability of phase values across trials is indeed a relevant predictor of subsequent memory performance, the effect of absolute single-trial phase values has been little explored. Here, we reanalyzed intracranial EEG recordings from the mediotemporal lobe of 27 epilepsy patients performing a continuous word recognition paradigm. Two-class classification using a support vector machine was performed to predict subsequently remembered vs. forgotten trials based on individually selected frequencies and time points. We demonstrate that it is possible to successfully predict single-trial memory formation in the majority of patients (23 out of 27) based on only three single-trial phase values given by a rhinal phase, a hippocampal phase, and a rhinal-hippocampal phase difference. Overall classification accuracy across all subjects was 69.2% choosing frequencies from the range between 0.5 and 50 Hz and time points from the interval between −0.5 s and 2 s. For 19 patients, above chance prediction of subsequent memory was possible even when choosing only time points from the prestimulus interval (overall accuracy: 65.2%). Furthermore, prediction accuracies based on single-trial phase surpassed those based on single-trial power. Our results confirm the functional relevance of mediotemporal EEG phase for long-term memory operations and suggest that phase information may be utilized for memory enhancement applications based on deep brain stimulation.

Introduction

During recent years a growing body of studies has provided evidence for the impact of oscillatory phases of local field potentials (LFPs) and electroencephalographic (EEG) signals on neural processing. LFP/EEG phases interact with neural membrane potentials and thereby modulate the degree of excitability of neurons and influence their discharge times (Elbert and Rockstroh, 1987; Fröhlich and McCormick, 2010; Anastassiou et al., 2010). In this sense, LFP/EEG phases can be thought of as facilitating or impeding the occurrence of neural activity within a required time window or processing stage (e.g. Fell and Axmacher, 2011).

Indeed, several investigations have shown that LFP/EEG phases affect perceptual and cognitive operations. For instance, the phases of alpha oscillations of scalp EEG were reported to be predictive for visual perception of stimuli close to the detection threshold (Busch et al., 2009; Mathewson et al., 2009). Importantly, it has been demonstrated that transcranial alternating current stimulation modulates visual and acoustic detection thresholds depending on local phases and phase differences between regions suggesting a causal role of phase dynamics (Neuling et al., 2012; Helfrich et al., 2014).

With regard to memory operations it is well-known that the phase of theta oscillations within the hippocampus determines the direction and magnitude of synaptic plasticity. In rats, electrical stimulation at the peak of hippocampal theta oscillations facilitates long-term potentiation, whereas stimulation at the trough induces long-term depression (Pavlides et al., 1988; Huerta and Lisman, 1993). Moreover, stimulus-related phase reset of low-frequency oscillations has been reported to

* Corresponding author at: Department of Epileptology, University of Bonn, Sigmund-Freud-Str. 25, D-53105 Bonn, Germany.
E-mail address: juergen.fell@ukb.uni-bonn.de (J. Fell).

http://dx.doi.org/10.1016/j.neuroimage.2016.06.021
1053-8119/

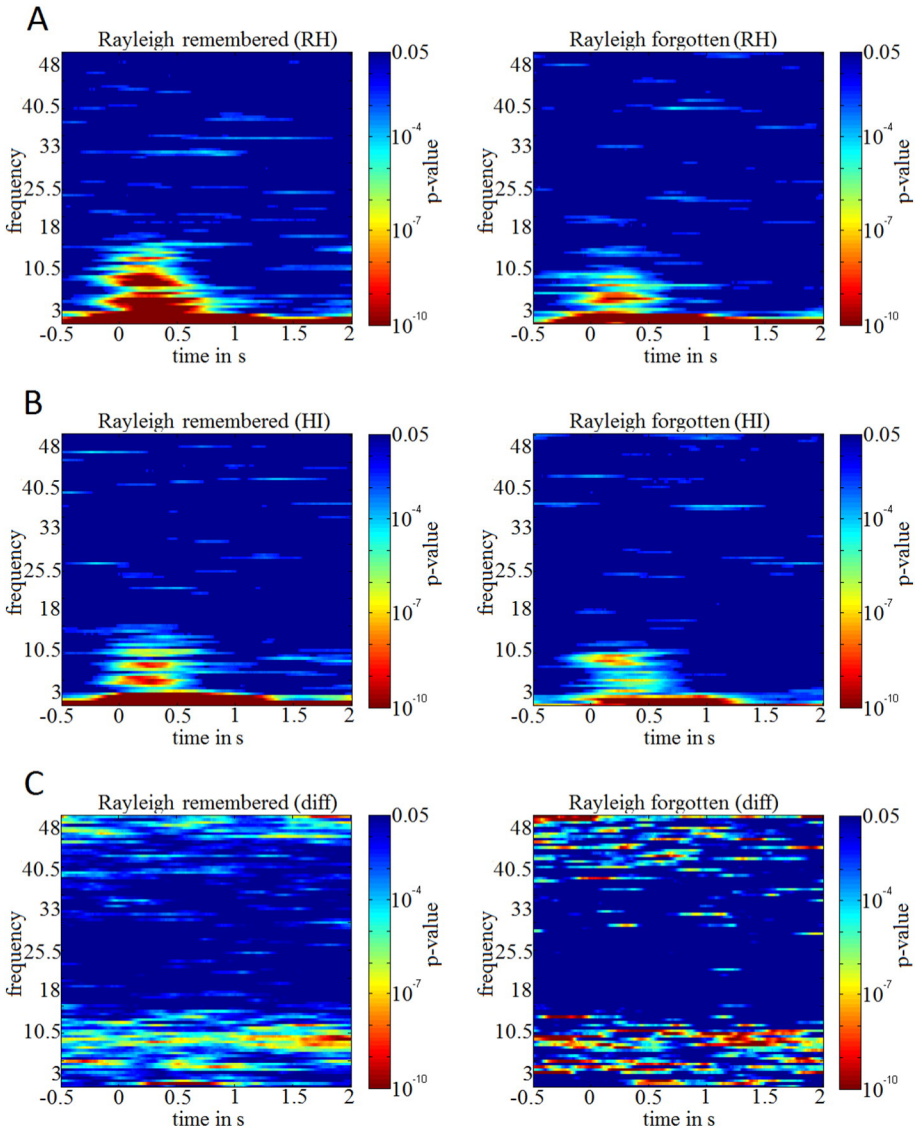

Fig. 1. Results of the Rayleigh tests. The figure shows the Fisher combined p-values of Rayleigh tests for the phase values within rhinal cortex (A) and hippocampus (B) as well as for the phase differences between rhinal cortex and hippocampus (C) under the conditions "later remembered" (left column) and "later forgotten" (right). Colors indicate p-values according to a logarithmic scale, with all values >0.05 colored in dark blue.

be an essential characteristic of memory operations (e.g. Rizzuto et al., 2003; Mormann et al., 2005; Haque et al., 2015). Furthermore, phase information derived from mediotemporal lobe (MTL) recordings in epilepsy patients was found to be superior to amplitude information for a classification of correct versus incorrect trials in a card-matching task (Lopour et al., 2013).

In a previous study, we have investigated how closely different mediotemporal EEG measures are related to memory formation (Fell et al., 2008). For this purpose, we analyzed intracranial data from 31 epilepsy patients performing a continuous word recognition paradigm. EEG measures comprised traditional average event-related potential (ERP) characteristics, rhinal and hippocampal power changes within different frequency bands, as well as inter-trial phase locking and rhinal-hippocampal phase synchronization. This analysis revealed that phase-based measures (i.e. inter-trial phase-locking and phase-synchronization), which reflect the stability of phase values and phase differences across trials, are better suited to distinguish subsequently remembered from forgotten trials than ERP or amplitude-based measures. Based on theoretical considerations there should be an optimal phase, as well as less optimal or unsuitable phases with regard to the facilitation of neural communication and plasticity (e.g. Fell and Axmacher, 2011). This suggests that phases for subsequently remembered compared to forgotten trials may be centered around different values, which, however, cannot be deduced from the previous finding that phases are more strongly accumulated for later remembered trials (they nevertheless could be centered around the same value). Thus, it remained an open question whether single-trial phase values per se are predictive for memory encoding.

For the present study, we therefore reanalyzed encoding-related responses for subsequently remembered and forgotten words in the same paradigm (Fell et al., 2008, 2011). In a first step, we identified time windows and frequencies with statistically significant phase clustering across patients. Then we determined for each patient time periods and frequencies for which the absolute phases and inter-electrode phase differences differ between the remembered and forgotten condition. Finally, a support vector machine (SVM) was trained by using the phases and phase differences from the most significant time windows and frequencies. Importantly, we aimed to employ a minimal set of features to predict subsequent memory, on the one hand, for ease of exposition, on the other hand, because such an approach is most closely related to possible practical applications (e.g. controlling one of the features by deep brain stimulation). Furthermore, we investigated whether prediction

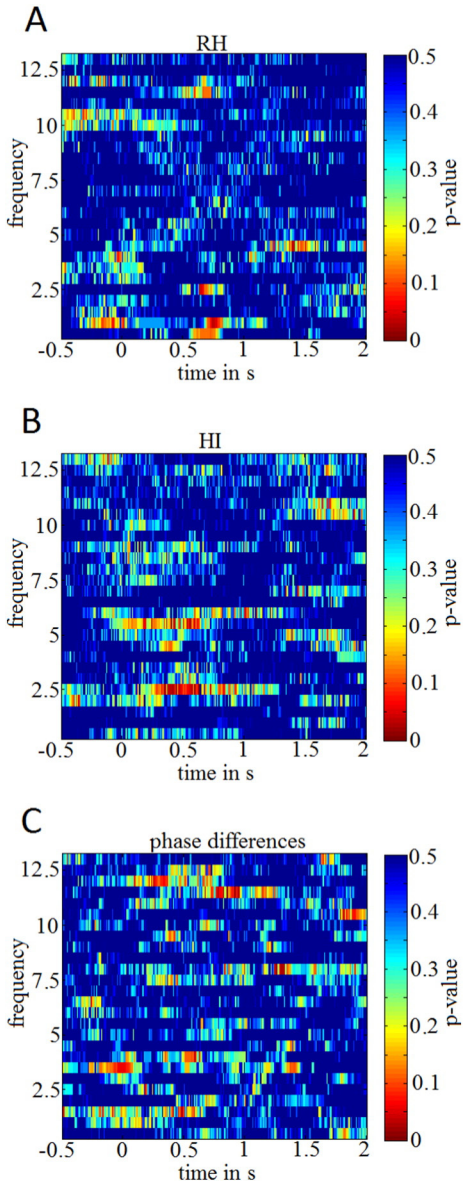

Fig. 2. Results of the testing for differences between the conditions "later remembered" and "later forgotten" for one exemplary subject (pat 13). The figure shows the p-values of the tests for frequencies up to 13 Hz for the phase values within rhinal cortex (A) and hippocampus (B) as well as for rhinal-hippocampal phase differences (C).

based on single-trial power outperforms prediction based on single-trial phase, as suggested by our previous findings (Fell et al., 2008). Our data reveal that in the majority of patients successful single-trial memory encoding can be predicted based on only three single-trial phase values given by a rhinal phase, a hippocampal phase, and a rhinal-hippocampal phase difference and that prediction accuracies based on single-trial phase surpass those based on single-trial power.

Materials and methods

Patients

We investigated data from 31 right-handed patients (14 females) with an average age of 40 years (from 16 to 61) who suffered since 4 to 57 years (mean 23 years) from pharmacoresistant unilateral temporal lobe epilepsies (see also Fell et al., 2008). Patients had been implanted with bilateral depth electrodes along the longitudinal axis of the hippocampus during presurgical evaluation. All patients had at least one electrode contact in the rhinal cortex and one in the hippocampus. The word recognition test was performed as part of the presurgical routine and all patients received anticonvulsive medication (plasma levels within the therapeutic range) at the time of the recordings. Each patient provided informed consent to participate in the study, which was approved by the ethics committee of the Medical Faculty at the University of Bonn. Post-surgical histological examinations or magnetic resonance imaging (MRI) scans indicated unilateral hippocampal sclerosis in 16 patients (left: 5; right: 11), unilateral extrahippocampal lesions without signs of hippocampal sclerosis in 9 patients (left: 3; right: 6), unilateral hippocampal sclerosis with additional extrahippocampal lesions on the same side in 3 patients (left: 2; right: 1) and no clear lesion in 3 patients. All but two patients underwent subsequent epilepsy surgery after implantation.

Experimental paradigm

The stimuli of the continuous word recognition paradigm consisted of 300 frequent German nouns and were consecutively presented with a duration of 300 ms per word. A total of 450 words were

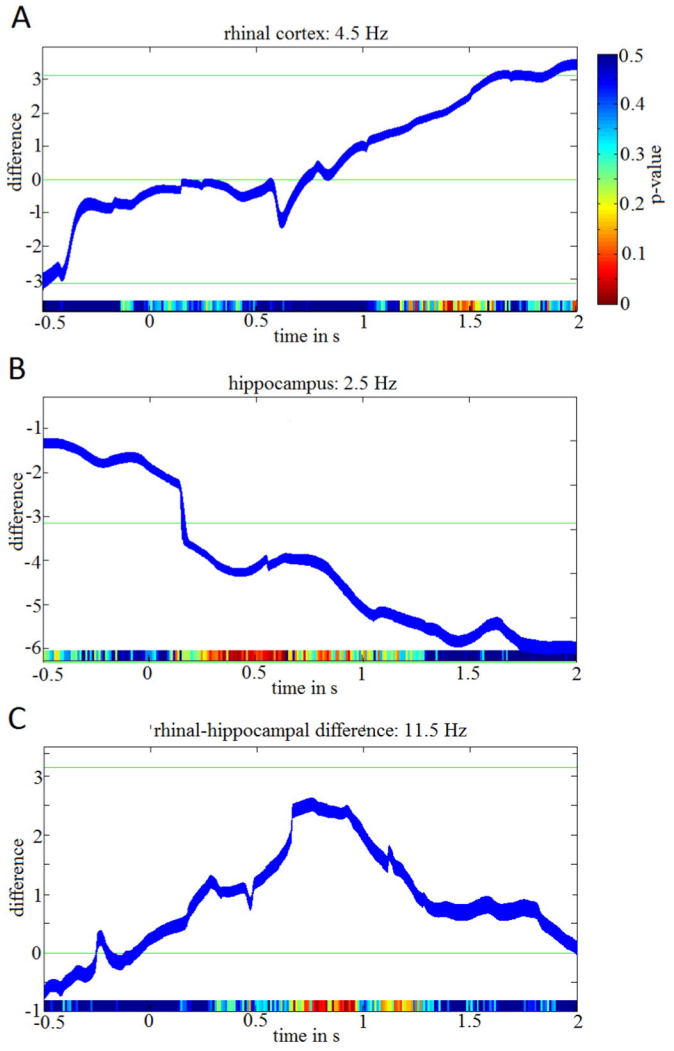

Fig. 3. Mean phase differences between the conditions "later remembered" and "later forgotten" over time for one exemplary subject (pat13). The phase differences are averaged over trials for the rhinal phase (A) and hippocampal phase (B) as well as rhinal-hippocampal phase difference (C) at the frequency that was selected for classification. Line width shows circular variance reduced by factor 5. The colored line at the bottom indicates the p-values of the tests for differences between conditions. Green lines mark zero and ± pi.

Table 1
Frequencies and time points chosen as features for classification in each patient (for at least 4 of 5 folds). The listed time points specify the starting point of the used 10 ms time interval. The left part of the table lists the selection for frequencies up to 13 Hz, the right part up to 50 Hz. Abbreviations: RH (rhinal cortex), HI (hippocampus), diff (difference).

	Up to 13 Hz						Up to 50 Hz					
Pat	Freq RH	Time RH	Freq HI	Time HI	Freq diff	Time diff	Freq RH	Time RH	Freq HI	Time HI	Freq diff	Time diff
1	9,5	150	7	530	4	1870	34	−290	27,5	1300	27	1180
2	7	290	7,5	460	8,5	120	28,5	400	48	1050	19	1130
3	7	1100	8	1910	5	1740	22,5	730	19,5	−130	45,5	1440
4	12	390	11	820	6,5	−90	12	390	33,5	1050	41	270
5	5,5	40	8	80	10	990	5,5	40	22	−280	10	990
6	7	1510	6,5	1950	0,5	1040	29	1420	16	690	13,5	−290
7	8	210	12	350	5	−360	48	1210	45	940	41	170
8	1,5	−370	3	640	6	1370	27	1340	37,5	−370	40	1230
9	11,5	−160	13	−320	0,5	1360	11,5	−160	13	−320	0,5	1360
10	10	770	6,5	−200	10	130	49	−120	24,5	380	10	130
11	0,5	1650	1,5	480	0,5	1630	0,5	1650	1,5	480	44,5	60
12	0,5	1200	3,5	−20	7	−80	24	100	3,5	−20	45,5	1360
13	4,5	1420	2,5	640	11,5	920	41,5	520	2,5	640	27	−130
14	2	1000	0,5	1220	13	1300	35	580	0,5	1220	13,5	310
15	13	1740	1,5	−410	6,5	−270	45	620	30	−440	29,5	550
16	11,5	890	0,5	1040	2,5	1730	11,5	890	0,5	1040	23,5	100
17	2	580	0,5	1920	2	1240	2	580	45,5	980	2	1240
18	9	30	0,5	−440	3,5	760	9	30	40,5	1270	43	1750
19	3	1320	9,5	−220	1,5	1420	47	280	9,5	−220	29,5	870
20	0,5	900	1,5	−330	5	1680	0,5	900	44	360	23	320
21	4,5	−150	10,5	−300	8	1670	4,5	−150	10,5	−300	15	1490
22	2	880	1,5	−410	13	1880	33	370	29,5	1780	13	1880
23	12	1870	10,5	1240	1,5	−50	12	1870	10,5	1240	1,5	−50
24	4	1660	4	1020	3	−440	39,5	1650	13,5	1020	23,5	510
25	9	550	5,5	1130	4,5	−170	45	570	40	1240	46,5	1890
26	9	1190	0,5	770	2,5	1850	49	540	0,5	770	29,5	620
27	12,5	420	1,5	950	10	110	15	540	1,5	950	25	1310

presented in white color on a black background. One hundred and fifty stimuli were presented twice and the other 150 words were only shown once. Between the first and the second presentation in 50% of the trials there was a short lag of 3 to 6 words and in 50% a long lag of 10 to 30 words. The length of the inter-stimulus interval was adjusted to the subjects' abilities (assessed from the responses in a few pilot trials) and was either short (1600 ± 200 ms; n = 6), middle (2000 ± 200 ms; n = 16) or long (2700 ± 200 ms; n = 9). After each presentation, subjects had to decide if they had seen the word before or not using one of two buttons which they pressed with their right (old) and left (new) forefingers. If performance was bad with only a small amount (<30 correctly recognized "old" or "new" words) of evaluable trials or if ERPs were contaminated by spikes or sharp waves, recordings were repeated with a parallel version of the recognition task on the following day. In these cases, the data of the second recordings were used for the analyses.

EEG recordings

Data were recorded at a sampling frequency of 200 Hz, referenced to linked mastoids and bandpass-filtered from 0.01 Hz (6 dB/octave) to 70 Hz (12 dB/octave). The placement of electrode contacts was ascertained based on the individual MRIs and comparison with standardized anatomical atlases (e.g. Duvernoy, 1988). Only recordings from the non-pathological MTL were included in the analysis. The hippocampal electrode was defined as the electrode located within the hippocampus (based on the MRI data) with the largest mean amplitude (new words) of the positive component between 300 and 1500 ms (e.g. Fernández et al., 1999). The rhinal electrode was defined as the electrode located within the anterior parahippocampal gyrus (based on the MRI data) with the largest N400 mean amplitude (new words) between 200 and 600 ms (e.g. Grunwald et al., 1999). Because lateralization of verbal memory in MTL epilepsy patients is variable due to functional shifts (e.g. Helmstaedter et al., 2006), EEG measures from right and left hemisphere were combined for statistical analyses and figures.

Artifact rejection

Trials that included abnormally high amplitudes as well as abrupt rises or falls were removed by an automated artifact rejection algorithm implemented in MATLAB (Mathworks, MATLAB 7.1). The segments of both contacts (rhinal and hippocampal) were eliminated if at least one segment showed data points or gradients (differences between two consecutive data points) diverging more than five standard deviations from the mean. On average, this resulted in removing 14% of the trials. Four patients were excluded from further analysis because their data still showed artifacts (observed by visual inspection) after the automatic artifact rejection leaving 27 patients for classification.

Categorization of trials

We analyzed the EEG responses to the first presentation of words shown with one repetition. Responses were classified into "later remembered" or "later forgotten" depending on whether the word was subsequently (i.e. at the second presentation) correctly identified (i.e. correctly recognized as "old") or not (i.e. wrongly labeled as "new").

Extraction of phase and power values and analysis of phase effects

The free FieldTrip toolbox for MATLAB was used for the extraction of phase values (Oostenveld et al., 2011). EEG responses were filtered in the frequency range from 0.5 Hz to 50 Hz (0.5 Hz steps) by a second order Butterworth filter with a bandwidth of 1 Hz. The complex discrete-time analytic signal was determined by the Hilbert transform of the signals to obtain the phase values. In order to avoid edge effects, EEG responses were segmented from −1000 ms to 2500 ms with respect to stimulus onset, and after Hilbert-transform 500 ms at both sides were discarded.

Based on the complex signals $w_{j,k,}$ the phases $\Phi_{j,k} = \arctan(\mathrm{Im}(w_{j,k})/\mathrm{Re}(w_{j,k}))$ and the phase differences between rhinal cortex and hippocampus $\Delta_{j,k} = \Phi_{j,k}(\mathrm{RH}) - \Phi_{j,k}(\mathrm{HI})$ were extracted for each time point j of each trial k. The phases spanned the range $[0, 2\pi)$ with zero representing the

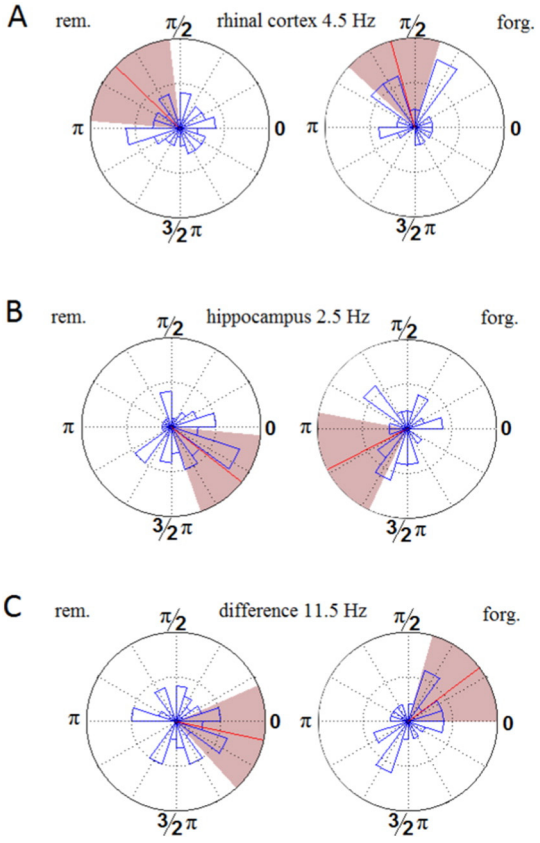

Fig. 4. Phase values from the rhinal cortex (A) and the hippocampus (B) as well as the rhinal-hippocampal phase difference (C) used for the training of the classifier for one exemplary subject (pat13). The figure shows rose diagrams of the values for the features selected for frequencies up to 13 Hz. The values of the condition "later remembered" can be found in the left column and "later forgotten" in the right one. Mean phases are marked with a red line; the angular deviation is displayed by shaded areas.

peak and π the trough of the oscillation. Additionally, rhinal and hippocampal power values $Pow_{j,k} = (Re(w_{j,k})^2 + Im(w_{j,k})^2)$ were extracted. Phase and power values were averaged for non-overlapping successive time windows of 10 ms duration from −500 to 2000 ms (250 windows in total) for each trial and frequency.

For circular statistics the free CircStat toolbox for MATLAB was used (Berens, 2009). First a Rayleigh test (function circ_rtest) was performed for each time window and each filter frequency separately for "later remembered" and "later forgotten" trials. A significant Rayleigh test indicates that phases are not uniformly distributed but exhibit significant phase accumulations. To identify overall effects Rayleigh tests were performed for each patient individually and were then combined with Fisher's method (Neuhäuser, 2011). This is a statistical procedure testing a hypothesis for a collective based on the results of independent statistical tests for the individuals of the collective.

Prediction of subsequent memory

To identify frequencies and time intervals with significant differences of rhinal and hippocampal phase values and of rhinal-hippocampal inter-electrode phase differences between "later remembered" and "later forgotten" trials, for each patient a non-parametric multi-sample test for equal circular medians similar to a Kruskal-Wallis test for linear data was performed on the training data set (60% randomly selected trials; 20% validation trials; 20% test trials; five-fold cross-validation; function circ_cmtest). In other words, only 60% of the data were used for testing for differences in median phase direction and these data had no overlap with the test data used for classification. Two frequency ranges were considered for classification and their results were compared. First the frequencies from 0.5 to 13 Hz were used based on the overall result of the Rayleigh test (most pronounced phase accumulations were found for the frequency range between 0.5 and 13 Hz, see Fig. 1). Second, all frequencies up to 50 Hz were included. For each patient the frequencies and time windows with the 10 most significant differences between conditions were preselected as features. Alternatively, the considered time windows were restricted to the prestimulus interval (−500 to 0 ms) and features were selected accordingly.

To further reduce the number of preselected features a support vector machine with a linear kernel was applied to a validation data set (20% trials) classifying the trials into the categories "later remembered" and "later forgotten". Based on the highest prediction accuracies in the validation data one rhinal phase value, one hippocampal phase value and one rhinal-hippocampal phase difference were selected as final features for classification of the test data (remaining 20% trials). Because phase is a circular quantity, the real and imaginary part Re(φ(t)) and Im(φ(t)) of the complex representation of the phases were entered as features instead of the phase φ(t), resulting in a doubling of the number of features.

Because of the unbalanced and small trial numbers the classification procedure was performed using five-fold cross-validation with adjusted numbers of randomly chosen training trials, i.e. for the condition with the higher number of trials n training trials were randomly chosen with n being the number of trials in the other condition. Average accuracies from these five cross-validations are reported.

Classification efficiency was evaluated by a non-parametric label permutation approach. Group labels were randomly shuffled 1000 times and then these surrogate trials were classified again for all five-folds. The statistical significance of above chance classification performance was evaluated by ranking the mean accuracy of the real data within the accuracies obtained from the label shuffled data. Additionally, to investigate whether prediction based on single-trial phase outperforms prediction based on power, the same procedures which were applied to rhinal and hippocampal phase values and phase differences were independently applied to rhinal and hippocampal power values.

Results

Behavioral responses

On average, presented words were successfully remembered in 66.7 ± 21.3% (mean ± s.d.) of all cases, i.e. 66.7% of the repeated words were correctly recognized as old (hits). 23.8% ± 30.7% of all new words were wrongly categorized as old (false alarms). Hit minus false alarm rate was significantly above zero (paired *t*-test; $p < 10^{-7}$). Reaction times at the time of encoding did not differ between subsequently remembered and forgotten words (remembered: 878 ± 161 ms; forgotten: 882 ± 232 ms; paired *t*-test $t_{30} = 0.175$, $p = 0.86$).

Phase accumulation

First, we located the frequency bands where phase accumulation occurred across subjects. Phase values within rhinal cortex and hippocampus and phase differences between the rhinal cortex and hippocampus were calculated for all frequencies from 0.5 Hz to 50 Hz (0.5 Hz steps) and for non-overlapping 10 ms time windows. Phase accumulations were identified by conducting a Rayleigh test for each patient individually and combining p-values with Fisher's method (Neuhäuser, 2011). For both, phase values and phase differences, the most pronounced accumulations were found for the low frequency range up to 13 Hz, predominantly in the time range between −200 ms and 800 ms (see Fig. 1). For phase differences additional accumulations were observed in

Table 2
Frequencies and time points chosen as features for classification in each patient limited to prestimulus time range (for at least 4 of 5 folds). The listed time points specify the starting point of the used 10 ms time interval. The left part of the table lists the selection for frequencies up to 13 Hz, the right part up to 50 Hz. Abbreviations: RH (rhinal cortex), HI (hippocampus), diff (difference).

	Up to 13 Hz						Up to 50 Hz					
Pat	Freq RH	Time RH	Freq HI	Time HI	Freq diff	Time Diff	Freq RH	Time RH	Freq HI	Time HI	Freq diff	Time diff
1	5,5	−240	3	−290	6,5	−50	34	−290	34	−230	29	−130
2	0,5	−410	8	−130	8,5	−10	0,5	−410	8	−130	20,5	−10
3	12,5	−50	9	−220	12	−60	31,5	−30	19,5	−130	50	−50
4	2	−290	13	−390	6,5	−90	30,5	−120	14,5	−400	20	−230
5	5,5	−10	10,5	−280	11	−70	5,5	−10	22	−280	11	−70
6	2	−230	4	−330	3,5	−180	48,5	−410	23	−220	13,5	−290
7	4	−430	11,5	−360	5	−360	4	−430	17,5	−350	5	−360
8	1,5	−370	11	−30	9,5	−190	1,5	−370	37,5	−370	43,5	−270
9	11,5	−160	13	−320	0,5	−320	11,5	−160	13	−320	0,5	−320
10	2	−290	6,5	−200	10,5	−200	49	−120	6,5	−200	39	−420
11	2	−50	9,5	−320	5	−110	2	−50	34,5	−230	18,5	−440
12	10,5	−390	3,5	−20	7	−80	44	−190	3,5	−20	7	−80
13	12	−380	2	−350	1,5	−80	20	−180	40	−10	27	−130
14	9,5	−50	1,5	−20	3,5	−440	19,5	−300	15,5	−10	41,5	−310
15	7,5	−360	1,5	−410	6,5	−270	26,5	−80	30	−440	26	−10
16	10,5	−80	4	−270	4	−330	17	−230	49,5	−320	49,5	−440
17	8	−70	1,5	−40	5,5	−270	22	−330	36,5	−420	16,5	−320
18	9	−30	0,5	−440	4,5	−420	9	−30	0,5	−440	37,5	−200
19	8,5	−400	9,5	−220	8,5	−90	8,5	−400	9,5	−220	46,5	−290
20	3,5	−60	1,5	−330	8,5	−150	46,5	−300	39	−40	45,5	−20
21	4,5	−150	10,5	−300	13	−30	4,5	−150	10,5	−300	13	−30
22	11	−240	1,5	−410	5	−140	19,5	−160	1,5	−410	25	−340
23	2,5	−180	2,5	−130	1,5	−50	21	−10	2,5	−130	1,5	−50
24	4	−320	5,5	−100	3	−440	48,5	−420	37	−320	3	−440
25	9	−210	2,5	−160	4,5	−170	49,5	−360	50	−150	34,5	−270
26	3,5	−290	5,5	−170	3,5	−380	24,5	−370	5,5	−170	39	−80
27	1,5	−20	8,5	−180	5	−370	1,5	−20	42,5	−90	14,5	−200

the gamma range between 40 Hz and 50 Hz indicating synchronization between rhinal cortex and hippocampus with a consistent coupling phase (phase lags were clustered around zero; see Supplementary material for data and control analyses addressing a possible influence of volume conduction).

Differences between conditions

Next, we used a nonparametric multi-sample test for equal medians (circular version of the Kruskal-Wallis test) to identify the frequencies and time intervals (width 10 ms) with significant differences between the conditions "later remembered" and "later forgotten" for each patient. On average, patients show significant differences between conditions ($p < 0.05$) in 2.9 ± 1.9 frequencies per measure (rhinal and hippocampal phase values, rhinal-hippocampal differences) with a mean length of significant intervals of 36 ± 24 ms regarding frequencies up to 13 Hz and in 8.9 ± 6.1 frequencies with a mean length of 38 ± 14 ms regarding the frequency range up to 50 Hz.

The test results for one exemplary patient are shown in Fig. 2 (frequency range up to 13 Hz). For this patient the most significant differences between conditions were detected at a frequency of 4.5 Hz for rhinal phase, at a frequency of 2.5 Hz for hippocampal phase, and at a frequency of 11.5 Hz for the rhinal-hippocampal phase difference (please see Supplementary material for exemplary results of two other patients, one with more pronounced condition differences and one with less pronounced differences).

Fig. 3 shows the differences between conditions averaged across trials for these three frequencies (please see Supplementary material for two other examples). In this example, the rhinal phase difference is slightly negative for times up to 600 ms and then drifts to increasingly positive values up to π. For the hippocampus, the phase difference drifts from close to zero during the prestimulus time range towards −π and further to −2π in the poststimulus range. The condition difference of rhinal-hippocampal phase differences starts from slightly negative values in the prestimulus range and then drifts to values up to π and afterwards back to zero in the poststimulus range.

Classification results

Based on the results of the Rayleigh tests, features for classification were first chosen from the frequency range up to 13 Hz and alternatively from the extended frequency range up to 50 Hz. For each measure, i.e. hippocampal phase, rhinal phase and rhinal-hippocampal phase difference, one frequency and time interval was selected based on the most significant differences between conditions in the circular version of the Kruskal-Wallis test and based on the highest classification accuracies in the validation data. Table 1 gives a list of the frequencies and time points chosen as features for each patient considering the whole time range. The selected phase values for the exemplary patient are shown in Fig. 4 (frequencies as above; please see supplementary material for two other examples). The rhinal phases concentrate at an angle of 2.37 ± 1.37 (average angle in radians ± angular deviation) for the "later remembered" and at 1.84 ± 1.12 for the "later forgotten" condition. Hippocampal phases concentrate at 5.62 ± 1.14 versus 3.61 ± 1.32 and rhinal-hippocampal phase differences at 6.07 ± 1.25 versus 0.65 ± 1.30. Table 2 gives a list of selected frequencies and time points with time windows considered for feature selection limited to the prestimulus range. We cannot exclude that the Butterworth filtering may have caused some temporal smearing of poststimulus activity into the prestimulus domain. For the chosen filter characteristics, based on tests with simulated signals, such temporal smearing may extend up to half the cycle length of the filter frequency (e.g. 100 ms for 5 Hz). Accordingly, this consideration applies to 39 (24.1%) of the $6 \times 27 = 162$ values listed in Table 2.

The individual accuracies of correct classifications into the categories "later remembered" and "later forgotten" with a support vector machine (see methods) are shown in Fig. 5. By using features from the frequency range up to 13 Hz the overall classification accuracy (averaged over all 27 subjects) achieved 66.2%. Based on comparison with label

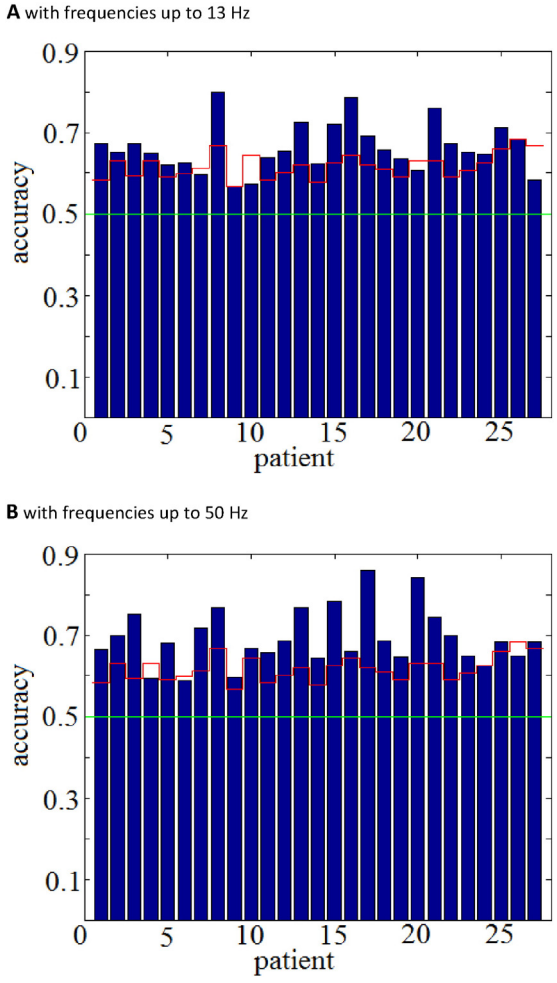

Fig. 5. Individual classification accuracies for each patient. Red lines mark the individual 95% thresholds; the green line marks the 50% accuracy. (A) Included frequencies up to 13 Hz. (B) Included frequencies up to 50 Hz.

shuffled surrogate data (see methods) the classifier gave individual classification results significantly above chance for 21 subjects. Including the frequencies up to 50 Hz for feature selection, above chance results were achieved for 23 subjects with an overall classification accuracy of 69.2%. Regarding only subjects with above chance classification, the average accuracy for frequencies up to 13 Hz was 67.9% and for frequencies up to 50 Hz it was 70.6%.

When the time range for feature selection was limited to the prestimulus interval, the average classification accuracy for the frequency range up to 13 Hz was 61.2% with above chance classification for 15 subjects and accuracy reached 65.2% for frequencies up to 50 Hz with above chance results for 19 subjects. The corresponding individual accuracies are shown in Fig. 6.

To assess the predictive capabilities of the three different measures we performed classifications based on inclusion of only one measure selected from the complete time range. For the frequency range up to 13 Hz, the ranking of classification accuracies revealed rhinal-hippocampal phase difference as most predictive measure (63.7%), followed by hippocampal phase (62.2%) and rhinal phase (61.9%). However, across subjects these accuracies are not significantly different from each other (repeated measures ANOVA: $F_{2,52} = 0.545$; $p > 0.5$). For the frequency range up to 50 Hz, hippocampal phase predicted successful memory performance most accurately (64.5%), followed by rhinal-hippocampal phase difference (63.6%) and rhinal phase (63.3%). Again, these accuracies are not significantly different from each other (repeated measures ANOVA: $F_{2,52} = 0.254$; $p > 0.70$). In accordance to previous results (e.g. Fell et al., 2001) rhinal-hippocampal phase differences were distributed around zero (see supplementary material for data and control analyses addressing a possible influence of volume conduction). For the selected frequency and time intervals the phase differences on average were slightly negative for later remembered trials (-0.25 ± 1.02) and slightly positive for later forgotten trials (0.35 ± 0.86, circular Kruskal-Wallis: $p = 0.057$).

Finally, we investigated the classification performance of single-trial power values by applying the same procedures as for the phase values. For the frequency range up to 50 Hz, overall classification accuracy was 60.4% for rhinal power (vs. 63.3% for rhinal phase) and 61.5% for hippocampal power (vs. 64.5% for hippocampal phase). Across subjects prediction accuracies for classification based on single-trial phase were significantly higher than those based on single-trial power (two-way repeated measures ANOVA, main effect for measure (phase/power), $F_{1,52} = 6.865$; $p = 0.012$; no main effect for locus (rhinal cortex/hippocampus) and no interaction measure × locus). Prediction accuracy surpassed chance level in 13 subjects for rhinal power (vs. 18 subjects for rhinal phase) and in 15 subjects for hippocampal power (vs. 18 subjects for hippocampal phase). Combining the two single-trial power-based features (rhinal and hippocampal power) and the three phase-based features (rhinal and hippocampal phase plus rhinal-hippocampal phase difference) resulted in an overall prediction accuracy of 71.2% compared to 69.2% for only the phase-based features (no significant difference across subjects; paired *t*-test, $p > 0.25$).

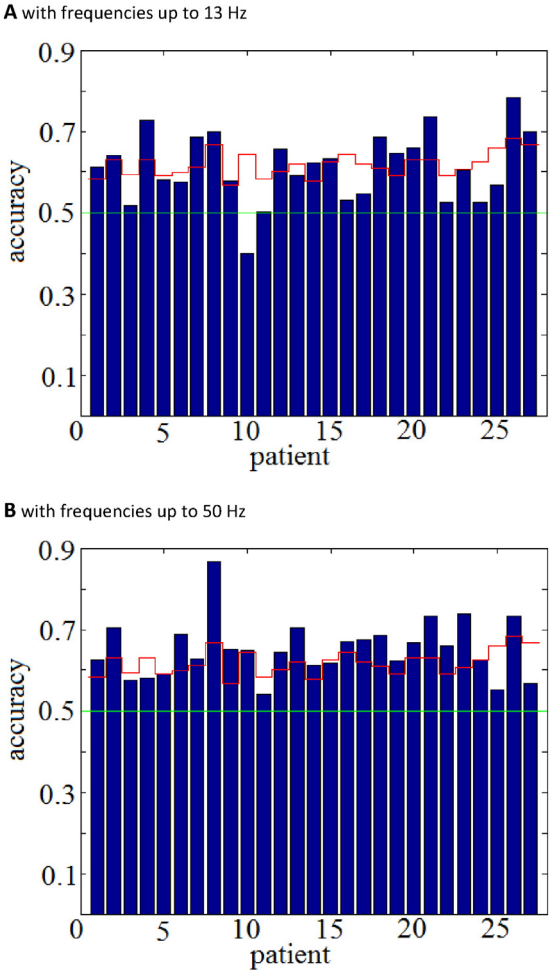

Fig. 6. Individual classification accuracies for each patient for prestimulus intervals. Red lines mark the individual 95% thresholds; the green line marks the 50% accuracy. (A) Included frequencies up to 13 Hz. (B) Included frequencies up to 50 Hz.

Discussion

In a prior investigation, we had analyzed the relation of different mediotemporal EEG measures to memory formation and had found that phase-based measures quantifying the stability of phases and of phase differences across trials outperform other measures in distinguishing subsequently remembered from forgotten trials (Fell et al., 2008). It remained open, however, whether single-trial phase values per se are predictive for memory encoding. The present data show that in 23 out of 27 patients (85%) single-trial memory formation can be predicted above chance based on only one rhinal phase value, one hippocampal phase value and a rhinal-hippocampal phase difference. Moreover, in the majority of patients (19 out of 27) prediction of successful memory encoding was even possible when only phase values from the prestimulus interval were used. As a note of caution, we can not exclude that the Butterworth filtering may have caused some temporal smearing of poststimulus activity into the prestimulus domain (see Classification results).

In accordance with our findings, several studies have shown that prestimulus electrophysiological activity predestinates memory formation. For instance, increased hippocampal theta activity before stimulus presentation is associated with successful memory encoding (Fell et al., 2011; Guderian et al., 2009). Other studies have demonstrated that prestimulus ERP measures are related to subsequent memory performance (for an overview, see Cohen et al., 2015). Recently, Haque et al. (2015) have reported based on intracranial EEG recordings that prestimulus power increases in the 2–4 Hz range and concomitant phase synchronization enhancements precede successful memory encoding. This activity pattern especially involved the temporo-parietal junction, bilateral prefrontal cortex and the medial temporal lobe.

Applying single-trial based classification methods, Noh et al. (2014) recently attempted to predict subsequent memory performance based on high-resolution surface EEG recordings. Data from an object recognition experiment using pictures of cars and birds as stimuli were analyzed. Features entering classification were pre- and peristimulus event-related potentials and EEG power in nine different frequency bands. Applying linear and SVM classifiers Noh et al. (2014) reported an average classification accuracy of 59.6% across 18 subjects (with a chance level of 50%). Here, we demonstrate an average accuracy of 69.2% across 27 subjects based on SVM classification using only rhinal and hippocampal phase values as features. Furthermore, in accordance with previous findings related to inter-trial phase-stability (Fell et al., 2008), we observed that prediction of subsequent memory based on single-trial phase significantly outperformed prediction based on single-trial power. Hence, our results confirm the importance of mediotemporal EEG phase for long-term memory operations.

But what might be the functional relevance of rhinal and hippocampal phase values for memory formation? First of all, EEG phases reflect – and potentially even influence – neural membrane potentials and are thus related to the amount of neural excitability (Elbert and Rockstroh, 1987; Fröhlich and McCormick, 2010; Anastassiou et al., 2010). In this sense, an optimal EEG phase may indicate that neural activity occurs within the time window required for a certain perceptual or cognitive processing step. In case of a non-optimal phase, processing may be hampered, as for instance has been shown for visual perception of stimuli close to the detection threshold (Busch et al., 2009; Mathewson et al., 2009).

With regard to memory operations there are other additional putative functions of EEG phase, in particular within the MTL. It has been demonstrated in rodents that the phase of low-frequency hippocampal oscillations correlates with the direction of synaptic changes (Pavlides et al., 1988; Huerta and Lisman, 1993). Furthermore, the synchronization of phases between rhinal cortex and hippocampus is closely related to long-term memory formation (Fell et al., 2001, 2008). Two complementary mechanisms may contribute to this phenomenon. Rhinal-hippocampal phase differences may modulate communication between rhinal cortex and hippocampus, as well as initiate spike-timing dependent plasticity via spike-field coupling (Fries, 2005; Fell and Axmacher, 2011).

Recently, the prospects of memory enhancement by deep brain stimulation have gained increasing interest (e.g. Lee et al., 2013; Suthana and Fried, 2014; Reardon, 2015). Can the notion of the relevance of rhinal and hippocampal EEG phases be utilized for the purpose of memory enhancement? Since oscillatory phases continuously progress, the control of rhinal and hippocampal EEG phases requires knowledge of the exact time point at which a stimulus occurs. Hence, controlling rhinal or hippocampal EEG phase in a stimulus-related manner is only feasible in experimental settings, but not in ecologically realistic situations, where the exact time point of stimulus appearance is uncertain. Controlling the rhinal-hippocampal phase difference is a more viable option, because phase difference may remain relatively stable for longer time intervals. Indeed, we have shown that memory formation can be modulated by controlling the phase differences between rhinal cortex and hippocampus via deep brain stimulation (Fell et al., 2013). However, the same frequency (40 Hz) and phase differences (zero vs. 180 degree) were chosen for all subjects in this study. The present findings suggest that memory enhancement and inhibition effects may be augmented by classification analyses enabling the individual selection of optimal stimulation frequencies and phase differences.

Acknowledgements

This study was supported by the German Research Foundation via the collaborative research center SFB 1089. Nikolai Axmacher received additional support via Emmy Noether grant AX82/2, DFG project AX82/3, and SFB 874. We would like to thank Eva Ludowig for her contributions to data acquisition and preprocessing and Markus Neuhäuser for helpful comments regarding statistical analyses.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2016.06.021.

References

Anastassiou, C.A., Montgomery, S.M., Barahona, M., Buzsáki, G., Koch, C., 2010. The effect of spatially inhomogeneous extracellular electric fields on neurons. J. Neurosci. 30 (5), 1925–1936. http://dx.doi.org/10.1523/JNEUROSCI.3635-09.2010.

Berens, P., 2009. CircStat: a Matlab toolbox for circular statistics. J. Stat. Softw. 31(10), 1–21. doi: 10.18637/jss.v031.i10

Busch, N.A., Dubois, J., VanRullen, R., 2009. The phase of ongoing EEG oscillations predicts visual perception. J. Neurosci. 29 (24), 7869–7876. http://dx.doi.org/10.1523/JNEUROSCI.0113-09.2009.

Cohen, N., Pell, L., Edelson, M.G., Ben-Yakov, A., Pine, A., Dudai, Y., 2015. Peri-encoding predictors of memory encoding and consolidation. Neurosci. Biobehav. Rev. 50, 128–142. http://dx.doi.org/10.1016/j.neubiorev.2014.11.002.

Duvernoy, H.M., 1988. The Human Hippocampus. An Atlas of Applied Anatomy. J. F. Bergmann Verlag, München, pp. 25–43.

Elbert, T., Rockstroh, B., 1987. Threshold regulation - a key to the understanding of the combined dynamics of EEG and event-related potentials. J. Psychophysiol. 4, 317–333.

Fell, J., Axmacher, N., 2011. The role of phase synchronization in memory processes. Nat. Rev. Neurosci. 12 (2), 105–118. http://dx.doi.org/10.1038/nrn2979.

Fell, J., Klaver, P., Lehnertz, K., Grunwald, T., Schaller, C., Elger, C.E., Fernández, G., 2001. Human memory formation is accompanied by rhinal–hippocampal coupling and decoupling. Nat. Neurosci. 4 (12), 1259–1264.

Fell, J., Ludowig, E., Rosburg, T., Axmacher, N., Elger, C.E., 2008. Phase-locking within human mediotemporal lobe predicts memory formation. NeuroImage 43 (2), 410–419. http://dx.doi.org/10.1016/j.neuroimage.2008.07.021.

Fell, J., Ludowig, E., Staresina, B.P., Wagner, T., Kranz, T., Elger, C.E., Axmacher, N., 2011. Medial temporal theta/alpha power enhancement precedes successful memory encoding: evidence based on intracranial EEG. J. Neurosci. 31 (14), 5392–5397. http://dx.doi.org/10.1523/JNEUROSCI.3668-10.2011.

Fell, J., Staresina, B.P., Do Lam, A.T., Widman, G., Helmstaedter, C., Elger, C.E., Axmacher, N., 2013. Memory modulation by weak synchronous deep brain stimulation: a pilot study. Brain Stimul. 6 (3), 270–273. http://dx.doi.org/10.1016/j.brs.2012.08.001.

Fernández, G., Effern, A., Grunwald, T., Pezer, N., Lehnertz, K., Dümpelmann, M., Van Roost, D., Elger, C.E., 1999. Real-time tracking of memory formation in the human rhinal cortex and hippocampus. Science 285 (5433), 1582–1585.

Fries, P., 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn. Sci. 9 (10), 474–480.

Fröhlich, F., McCormick, D.A., 2010. Endogenous electric fields may guide neocortical network activity. Neuron 67 (1), 129–143. http://dx.doi.org/10.1016/j.neuron.2010.06.005.

Grunwald, T., Beck, H., Lehnertz, K., Blümcke, I., Pezer, N., Kurthen, M., Fernández, G., Van Roost, D., Heinze, H.J., Kutas, M., Elger, C.E., 1999. Evidence relating human verbal memory to hippocampal *N*-methyl-D-aspartate receptors. Proc. Natl. Acad. Sci. U. S. A. 96 (21), 12085–12089.

Guderian, S., Schott, B.H., Richardson-Klavehn, A., Düzel, E., 2009. Medial temporal theta state before an event predicts episodic encoding success in humans. Proc. Natl. Acad. Sci. U. S. A. 106 (13), 5365–5370. http://dx.doi.org/10.1073/pnas.0900289106.

Haque, R.U., Wittig, J.H., Damera, S.R., Inati, S.K., Zaghloul, K.A., 2015. Cortical low-frequency power and progressive phase synchrony precede successful memory encoding. J. Neurosci. 35 (40), 13577–13586. http://dx.doi.org/10.1523/JNEUROSCI.0687-15.2015.

Helfrich, R.F., Knepper, H., Nolte, G., Strüber, D., Rach, S., Herrman, C.S., Schneider, T.R., Engel, A.K., 2014. Selective modulation of Interhemispheric functional connectivity by HD-tACS shapes perception. PLoS Biol. 12 (12), e1002031. http://dx.doi.org/10.1371/journal.pbio.1002031.

Helmstaedter, C., Fritz, N.E., González Pérez, P.A., Elger, C.E., Weber, B., 2006. Shift-back of right to left hemisphere language dominance after control of epileptic seizures: evidence for epilepsy driven functional cerebral organization. Epilepsy Res. 70 (2–3), 257–262.

Huerta, P.T., Lisman, J.E., 1993. Heightened synaptic plasticity of hippocampal CA1 neurons during a cholinergically induced rhythmic state. Nature 364 (6439), 723–725.

Lee, H., Fell, J., Axmacher, N., 2013. Electrical engram: how deep brain stimulation affects memory. Trends Cogn. Sci. 17 (11), 574–584. http://dx.doi.org/10.1016/j.tics.2013.09.002.

Lopour, B.A., Tavassoli, A., Fried, I., Ringach, D.L., 2013. Coding of information in the alpha phase of local field potentials within human medial temporal lobe. Neuron 79 (3), 594–606. http://dx.doi.org/10.1016/j.neuron.2013.06.001.

Mathewson, K.E., Gratton, G., Fabiani, M., Beck, D.M., Ro, T., 2009. To see or not to see: prestimulus alpha phase predicts visual awareness. J. Neurosci. 29 (9), 2725–2732. http://dx.doi.org/10.1523/JNEUROSCI.3963-08.2009.

Mormann, F., Fell, J., Axmacher, N., Weber, B., Lehnertz, K., Elger, C.E., Fernández, G., 2005. Phase/amplitude reset and theta–gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. Hippocampus 15 (7), 890–900.

Neuhäuser, M., 2011. Nonparametric Statistical Tests: A Computational Approach. CRC Press, Boca Raton, pp. 167–170.

Neuling, T., Rach, S., Wagner, S., Wolters, C.H., Herrmann, C.S., 2012. Good vibrations: oscillatory phase shapes perception. NeuroImage 63 (2), 771–778. http://dx.doi.org/10.1016/j.neuroimage.2012.07.024.

Noh, E., Herzmann, G., Curran, T., de Sac, V.R., 2014. Using single-trial EEG to predict and analyze subsequent memory. NeuroImage 84, 712–723. http://dx.doi.org/10.1016/j.neuroimage.2013.09.028.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869. http://dx.doi.org/10.1155/2011/156869.

Pavlides, C., Greenstein, Y.J., Grudman, M., Winson, J., 1988. Long-term potentiation in the dentate gyrus is induced preferentially on the positive phase of theta-rhythm. Brain Res. 439 (1–2), 383–387.

Reardon, S., 2015. Memory-boosting devices tested in humans. Nature 527 (7576), 15–16. http://dx.doi.org/10.1038/527015a.

Rizzuto, D.S., Madsen, J.R., Bromfield, E.B., Schulze-Bonhage, A., Seeling, D., Aschenbrenner-Scheibe, R., Kahana, M.J., 2003. Reset of human neocortical oscillations during a working memory task. Proc. Natl. Acad. Sci. U. S. A. 100 (13), 7931–7936.

Suthana, N., Fried, I., 2014. Deep brain stimulation for enhancement of learning and memory. NeuroImage 85 (3), 996–1002. http://dx.doi.org/10.1016/j.neuroimage.2013.07.066.